An Autosegmental Approach to Problems in Reconstructing Old Chinese

Joshua Anderson
Senior Honors Thesis for
April 26, 2004

Introduction

The aim of this thesis[*] is to apply the insights of autosegmental phonology to problems in the reconstruction of Old Chinese. The first problem examined is the licensing of Old Chinese affixes, as proposed by Laurent Sagart. The merits of a prenasalizing prefix, as Sagart proposes, are compared with a hypothesis based solely on voicing. The licensing of Sagart's *-ŋ suffix is also compared with that of his other suffixes. The next problem considered is how the underspecification of vowels can be used to gain insight into the necessity of the rounding dissimilation rule William Baxter proposes for the few words he reconstructs with labialised velar initials and round vowels.

The final question examined is the origin of Middle Chinese palatalizing finals. This paper proposes an autosegmental counterpart of Baxter's proposal that projects a palatal approximant segment back to Old Chinese. Jerry Norman's proposal about the pharyngealization of initials in non-palatalized reflexes is then compared with this hypothesis. A synthesis of these two hypotheses and Sagart's Old Chinese morphology is attempted to explain odd fǎnqiè spellings found in the *Qièyùn*.

Since most readers will not be familiar with the theories and reconstructions presented by Baxter, Sagart, and Norman, these are each introduced in the text. The first

[*] I would like to thank Sean Crist, professor of Linguistics at Swarthmore College, for his invaluable help in preparing this text and for his indefatigable spirit in reading all of the drafts presented to him, Shizhe Huang, professor of Chinese and Linguistics at Haverford College, for her fresh eyes and insightful comments as the draft of this text approached completion, Ted Fernald, professor of Linguistics at Swarthmore College, for his comments on the very first draft of the text, for guiding me throughout the process of writing a thesis, and for his observations after the various presentations developing the analysis present here, Lillie Dremeaux for her Herculean effort in attempting to make my unintelligible first draft readable, Sarah Tan for her comments at a busy time in the semester, and finally, my friends Kathleen Kristian, Meredith Leigh, and LiErin Probasco for supporting me and always being willing to listen to my tales of woe about completing this text. The responsibility for any errors to be found in this text falls squarely on the shoulders of the author.

section defines Old Chinese and explains what sources are used to reconstruct it. The second introduces Baxter and Sagart's reconstructions and points out problems that will be examined later. The third section analyzes the underspecification of Old Chinese vowels and the licensing of affixes. The fourth section introduces problems in analyzing the origin of Middle Chinese palatalizing finals and then presents the analyses sketched above.

1. Methodology Used to Reconstruct Old Chinese

1.1 Defining Old Chinese

Old Chinese describes the oldest reconstructible language spoken in China for which there are written records. In its broadest sense, the term describes the language spoken during the 周 Zhōu[1] dynasty, 11th century to 211 B.C. There is also a narrower sense in which Old Chinese is used, to label the language spoken during the first part of the Zhōu dynasty, 西周 the Western Zhōu period, 11th century to 771 B.C. (Sagart, 1999: 5). The language of this period is believed to be ancestral to all of the modern Chinese dialects, to Middle Chinese, and to the language captured in later writings of the Zhōu dynasty (Baxter, 1992: 24-25).[2]

1.2 Written Records in Early China

There are two genres of writings that come from the Western Zhōu and early 春秋 Spring and Autumn (770 – 476 B.C.) periods, the inscriptions on bronze vessels and the 詩經 *Shījīng*. The first bronze inscriptions are short prose passages commemorating an honor bestowed upon a family. Later inscriptions are much longer and contain some

[1] All romanizations represent the Modern Mandarin Chinese pronunciations of the characters according to the Hànyǔ Pīnyīn system.
[2] Table 1 at the end of the text charts the periodization of Chinese.

rhymed sections (Baxter, 1992: 345). The *Shījīng* (often translated as *The Book of Odes* or *The Classics of Poetry*) is a collection of 305 poems written between 1000 and 600 B.C. (Baxter, 1992: 355).

The Chinese did not begin writing in the Zhōu dynasty; the oracle bones of the 商 Shāng dynasty (15th century to 11th century B.C.) represent the oldest Chinese corpus (Sagart, 1999: 7). These texts record the question a Shāng king asked, the divination response, and, occasionally, the outcome of the event. While priceless from a historical perspective, the oracle bones themselves contain no rhymed passages and offer little evidence about the phonology of the language of the Shāng rulers (Baxter, 1992: 344-345). There are parallels in the grammar and lexicon of Shāng and Zhōu Chinese. To a certain extent, the phonetic elements of the scripts bear certain similarities (Sagart, 1999: 7), but our understanding of this phenomenon is in terms of the phonology reconstructed for Zhōu Chinese.

1.3 Evidence for Reconstructing Old Chinese

Historical reconstructions of Old Chinese are primarily based on three types of evidence. The first of these is the rhymes in the *Shījīng* and, to a lesser extent, those of the bronze inscriptions. The second type of evidence consists of the characters themselves. Certain Chinese characters, called 諧聲 xiéshēng or phonetic compound characters, are composed of two radicals, one of which serves as a semantic marker and the other as a phonetic indicator. Finally, reconstructions are also based on the distinctions attested in the Middle Chinese rhyme books, the most important of which is the 切韻 *Qièyùn*, completed in A.D. 601, and rhyme tables (Baxter, 1992: 33). Since the language captured in the *Shījīng* may not have evolved directly into the language of the

Qièyùn, Old Chinese is defined, in its narrowest sense, as the language from which all attested forms of Chinese derive. Each of these types of evidence is discussed more fully below.

1.3.1 Rhyming Evidence

As noted above, the corpus of Zhōu writings contains many rhymed texts, namely the poems of the *Shījīng* and some passages in the bronze inscriptions. The poems of the *Shījīng* are quite varied in content, describing feasts, marriages, love, hunting, being away at war, drinking bouts, and everyday life (Karlgren, 1949: 102-103). Since some of the poetry has rhymed in all the historical stages of Chinese (Baxter, 1992: 150), it is not much of a leap to imagine that all of the poetry rhymed when it was written. In fact, early scholars such as 陸德明 Lù Démíng (circa A.D. 550 – 630) simply believed that the lines which did not rhyme simply showed that the poets of *Shījīng* used loose rhyming standards (Baxter, 1992: 40, 153).

The *Shījīng* is divided into six sections: the 國風 *Guófēng*, 小雅 *Xiǎoyǎ*, 大雅 *Dàyǎ*, 周頌 *Zhōusòng*, 魯頌 *Lǔsòng*, and商頌 *Shāngsòng*. The *Guófēng* is largest and most heterogeneous of the sections. The translation of this section title is *Airs of the States*; the poems of this section are believed to have been written by authors from many different places of the Zhōu kingdom and from different times. The poems of the *Xiǎoyǎ* may have been written by members of the Zhōu court and are generally political in nature. The *Dàyǎ* contains dynastic hymns from the Western Zhōu period, including legendary histories of early Zhōu ancestors. The *Zhōusòng* is composed of hymns praising Heaven and the Zhōu ancestors, several of which are unrhymed. The *Lǔsòng* and *Shāngsòng* sections were the last to be written (during the 7th century B.C.). The names of these

sections, like that of the *Zhōusòng*, are related to states from the Spring and Autumn period, 魯 Lǔ and 宋 Sòng (whose rulers were descendants of the Shāng kings) (Baxter, 1992: 355-356).

According to tradition, the cannon of the *Shījīng* was chosen by Confucius, who selected it from three thousand poems written during the Zhōu dynasty (Rhew)[3]. Even though versions of the *Shījīng* were written during the Zhōu dynasty, these manuscripts did not survive the book burnings of 秦始皇 Qín Shǐhuáng, the first emperor of the 秦 Qín dynasty (221 - 207 B.C.) and the first emperor of united China. The *Shījīng* was also passed on orally, and later written down with commentary on difficult passages and glosses for obscure characters. During the 漢 Hàn dynasty (206 B.C. - A.D. 220), there were four different commentaries and glosses of the *Shījīng*. Three of these versions have survived to the present only as fragments and in quotations in other texts. The 毛 Máo version of the *Shījīng* and the accompanying commentary is the school of interpretation on which the present version of the *Shījīng* is based (Baxter, 1992: 357).

Based on the divergences of the extant fragments of the other three schools from the Máo version of the *Shījīng*, many scholars have reached the conclusion that the *Shījīng* was not well understood even as early as the Hàn dynasty. These scholars point to the vast number of glosses and copious commentary provided to explain passages of the text (Baxter, 1992: 358). Comparing the situation to American children's idiosyncratic understandings of the Pledge of Allegiance, Baxter asserts that if the people memorizing the poems of the *Shījīng* had a good understanding of what the poems meant, their understanding would have acted as a self-correction mechanism. In the absence of

[3] This can be found on-line at http://academic.reed.edu/chinese/chin-hum/materials/shijing/shijing.html.

this understanding, the text changed because the errors in transmission were not corrected. The commentary then attempted to explain and legitimize these errors (Baxter, 1992: 360).

The texts of the bronze inscriptions do not present the problems of textual transmission that the poems of the *Shījīng* do. Their origins in space and time can also be determined quite precisely. However, there are relatively few rhymed passages in the small corpus. These passages are also mixed together with prose passages on the same vessel, making their boundaries difficult to determine. In addition, the rhymes used in these passages tend to use words that are firmly reconstructible using other sources, such as the *Shījīng* (Baxter, 1992: 345).

1.3.2 Evidence from Xiéshēng characters

Most Chinese characters are xiéshēng characters, being composed of a phonetic and semantic component (Yao et al., 1997: 15). According to Baxter's formulation of the requirements for the use of phonetic elements, a new character can take a certain phonetic element if its rime is the same as the phonetic element and the initial consonant has the same place of articulation. Also, differences in the nasality of the initial consonants generally prevent the phonetic element from being used. Some other differences in the manner of articulation of the initial segment do not prevent using the phonetic element (Baxter, 1992: 348). If characters meet these criteria, then they can be written with the same phonetic element and form a xiéshēng series, some of which still share similar pronunciations. Two xiéshēng series are given below:

1) a) 里 lǐ 'a measure of distance (about ⅓ mile)'
   b) 哩 lī 'to speak indistinctly'
   c) 狸 lí 'fox'
   d) 俚 lǐ 'rustic'

e) 悝 lǐ ‘sad, worried’
f) 理 lǐ ‘reason, logic’
g) 裡 lǐ ‘within, inside’
h) 鯉 lǐ ‘carp’

2) a) 乍 zhà ‘suddenly’
b) 炸 zhá ‘to fry,’ zhà ‘to explode’
c) 窄 zhǎi ‘narrow’
d) 作 zuò ‘to do’
e) 昨 zuó ‘yesterday’
f) 怎 zěn ‘why, how’

This area of study has fascinated many scholars, and there exists a great volume of literature explaining the form of these types of characters. The first of these is the 說文解字 *Shuōwén jiězì* ‘Clarifying texts and explaining characters,’ a Hàn dynasty dictionary written in A.D. 100 by 許慎 Xǔ Shèn. For each xiéshēng character in the dictionary, the author explains which radical is the semantic element and which is the phonetic.

A famous dictum in historical Chinese phonology is 同聲必同部 (Words with the same phonetic element must belong to the same rhyme group.) (Baxter, 1992: 141). This principle is applied when reconstructing Old Chinese, though not indiscriminately. Consider the following reconstructions (Sagart, 1999: 128)

3) 各 Baxter *kak, Sagart *^{A}k-lak > kak > gè ‘each, every’[4]
4) 骼 Baxter *krak, Sagart *^{A}kr-lak > kæk > gé ‘hind-leg’
5) 落 Baxter *C-rak, Sagart *^{A}kə-lak > lak > luò ‘to fall’
6) 貉 Baxter *gak, Sagart *^{A}Nk-lak > hak > hé ‘badger, marmot’

[4] These data show the development of Old Chinese into Middle Chinese and then into Modern Mandarin Chinese. In general, any word preceded by an asterisk is the Old Chinese reconstruction; the Middle Chinese and Modern Mandarin forms follow. Reconstructed or attested words from other periods will be marked accordingly. Whether the reconstruction is Baxter’s or Sagart’s should be clear from context and both are used without special reference in the text, though all of Sagart’s reconstructions are preceded by a superscript A or B, whereas Baxter’s are not. The superscripts in Sagart’s reconstructions correspond to the presence or absence of a palatal initial, j, or i in the Middle Chinese reflexes. This issue is introduced in section 1.3.3 and is explored more fully in section 4 of the text.

The characters above all have xiéshēng contact; they share the same phonetic element, 各. Both Baxter and Sagart's reconstructions are similar to each other and all have the rime *ak.

However, xiéshēng characters can be created at any time. Such characters reflect only the phonology of the time at which they were created. After the Communist government came to power, 毛澤東 Máo Zédōng (simplified to 毛泽东) initiated a reformation of the Chinese script to increase literacy. In general, the simplified characters used in the People's Republic of China today make the characters easier to write by reducing the number of strokes required to write them, replacing obscure phonetic elements with more common ones, and conflating rarely used characters with similar ones written almost identically. The simplification also replaced the phonetic elements of a limited number of characters with other phonetic elements that reflected similarities in the Modern Mandarin pronunciations, as Table 1.1 shows.

During 戰國 the Warring States (475 – 221 B.C.) period, each state established its own writing system. After the unification of China under Qín Shǐhuáng, the Chinese script was reformed to reflect the writing of the conquering state of Qín. During this period of disruption and the subsequent evolution of the Qín characters in the Hàn dynasty, many of the original phonetic elements of the Zhōu characters were replaced by phonetic elements that were more appropriate based on the phonology of the time. It is, in fact, the Hàn dynasty script that was used to write the text of the *Shījīng* as it has been passed down to the present. This fact makes the script of the bronze inscriptions even more valuable for reconstructing Old Chinese, as the characters contain the original

Table 1.1

Simplified Character	Traditional Character	Pronunciation	Gloss[5]	Simplified Phonetic Element	Pronunciation	Traditional Phonetic Element	Pronunciation
迟	遲	chí	late, slow; to delay	尺	chǐ	犀	xī
态	態	tài	attitude, manner	太	tài	能[6]	néng
犹	猶	yóu	like, similar to	尤	yóu	酋	qíu
战	戰	zhàn	war; to fight	占	zhàn	單	dān
证[7]	證	zhèng	evidence, proof	正	zhèng	登	dēng
钟	鐘	zhōng	bell; clock	中	zhōng	童	tóng

phonetic elements that reflect the phonology of Zhōu dynasty Chinese (Baxter, 1992: 345-347).

Not only did new phonetic elements replace older ones that did not match the phonology of Hàn dynasty Chinese, but older characters may have also justified the use of phonetic elements that were not really similar enough to fit Baxter's parameters of xiéshēng contact. For example, the following pair of characters from the Old Chinese period

7) 殘 *dzan > dzan > cán 'to hurt'
8) 踐 * dzjanʔ > dzjenX > jiàn 'to tread, to trample'

[5] The glosses are taken from Harbaugh, Rick. 1998. *Chinese Characters: A Genealogy and Dictionary*. Taipei: Zhongwen.com.

[6] 態 *hnis > thojH > tài was not considered to be the phonetic element of 能 *nɨ > nong > néng very early. The *Shuōwén jiězì* explains 態 as 意也从心从能徐鍇曰心能其事然後有態度也 '態 is intentions. It is from 心 [heart] and from 能 [to be able to]. Xú Kǎi says that the heart is able to [show] its affairs, then there certainly is attitude.' This is in sharp contrast to the brief explanations of xiéshēng characters, such as 忽 hū 'to forget; to disregard, to neglect.' 忽 is explained as 忘也从心勿聲 '忽 is to forget. It is from 心 [heart]; the phonetic element is 勿 wù' Xǔ Shèn 許慎. 1963. *Shuōwén jiězì* <說文解字>. Beijing: Zhōng huá shū jú 中華書局. p. 220. Baxter treats the Middle Chinese reading of 能 as unexpected. Baxter, William Hubbard. 1992. *A Handbook of Old Chinese Phonology*: Trends in Linguistics. Studies and Monographs: 64. New York: Mouton de Gruyter. pp. 192-193.

[7] The simplification of the semantic elements of the last two characters represents the regular simplification of these radicals according to the 简化字总表 Jiǎn huà zì zǒng biǎo 'Complete Table of Simplified Characters.'

may have justified the use of 9) as the phonetic element of 10) in the Middle Chinese period

9) 散 *sanʔ > sanX > sǎn ‘scattered’
10) 霰 *s(k)ens > senH > xiàn ‘sleet’

Originally, Baxter asserts, the character for 10) was written as 雨見 with the phonetic 11)

11) 見 *kens > kenH > jiàn ‘to see’ (Baxter, 1992: 354)

The original character does fit the principle of xiéshēng contact since 10) and 11) have the same rime.[8] Thus, the reformation of the Chinese writing system during the Qín and Hàn dynasties presents a major problem for the reconstruction of Old Chinese. Even though there are serious questions about the extent to which xiéshēng characters systematically reflect Chinese phonology (Branner: 1)[9], the phonetic components of xiéshēng characters are the only information from the period about the initial of Old Chinese (Sagart, 1999: 5).

1.3.3 Evidence from Middle Chinese rhyme books and rhyme tables

The final piece of information used to reconstruct Old Chinese are the Middle Chinese rhyme books and rhyme tables. The most important of the rhyme books is the *Qièyùn*. Though a complete version of the *Qièyùn* itself has not survived intact to the present, other rhyme books based on the *Qièyùn*, such as the 廣韻 *Guǎngyùn* (A.D. 1007 - 1008) and the 集韻 *Jíyùn* (A.D. 1038 - 1039) are extant. According to the *Qièyùn* tradition, characters are first grouped according to their tone. There were four tones in

[8] If 10) and 11) are part of the same xiéshēng series, 10) must be reconstructed with the k in parentheses so that the initials have the same place of articulation. In the reconstruction of 10), *s is a pre-initial (see section 2.2.1) and is not considered in the choice of phonetic element.
[9] Page numbers for Branner are taken from a pdf of the article that appeared in the book listed in the bibliography, not from the book itself.

Middle Chinese: 平聲 píngshēng ‘level tone,’ 上聲 shǎngshēng ‘rising tone,’ 去聲 qùshēng ‘departing tone,’ and 入聲 rùshēng ‘entering tone.’ The phonetic quality of the tones is not known, but these words were likely chosen because they represent the various tones. Rùshēng words were certainly characterized by oral stops in the coda (Baxter, 1992: 32), as they are in conservative dialect families such as 粵 Yuè (which includes Cantonese) (Bauer and Benedict, 1997: 154). Words with final nasals, approximant codas, or no coda consonant could belong to any of the first three tones. It is believed that these three labels may have been chosen to reflect the pitch contour (level, rising, or falling) of the words with that tone. Words with the same tone were then separated according to their rhymes. Finally each rhyme group is split into subdivisions of homophonous words (Baxter, 1992: 32-35). Each group of homophones, 小韻 xiǎoyùn ‘small rhyme’, in a given rhyme group is taken to be contrastive. There was no standard arrangement of xiǎoyùn across rhyme groups; their arrangement in a single work is “haphazard” (Branner: 8).

The rhyme groups did not always correspond to the words’ rime, in its linguistic definition of the syllable nucleus and coda (Branner, p.c., 2004). In general, the words had the same final, a term used to describe everything in the syllable that is not the syllable-initial segment, which is simply called the initial. The final includes the nucleus, the coda, any onset consonants between the syllable-initial segment and the nuclear vowel, as well as the syllable’s tone. Words could have different finals and still rhyme and possibly be assigned to the same rhyme group. The 東 Middle Chinese tuwng > dōng rhyme group is reconstructed with both the finals *uwng* and *juwng*. Other rhyme groups split words with different finals, such as the 陽 Middle Chinese yang > yáng

rhyme group, which contains *jang* and *jwang*, and the 唐 Middle Chinese dang > táng rhyme group, which is composed of words ending in *ang* and *wang*. These two rhyme groups are placed next to each other in the *Qièyùn*, an indication of their similarity. In fact, all four finals rhymed freely in Middle Chinese poetry (Baxter, 1992: 34) The division of finals into rhyme categories is not systematic across rhyme books; some may take all or some of the medials into account when dividing finals into rhyme groups, others may not take them into account at all (Branner, p.c., 2004)[10].

The Middle Chinese rhyme books primarily functioned as pronunciation dictionaries. Later books in this tradition became more like dictionaries, providing information not only about pronunciation, but also meanings and variant written forms of the character. In creating a pronunciation dictionary for Chinese characters, the authors of these works employed a spelling system using Chinese characters to indicate the pronunciation of homophonous characters in every homophonous group of the *Qièyùn*. The first character in the xiǎoyùn would be spelled in terms of two well-known characters. The first represented the initial (the syllable-initial segment) of the homophonous words; the second represented the final in addition to the tone. This system is called 反切 fǎnqiè. In the absence of an alphabetic system, the fǎnqiè system, along with providing common homophonous characters, was a standard way of representing a character's pronunciation (Baxter, 1992: 33).

Fǎnqiè, most importantly, was still not a standard alphabetic system or syllabary. It is only relative, creating equivalence classes of the initials and finals. The compilers of

[10] David Prager Branner is a professor in the Department of Asian and East European Languages and Cultures at the University of Maryland's School of Languages. He is also the coordinator of the Chinese Program there.

the rhyme books chose any of the common characters that began with that initial, so one word had many fǎnqiè spellings in the different rhyme books. Even in one redaction of the *Qièyùn*, the same sounds are spelled with different characters. Even stranger, there are also instances in which one fǎnqiè spelling is used in one work to spell words with different pronunciations. This has led to disagreements about the number of initials and finals represented in the *Qièyùn*. This disagreement is possible because there are no guidelines for the pronunciation of any of the fǎnqiè spellers in any of the texts, so the readers must have relied on their own knowledge of literary accents to determine the pronunciation (Branner: 4-8).

In the preface to the original *Qièyùn*, 陸法言 Lù Fǎyán, its compiler, wrote about the importance of a good literary accent for both reading poetry and writing while bemoaning the fact that accents and rhymes across time and place vary (Branner: 9-10). It seems likely that one goal of the authors of the *Qièyùn* was to establish a standard for pronunciation. In accordance with this perscriptivist goal, the language of the *Qièyùn* is not taken to represent the speech of any particular place at any particular time. The authors made the finest distinctions possible when deciding which characters to use as fǎnqiè spellers (Baxter, 1992: 37). By drawing on other older works that have not survived to the present (Branner: 11) and the dialects of people from several different areas, the authors drew distinctions that were not representative of the speech of their dialects or even their own time (Baxter, 1992: 37). The *Qièyùn* is not believed to accurately represent common Middle Chinese spoken around A.D. 600, but, more likely, a combination of diverse prestigious traditions of reading literature aloud (Branner: 14). In any event, the distinctions are believed to have been important at some point in historical

Chinese and, thus, are used in reconstructing the origins of various Chinese dialects and ultimately Old Chinese (Baxter, 1992: 37).

The rhyme table tradition, as represented in the 韻鏡 *Yùnjìng* (A.D. 1161 and A.D. 1205) and the 七音略 *Qīyīn lüè* (circa A.D. 1150), developed later and more closely reflects Middle Chinese phonology. Every word in the rhyme table is found at the intersection of what generally corresponds to its phonetic rime and its onset (Baxter, 1992: 42). More specifically, the rhyme tables place characters at the intersection of their initial and final, as defined above. This provides a more detailed analysis of the phonology in the rhyme tables than does the fǎnqiè of the rhyme book tradition by providing a standard set of initials and finals and by describing the initials in phonetic terms generally analogous to the Sanskrit system in the prefaces to the rhyme tables.

However, scholars increasingly question the extent to which the rhyme tables, especially early rhyme tables, are descriptions of some form of Middle Chinese. The rhyme tables are increasingly viewed as a rationalization of the diverse fǎnqiè found in the rhyme book tradition. For instance, even though the rhyme books differ from each other in their treatment of the finals, the tradition in the rhyme tables is to sort finals containing a medial *w* from those that do not; this is the 合口/開口 hékǒu/kāikǒu distinction in the rhyme tables. Since these finals were obviously related, the hékǒu and kāikǒu tables were placed next to each other (Baxter, 1992: 42-43). Also, as we will see below, the rhyme tables may have also separated finals beginning with *j* from those that do not have *j*. To this end, they are believed to represent a formal system, possibly promulgating a standard literary accent (Branner, p.c., 2004).

There is a further division of Middle Chinese finals into four different divisions (等 děng) in the rhyme tables. What the word means děng means in terms of Chinese phonology is controversial. One traditional view is that it is related to the vowel quality, specifically vowel height (Branner, p.c., 2004), a position to which Baxter generally subscribes. Division I words have back vowels as nuclei and all coronal initials are dental.[11] Division II words have one of the two lax vowels [æ] and [ɛ]. Also, the coronal initials are all retroflex. Division III words are characterized by at least one of the following: [i] as the nucleus, [j] as a medial, and palatal coronal initials. Division IV words have the vowel [e], the only other front vowel in Middle Chinese, and dental coronal initials.

Some finals with [i] or [j] are also found in Divisions II and IV. The division II words have retroflex initials followed by [i] or [j]. The Division IV words may have initials with any place of articulation. The coronal initials of such Division IV words are all dental.[12] The reasons for placing words with labial or velar initial into Division IV are not clearly understood. It may represent an archaism preserved in the rhyme table tradition (Baxter, 1992: 65-75). Also, some finals begin with [j], but also contain the Division II vowel [æ]; they are *jæ*, *jæm*, and *jæng*. The final *jæ* is the only one that occurs in Division II of the rhyme tables and only with retroflex initials. The final *jæm* is placed only in Division III of the rhyme table and never in Division II. Likewise, *jæng* is found only in Divisions III and IV, never in Division II (Baxter, 1992: 72, 74, 80).

---

[11] Tables 2-4 at the end of the text show words from divisions I, II, and IV.

[12] Tables 5-7 show words with [i] or medial [j] paired with all the possible initials. Since they are sorted by place of articulation, it will be easy for the reader to divide the words with coronal initials into divisions. Since the placement of words with labial and velar initials is not well understood and has not been directly incorporated into the Old Chinese reconstructions, these words have not been labeled for which division they are assigned to in the *Yùnjìng*.

It is natural to assume that all of the finals that have [i], [j], or a palatal initial should be analyzed as one group, even though such words may be listed in Division II, III, or IV. Hence, it is useful to distinguish them, regardless of their divisional classification in the Middle Chinese rhyme tables, from the finals that do not have any of those characteristics. Though the term Division III finals has been used to refer to all such finals since most are placed in Division III of the rhyme tables, it is clearer to use another term. Pulleyblank refers to these as type B finals and to Middle Chinese finals without [i], [j], or palatal initials as type A finals (1977-1978: 183). Sagart projects this distinction back to Old Chinese by marking each word he reconstructs with the superscript letter corresponding to the classification of the Middle Chinese reflex.

More recent studies of the rhyme tables place much less emphasis on the divisions. Some linguists believe that a four-way distinction in one small class of words was expanded to the entire lexicon. Most linguists studying the rhyme tables are more interested in the set of initials that can precede any given final (Branner, p.c., 2004). Baxter, in some ways, considers this, but, for his analysis, the crucial distinction of the divisions still remains the vowel.

Even though, there is a tradition of phonological study in Chinese, Middle Chinese is still reconstructed. The Chinese philologists never divided the final, with all its information, into smaller analyzable units. The reconstruction of Middle Chinese is certainly not agreed upon by scholars in the field, especially as seen in the varied interpretations of the four děng (Branner, p.c., 2004). This is, of course, a major stumbling block for a credible reconstruction of Old Chinese. In fact, there are major concerns about all of the evidence on which Old Chinese reconstructions are based. Even

given this deficiency, I believe it is still useful to examine these reconstructions of Old Chinese, bearing in mind that the foundations on which they are based are by no means firm.

2. Previous Reconstructions

2.1 The Beginnings of Chinese Historical Phonology: Qīng Dynasty Studies

Chinese scholars were traditionally confused by the rhyming practices of the *Shījīng*. The first theories developed held that the ancient poets were less concerned about phonological similarity in their rhymes. Later scholars, particularly in the 宋 Sòng dynasty (A.D. 960 - 1279), believed that the characters were read with different pronunciations to make them rhyme. Since a character was given several readings, many specific to only one poem, later scholars found the arbitrariness of the theory's application unsatisfying. The 明 Míng dynasty (A.D. 1368 - 1644) scholar 陳第 Chén Dì (1541 - 1617) is credited with being the first person to recognize that historical and geographic sound changes affected the rhyming pattern of the *Shījīng*. Chén Dì also recognized that only one pronunciation was necessary to account for Old Chinese rhyming practices (Baxter, 1992: 152-155).

Systematic study of the rhyming patterns of the *Shījīng* flourished during the 清 Qīng dynasty (A.D. 1644 - 1911). As part of the imperial exams, scholars simply memorized the corpus of classical texts, including the *Shījīng*. Based on this immense knowledge and familiarity with the texts, scholars began dividing the rhyme words of the *Shījīng* into rhyming categories (Baxter, 1992: 139). Two major figures in the history of this scholarship are 段玉裁 Duàn Yùcái (1735 - 1815) and 王念孫 Wáng Niànsūn (1744 - 1832). Duàn Yùcái was the first person to sort the words of the *Shuōwén jiězì* into Old

Chinese rhyme groups, whether or not the characters rhymed in the *Shījīng*. He was the scholar who first formalized that characters with the same phonetic element must belong to the same rhyme group (Baxter, 1992: 160-161).

Wáng Niànsūn divided Old Chinese words into twenty-two different rhyme groups. His work is essentially the starting point for modern linguists reconstructing Old Chinese (Baxter, 1992: 167-169). Incorporating the ideas of 戴震 Dài Zhèn (1724 - 1777), who separated oral and nasal final stops (Baxter, 1992: 162), many linguists have not questioned the accuracy of the rhyme groups proposed by these Qīng scholars. Bernhard Karlgren, Edwin Pulleyblank, and Li Fang-kuei are just some of the linguists who reconstruct only one rime for each of the traditional rhyme groups.

2.2 Baxter's Analysis

Baxter, while appreciating the traditional analysis of Old Chinese, does not follow the work of the Qīng scholars as closely as previous reconstruction do. Baxter believes that the Qīng scholars' analysis, even at its best, predicts rhymes that do not actually occur. He believes that the thirty-one rhyme groups used by linguists like Li Fang-kuei still overlook certain distinctions that existed in Old Chinese (Baxter, 1992: 140). Baxter's text ends with tables that explain which rhymes were conflated in the traditional analysis.

Since proof that two groups of words do not rhyme with each other cannot be determined simply by examining the text of the *Shījīng*, Baxter does a statistical analysis of rhyming patterns for his rhyme groups. In fact, words that must be reconstructed with one vowel according to his hypotheses, statistically speaking, do not rhyme with words that must be reconstructed with a different vowel, even though the words are part of the

same traditional rhyme group. His statistical analysis confirms that most of his proposed rhyming distinctions are in fact statistically significant (Baxter, 1992: 369).

2.2.1 Sketch of Baxter's reconstruction

In Baxter's reconstruction, an Old Chinese syllable may have any of the following: a pre-initial, an initial, a medial (or two), a vowel, a coda, and a post-coda. There are only three pre-initials: *s, *ɦ, and a nasal consonant. Clusters with *s were generally simplified by the loss the following initial, though it sporadically metathesized with the following stop to form an affricate. *ɦ is reconstructed only before voiceless initials. Its effect was to cause these initials to become voiced in Middle Chinese. The nasal pre-initial occurs before oral stops and assumes the stop's place of articulation through homorganic nasal assimilation. The pre-initial then forced the deletion of the following oral stop, which may have been reanalyzed as excrescent, and so the word fell into the same homophonic group as the original initial nasals. The last two pre-initials are reconstructed to account for xiéshēng and morphological connections between words that alternate between oral and nasal or voiced and voiceless oral stops (Baxter, 1992: 176).

Baxter reconstructs a symmetric set of initial segments for Old Chinese. For stops, there are three laryngeal contrasts (voiceless unaspirated, voiceless aspirated, and voiced) at four places of articulation (bilabial, dental, velar, and labialised velar). For each place of articulation, there are also voiced and voiceless nasals. There are three dental affricates corresponding to the three contrasting laryngeal states. Baxter also reconstructs both simple and labialised glottal stops. There are four fricatives: *s, *z, *x, and *ɦ. There are also eight approximants: *l, *r, *j, *w, and their voiceless counterparts

(Baxter, 1992: 177). Baxter's medials are *r, *j, and rarely *l. The medial *j can occur with either of the liquid medials (1992: 178-179).

Baxter reconstructs six vowels: *i, *e, *a, *o, *u and another high, unrounded vowel, *ɨ, to which he assigns the feature [+ back] (1992: 180-181). Baxter's system has nine possible coda consonants: *j, *w, *m, *n, *ng, *p, *t, *k, and also the sequence *wk. While Baxter admits that *wk may be a final voiceless labialised velar, he does not reconstruct any words with the corresponding final labialised velar nasal and so does not know what to make of its possible phonetic realization (Baxter, 1992: 181).

Baxter's post-codas (*s and *ʔ) are responsible for the development of tones in Middle Chinese (qùshēng and shǎngshēng respectively) (1992: 182-183). Baxter's hypothesis is parallel to a theory about the formation of two tones in Vietnamese, as determined by the internal reconstruction of proto-Vietnamese. In fact, there is a correspondence between early loanwords from Chinese with qùshēng readings and the Vietnamese tone that reflects the proto-Vietnamese *s in coda position (Baxter, 1992: 308). Likewise, loanwords with shǎngshēng pronunciations correspond to Vietnamese reflexes of words with final glottal stops in the prehistory of Vietnamese (Baxter, 1992: 320-321). Also, early Chinese transcriptions of names that have /s/ in a syllable coda in the loan language correspond to qùshēng syllables (marked with *H* in Middle Chinese)

12) 央匱 Middle Chinese ʔjang gwijH > yāng kuì Tocharian B ankwaṣ 'asafoetida'
13) 阿迦貳吒 Middle Chinese ʔa kja nyiH træ > Ē jiā èr zhā for Akaniṣṭha
14) 都賴 Middle Chinese tu lajH > Dū lài for Talas (name of a river) (Baxter, 1992: 313).

Thus, it seems likely that Chinese tones may have been a development subsequent to the writing of the *Shījīng*.

2.2.2 Concerns with Baxter's reconstructions

The voiceless continuants reconstructed are a bit of a rarity. In fact, some linguists believe that no language should have a contrast between voiced and voiceless nasals (Crist, p.c., 2002). Furthermore, the place of articulation of voiceless nasals can be differentiated only by the transition to the following voiced segment (Maddieson, 1984: 71). In line with this, Baxter does not reconstruct any voiceless nasals before voiceless segments. The segment *hngw (a voiceless labialised velar nasal) does not occur in any of the languages documented in UPSID[13] (Maddieson, 1984: 239). Even though the reconstructed *hng, [ŋ̊], is found in seven UPSID languages, it is not clear that it is phonemic in all of those languages. Maddieson cites it as a phoneme of Burmese (1984: 239), which is distantly related to Old Chinese, but other sources treat Burmese voiceless nasals as allophones of voiced nasals after /h/ (Crist, p.c., 2002). Such exotic sounds as *hngw are not likely segments of Old Chinese, though they certainly may be. However, Baxter can find very few instances when he can unequivocally reconstruct *hngw (1992: 216). The underspecification of voiceless nasals and their appropriateness is further discussed in section 3.1.

Another interesting phenomenon in the reconstruction is the different effects of *j and *i. Most phonologists hold that [i] and [j] are the phonetic realization of the same feature bundle; [i] is the realization of the sound as a vowel and [j] is the realization of the same sound in a non-vocalic position (Buckley, 1994: 56). However, in Baxter's reconstruction of Old Chinese, the vowel *i does not cause palatalization the way that the

[13] UPSID is an acronym for the UCLA Phonological Segment Inventory Database. The database lists the phonemes of a representative sample of the world's languages. Maddieson's work is a statistical analysis of the frequency and co-occurrence of various phonemes in the inventories of these languages. Using these analyses, phonological hypotheses can be tested empirically test against the data in UPSID. Maddieson, Ian. 1984. *Patterns of Sounds*. New York: Cambridge University Press. pp. 1-3.

medial *j does. The dental initial in 15) is maintained lautgesetzlich, according to regular sound changes.

15) 替 *thij/ts > thejH > tì ‘to replace, to substitute’ (Baxter, 1992: 792)

This discrepancy is not simply a matter of rule ordering. Baxter orders dental palatalization (rule A.7) well before Hi > Mid (rule A.24), which lowered *i to Middle Chinese /e/. A difference between [j] and [i] has been proposed for Kashaya (Buckley, 1994: 56), a Native American language spoken in northern California (1994: 2). In Kashaya, the vowel [i] is completely unspecified for any vocalic features, such as rounding, height, or backness (Buckley, 1994: 32) and is the default vowel inserted epenthetically (1994: 103). The glide [j], however, is specified as Coronal, [+ distributed], and Dorsal (Buckley, 1994: 19). Perhaps, Old Chinese also had two different phonemes, *i and *j, with superficially the same phonetic realization, but different features underlyingly, though there seems to be no need, such as vowel epenthesis, to treat *i as unspecified by any vowel features. Baxter includes a lengthy discussion of the nature of medial *j, making many references to previous reconstructions in which the realization of *j was not phonetic at all, instead corresponding to some prosodic feature. Baxter fundamentally believes that *j was in fact a high front glide (1992: 287-290), but does not recognize or try to explain why it behaves differently than *i. In section 4.3, an analysis of Baxter’s *j as a palatalizing autosegment is discussed. If *j is such an autosegment, it follows naturally that the effects of *i and *j are different.

Baxter's reconstruction is also notable for the length of the syllables that are permissible in Old Chinese. It is possible for a word to have five onset consonants

because of the pre-initials and medials allowed by the syllable template. However, the longest syllables I found in Baxter's list of reconstructed words had the form CCCCVCC:

16) 順 *ɦsKjuns > zywinH > shùn ‘smooth, to arrange, to obey’ (1992: 789)
17) 碾 *Ntrjenʔ >nrjenX > niǎn ‘trample’ (1992: 221)

The capital K in 16) represents an undetermined voiceless velar, either aspirated or unaspirated, and the capital N in 17) represents the nasal pre-initial. The pre-initial and medials are responsible for pushing the reconstructed onsets of Old Chinese to four initial consonants, which is more than even English or Russian allow as a complex onsets.

18) squall [skwɑɬ]
19) straight [stɹɛɪt]
20) spew [spjuw]
21) Russian страна [s̪t̪r̪ʌ.ˈn̪ą̣] ‘nation’
22) Russian складно [ˈs̪kɫ̪ą̣d̪.n̪ə] ‘smoothly, coherently’
23) Russian здравствуйте [ˈz̪d̪r̪ą̣s̪.t̪vuj.t̪ʲɛ̣] ‘how do you do, how are you’[14]

Onset clusters of this length, however, are not completely unattested, as in Bella Coola.

24) [st’qʷlus] ‘black bear snake’
25) [qpsta] ‘taste’
26) [pɬtkn̩] ‘bark of bitter cherry tree (Kenstowicz, 1994: 290)

Not only is the length of the syllables surprising, especially given the Middle and Modern Chinese reflexes, which only have at most two initial segments that are the reflexes of the onset cluster, but so is the distribution of the pre-initials. The pre-initial *ɦ only precedes voiceless sounds. Thus, the initial is surrounded by voiced sounds since it is followed by a medial or vowel, both of which are always voiced. While this is a

[14] The transcriptions are based on the phonemic analysis in Maddieson, Ian. 1984. *Patterns of Sounds*. Cambridge, Cambridgeshire; New York: Cambridge University Press. p. 266. and adjusted after listening to readings by Kimberly Fedchak, professor of Russian at Swarthmore College. The glosses are taken from Falla, P. S., Wheeler, Marcus, Unbegaun, Boris Ottokar, and Howlett, Colin. 1993. *The Oxford Russian Dictionary*. New York: Oxford University Press.

perfect environment for voicing assimilation, it seems to form an odd sort of syllable, which seems to have no parallels. In all of the examples from English, Russian, and Bella Coola, the initial segments are either continuously voiced or go strictly from voiceless to voiced. When the vocal cords begin vibrating in a syllable, they do not stop until, possibly, during the coda. Even in the Imdlawn Tashliyt dialect of Berber, in which every phoneme can be a nucleus, if the nucleus is voiceless, the syllable contains no voiced segments.

27) [tḍtḳt] 'you sprained'
28) [tmṣsχt] 'you transformed' (Kenstowicz, 1994: 278)

This is in marked contrast to the syllables that Baxter reconstructs, which begins with a voiced segment, followed by necessarily voiceless segments, and then once again becomes voiced.

Russian obstruent clusters in onset position undergo progressive voicing assimilation, as can be seen below.

29) Russian сплошать [s̬pɬ̬ʌ.ˈʃˠa̬ʲt̬ʲ] 'to make a mistake, slip up (colloquial, perfect)'
30) Russian сближать [z̬bl̬i.ˈʒˠa̬ʲt̬ʲ] 'to bring together (imperfect)'
31) Russian сгнить [z̬gn̬ʲi̬ʲt̬ʲ] 'to rot, decay (perfect)'

The voicing of obstruent segments closer to the nucleus spreads to obstruents further form the nucleus, even in words that are not monomorphemic (Fedchak, p.c., 2003). Underlying, the first segment is voiceless, or unspecified for voicing, in the three example above. In 30) and 31), the initial fricative takes on the voicing of the following stops. Voiceless is also perceived as spreading across obstruents.

32) Russian вторник [ˈf̬to̬r̬.nʲik] 'Tuesday'
33) Russian вперёд [fpʲi.ˈr̬ʲo̬t̬] 'forward(s), ahead'

The initial fricative is written with the letter that normally represents [v] and speakers of Russian consider the phonetic realization to be devoicing the initial fricative. Even though the actual phonological process is not the spreading of voicelessness, in accordance with the representation of voiced and voiceless obstruents in autosegmental theory. Examples 18) through 33) all show that obstruent clusters very often have the same voicing.[15] If the sonority hierarchy were divided into two columns, one for voiced sounds and one for voiceless ones, it seems that moving from voiceless column to the voiced one can happen only once as in the onset as it approaches the syllable nucleus. More specifically, all obstruent clusters that violate the sonority hierarchy have the same voicing.

This principle is violated by 16) and 17). To bring these reconstructed words in line with it would require that the stops be voiced. If the stops are already voiced, then the appropriateness of *ɦ is brought into question. Further, the principle of xiéshēng contact only requires that the initials have the same place of articulation, not necessarily the same manner of articulation. There seems to be nothing that precludes voiced and voiceless stops from being in the same xiéshēng series. The voiceless nasal series is itself reconstructed only in xiéshēng series with voiced nasals (Baxter, 1992: 189, 193-196, 208-209, 212, 216-217). The reflexes of the voiceless nasals explain the alternation between Middle Chinese nasal and oral initials. Since xiéshēng contact occurs across these pairs of voiced and voiceless sounds, there seems to be no reason why voiced and

[15] Readers will likely notice an exception to this in the second syllable of 23), where [v] follows [t]. According to Professor Fedchak, the Russian /v/ is affected by progressive voicing assimilation, but does not affect other sounds, as is seen most clearly in the phrase в квартире [fkvʌɾ̪.ˈtʲi.ɾ̪ʲɛ̝] 'in the apartment'. In any case, the voicing of the syllable still proceeds strictly from voiceless to voiced sounds.

voiceless oral stops should not be in the same xiéshēng series. The reconstruction of words such as 16) and 17) is discussed further in section 3.2.1.

Even though Baxter's reconstruction raises several questions about the appropriateness of some reconstructed segments, it is a great advancement over previous reconstructions. It is phonologically motivated, as opposed to Bernhard Karlgren's, by a list of ordered rules of sound changes from Old Chinese into Middle Chinese. Open syllables are reconstructed, unlike in the reconstruction of Li Fang-kuei, bringing Old Chinese in line with every other observed language.

2.3 Sagart's Proposals

Some of Baxter's pre-initials and post-codas could also be morphemes in Old Chinese. The pre-initial *ɦ can make a passive or intransitive form of a verb from the active or transitive form, sometimes both represented with the same character (Baxter, 1992: 218). The suffixed post-coda *s could either make a noun out of a verb or a verb out of a noun. Baxter notes that there is a similar suffix operating in Tibetan, which is historically related to Old Chinese. This suffix can also explain rhymes in the *Shījīng* that are unexpected because only one tonal reading is attested in Middle Chinese. According to Middle Chinese sources, 害 only has a qùshēng reading, hajH > hài, reflecting an earlier post-coda *s, both as a noun and a verb meaning 'harm.' In the *Shījīng*, the verbal sense of 害, 'to suffer harm,' rhymes strictly as *at, without the post-coda *s. This reading would give rise to a rùshēng reading, but was lost.[16] The nominal sense rhyme strictly as *ats, giving rise to the documented qùshēng reading lautgesetzlich

[16] 害 also has another rùshēng reading, hat > hé, meaning 'what.' In this sense, 害 is also written 曷 *gat > hat > hé 'what' Baxter, William Hubbard. 1992. *A Handbook of Old Chinese Phonology*: Trends in Linguistics. Studies and Monographs: 64. New York: Mouton de Gruyter. p. 317.

(Baxter, 1992: 317). The morphological function of the post-coda *s explains how this rhyming pattern could occur, even in the absence of a historical record documenting the readings proposed.

Word classes in Classical (Late Zhōu) Chinese were very fluid; this may have been due to an affixal system. There are other variant tonal readings of the same character as a noun and a verb

34) 王 wáng ‘king’ and wàng ‘to rule’
35) 衣 yī ‘clothes’ and yì ‘to wear’

Also, in the following passage, nouns 君 jūn ‘ruler,’ 臣 chén ‘minister,’ 父 fù ‘father,’ and 子 zǐ ‘son’ all serve as nouns and verbs. The proof that they are verbs is that each is negated by 不 bù ‘not’ and not by 非 fēi ‘is not.’

36) 齊景公問政於孔子。孔子對曰君君臣臣父父子子。公曰善哉，信如君不君，臣不臣，父不父，子不子，雖有粟，吾得而食諸。
Duke Jǐng of Qí asked Confucius about governing. Confucius replied, “The ruler acts as the ruler, and the minister acts as the minister; the father acts as the father, and the son acts as the son.” The duke replied, “Wonderful! Truly, if the ruler does not act as the ruler, the minister not act as the minister, father not act as father and son not act as son, even though there is grain, will I get to eat it?” (Fuller, 1999: 41-42)

Such a system could easily be evidence for some sort of morphology, though Modern Chinese retains some of the fluidity almost without morphology.

Based on this evidence and comparison with neighboring Tibeto-Burman languages, Sagart proposes a reconstruction of Old Chinese with a largely prefixing morphological system that is typologically more like Gyarong or Khmer than any modern Chinese dialect or Middle Chinese itself (1999: 13). This is not surprising given that Old Chinese is believe to be toneless, unlike Middle Chinese and modern Chinese dialects, and that Old Chinese is believed to be related to the Tibeto-Burman languages.

2.3.1 Overview of Syllables in Sagart's Reconstruction

Phonetically, Sagart's reconstruction is very similar to Baxter's. Sagart accepts all of the same stops, affricates, and nasals. Sagart does not reconstruct *z, *ɦ, *hj, or *j as initials, though *j can be a coda consonant (1999: 28). The function of Baxter's pre-initial *ɦ is generally taken over by a nasalizing prefix *N-. Sagart uses alternative reconstructions with one of his prefixes for words that Baxter reconstructs with *N. Sagart does not reconstruct any medial consonants. He treats Baxter's medial *r as an infix and uses alternatives to reconstruct words with Baxter's medial *l. Sagart pushes Baxter's medial *j out of the segmental analysis, believing it to represent some prosodic feature. He simply labels words as being type A or B with a preceding superscript as discussed above (Sagart, 1999: 49). Sagart's system allows for an *r in coda position in addition to the coda segments reconstructed by Baxter (1999: 51).

Sagart's syllable template is thus simplified to CV(C)(ʔ). In his analysis, *ʔ did not serve a morphological function (Sagart, 1999: 134), though it also conditioned the development of shǎngshēng in Middle Chinese. In this way, Sagart eliminates certain problems in Baxter's work by removing some of the more exotic sounds from the segment inventory. Also, by splitting Baxter's monomorphemic reconstruction into several morphemes and shortening of the root syllable, the reconstructed syllables are more natural. Languages often allow consonant clusters separated by a morpheme boundary that would not be permissible if the cluster were found inside a single morpheme.

37) suitcase [sutkeɪs] suit+case
38) hotcake [hɔtkeɪk] hot+cake
39) wives [waɪvz] wife+s

40) cliffs [klɪfs] cliff+s

For example, the sequence [tk] in 37) and 38) does not occur in English morphemes. Likewise, the fricative sequences in 39) and 40) are not permitted in monomorphemic forms.

The affixes of Old Chinese were added to these roots. Prefixes could take two forms, either fused or iambic forms. In the fused forms, the prefixed consonant is added directly before the root onset. The iambic forms have a schwa or reduced vowel inserted between the prefixed consonant and the initial (Sagart, 1999: 17-18). Sagart believes that the iambic prefixes were simply lost without affecting the root initial, whereas the fused forms created consonant clusters whose development radically altered xiéshēng connections between words. Sagart proposes that iambic prefixes fell off in stages in order to account for xiéshēng connections like

41) 律 *^{B}Cə-lut > lwit > lǜ ‘follow a model’
42) 聿 *^{B}(Cə)-lut > ywit > yù ‘writing brush’ (1999: 129)

In the first word, the iambic prefix prevented the root initial *l from undergoing regular sound changes. In the second, the iambic prefix is inferred from the xiéshēng contact (Sagart, 1999: 18). Sagart believes that the iambic prefixes and their fused equivalents were in free variation and that their use of one as opposed to the other was a choice based on social and stylistic decisions. While the loss of iambic prefixes has diffused through the lexicon of most dialects, this type of prefixation is still preserved in some peripheral dialects (晉 Jìn, 閩 Mǐn, and 粵 Yuè) (Sagart, 1999: 19).

Infixal *-r- is inserted after the onset in unprefixed forms and after the prefix, but before the onset in prefixed forms (Sagart, 1999: 21)

43) 眼 *^{a}ŋrɨnʔ > ngɛnX > yǎn ‘eyeball’ and *^{a}ŋɨnʔ > ngonX >wěn ‘bulge, knob’
44) 闇 *^{a}q-lɨmʔ(-s) > ʔomX, ʔonH > àn ‘dark’
45) 黯 *^{a}qr-lɨmʔ(-s) > ʔɛmX > yǎn ‘deep black’ (Sagart, 1999: 115)

The application of Chinese morphology in this manner does not resolve the concern with Baxter’s reconstruction about the voicing of consonant clusters (Sagart, 1999: 129).

46) 謙 *^{A}k-hlem > khem > qiān ‘modest’
47) 歉 *^{A}k-hlemʔ > khemX > qiàn and *^{A}kr-hlemʔ > khɛmX > qiàn ‘modest’[17]

In the second reading of 47), voiced *r precedes voiceless *hl. Suffixes are simply applied to the end of the words and do not interact with other affixes.

2.3.2 Sagart’s Affixes

Sagart reconstructs seven different prefixes. The most well established is *s-. Sagart describes its function as deriving nouns from verbs, as well as deriving causative, inchoative, and directive verbs from other verbs and adjectives (which function like verbs in Chinese grammar). The directive verbs are applied to situations where the acts or states are directed towards external conditions.

48) 忸 *^{b}nruk > nrjuwk > niǔ ‘bashful, shy’
49) 羞 *^{b}s-nu > sjuw > xiū ‘ashamed of’ (Sagart, 1999: 71)

The prefix *N- derives intransitive verbs from transitive ones (Sagart, 1999: 75), examples of its use are given below in section 3.2.1 and Table 8 at the end of the text. Its phonetic nature is also discussed in 3.2.1. The prefix *m- is used in certain agentive nouns and names of small animals. It also prefixes verbs to express controlled action by volitional agents and imperatives (Sagart, 1999: 82-85). The prefix *t- marks some [- count] nouns (and a limited number of [+ count] nouns) (Sagart, 1999: 96). It also

[17] Sagart uses capital A and B as the superscript preceding words that do not have xiéshēng contact with words of the other type. If there is xiéshēng contact between type A and type B words, the Old Chinese reconstructions have a lowercase a or b, depending on the Middle Chinese reflex. Sagart, Laurent. 1999. *The Roots of Old Chinese*. Philadelphia: John Benjamins Pub. Co. p. 49.

marks various intransitive verbs, such as those describing involuntary physiological reactions, and adjectives, some of which are derived from unprefixed intransitive verbs (Sagart, 1999: 93-94). Sagart describes *k- as marking "actions and objects that are well-delineated in time and space" (1999: 107) such as [+ count] nouns, certain ethnic groups (1999: 107), verbs of momentary and continued actions derived from other verbs (1999: 100-101), and intensifying the unpleasant feelings of certain states (1999: 105). For two prefixes, *p- and *q- (the second of which may have phonetically been *ʔ-), their function is unclear (Sagart, 1999: 89, 109).

The infix *-r- can be applied to almost any word class. In adjectives, the effect is intensification (Sagart, 1999: 115-116). In verbs, *-r- marks a repeated and possibly durative action, an action occurring in multiple locations, or an action involving a collective participant (Sagart, 1999: 112-113). The infix is also used in the names of double or multiple objects, especially paired or multiple body parts (Sagart, 1999: 112, 115). Though there are still many unexplained situations in which *-r- is reconstructed, Sagart summarizes the use of *-r- as marking actions or objects distributed across time or space and treats the intensification aspect in adjectives as secondary (1999: 117).

Sagart is confident in reconstructing only one of his three suffixes, *-s. The main use of *-s was to derive nouns from verbs. A secondary use, deriving verbs from other verbs, is agency reversal, as in the examples below.

50) 學 *^{a}gruk > hæwk > xué 'to study' and *^{a}gruk-s > hæwH > xiào 'to teach'
51) 乞 *^{b}khɨt > khjɨt > qǐ 'to beg' and *^{b}khɨt-s > khjɨjH > qì 'to give alms'
52) 受 *^{b}duʔ > dzyuwX > shǒu 'receive'
53) 授 *^{b}du(ʔ)-s > dzyuwH > shòu 'transmit'

In each of these pairs, there is the same set of semantic roles: goal, source, and something moving between them. The difference between each is whether the goal (in the form without the suffix) or the source (the form with the suffix) is given agency over the situation. This is to say that the direction of the action changes. Translating 梅祖麟 Méi Zǔlín's Chinese terms, Sagart writes that *-s transforms "endodirectional" or "endoactive" verbs (內向動詞) into "exodirectional" or "exoactive" verbs (外向動詞) (Sagart, 1999: 133).

Sagart believes that the pair 買 mǎi and 賣 mài also fit this pattern, as one would expect based not only on the Middle Chinese and Modern Mandarin pronunciations, but also on the set of semantic roles involved. Sagart provides the following tentative reconstructions[18]

54) 買 *^{a}mreʔ > mɛɨX > mǎi 'to buy'
55) 賣 *^{a}mreʔ-s > mɛɨH > mài 'to sell'

However, Sagart notes that 賣, with the reading mɛɨH, is not used until relatively late. It is first written in *Shuōwén jiězì* as $\overset{出}{買}$, combining the characters 出 chū 'to go out' and 買. The use of directional verbs to relate written words for buying and selling is most strikingly seen in the following pair of verbs

56) 糴 *^{a}l[i,e]wk > dek > dí 'to buy grain'
57) 糶 *^{a}hl[i,e]wk-s > thewH > tì 'to sell grain' (Sagart, 1999: 206-207)

The characters representing these verbs are distinguished only by the radical in the upper left corner, which correspond to

[18] The existence of an earlier verb meaning 'to sell' and also written 賣, *^{a}l[iw]k > yuwk > yù makes Sagart hesitant to propose 54) and 55) as the reconstructions of these words. Ideally, Sagart would like the reconstructions of 54) and 55) to be related to this word because of the obvious relationship between them. The alternatives still reflect the *-s suffix. Ibid. p. 206.

58) 入 *^{b}nɨp > nyip > rù ‘to enter’ (Sagart, 1999: 133)
59) 出 *^{b}t-khut > tsyhwit > chū ‘to go out’ (Sagart, 1999: 88)

These two directional verbs combined with 米 mǐ, a word for cereal grains (Sagart, 1999: 178), form the semantic elements of these words. This graphic tradition in verbs of commerce fits well with the agency reversal, or shift in direction, of the pair of verbs.

The other two suffixes that may have been productive in Old Chinese were *-ŋ and *-n. The former suffix may have had a vocative function (Sagart, 1999: 135). It is discussed further in section 3.2.2 below; the data is also repeated in table 9. The latter suffix was possibly a nominative suffix.

Several aspects of Sagart’s reconstruction are considered below. First is the use of the prefix *N- and its relative merit against an autosegmental interpretation of Baxter’s pre-initial *ɦ. Then, the merits of adding *-ŋ to the inventory of Old Chinese affixes are considered in terms of suffix licensing. Finally, two autosegmental hypotheses about the differences between type A and type B syllables are compared with previous theories, including Baxter’s, and with each other.

3. Segmental Underspecification and Licensing of the Old Chinese Syllable

Autosegmental phonology is a theory that developed in response to the feature bundle analysis of Chomsky and Halle’s *The Sound Pattern of English*. This theory strips away the features associated with any segment and places them on separate tiers, leaving bare timing slots called skeletal positions. Association lines are drawn between the various tiers and the skeletal or CV tier. As the features are associated with different skeletal positions, the feature bundles of Chomsky and Halle’s analysis take shape. This analysis has been useful in explaining various phonological phenomena: tonal changes in

African languages, the morphology of Arabic, compensatory vowel lengthening, homorganic nasalization, and the indivisibility of geminate consonants.

Underspecification and licensing are two ideas that have been developed in autosegmental phonology. Goldsmith mentions two principles of underspecification. The first is that any redundant or non-contrastive features, such as voicing in vowels, should not be specified. The second is that any feature that can be specified later by a default rule should not appear underlyingly (Goldsmith, 1990: 219-220). The theory of underspecification fits well with the barebones autosegmental approach to phonology. One effect of underspecification is to reduce the amount of underlying information associated with each phoneme.

Licensing is a way of organizing the features that associate to the positions on the skeletal tier into syllables and morphemes. A licenser, such as a syllable or coda, can assign a set of autosegments or features to the skeletal positions it dominates. Also, any licenser can license a segment at most one time. This restriction prevents sequences like /sf/ from being initials because both /s/ and /f/ are specified as [+ continuant]. Not all licensers can license same set of autosegments. The syllable must license all of the distinctive features since any phoneme can be a syllable onset, but the same is not true of codas. Codas across languages generally license only a smaller set of autosegments. A consequence of using licensing in this way is that it explains why the onset and nucleus are assigned one mora, but a coda consonant may be assigned its own mora in quantity languages. In general, a mora may be assigned to each licenser. Since the onset and nucleus are both dominated by the same licenser, they are assigned only one mora. The

coda, as its own licenser may, or may not, be assigned a mora; this is determined in a language particular fashion (Goldsmith, 1990: 123-127).

3.1 Underspecification of Consonants

Using the ten features in the underspecification scheme for Sagart's Old Chinese initials in Table 3.1, any segment is uniquely identified by a combination of up to two features, with the exception of voiceless nasals. This table does not include Sagart's

Table 3.1

Segment	POA	Round	Aspiration	Affricate	Voice	Continuant	Nasal	Liquid	Lateral
*p	Labial								
*t	(Coronal)								
*k	Velar								
*kw		+							
*ph	Labial		+						
*th	(Coronal)		+						
*kh	Velar		+						
*kwh		+	+						
*b	Labial				+				
*d	(Coronal)				+				
*g	Velar				+				
*gw		+			+				
*ts				+					
*tsh			+	+					
*dz				+	+				
*s	(Coronal)					+			
*x	Velar					+			
*hw		+				+			
*hm	Labial				-		+		
*hn	(Coronal)				-		+		
*hng	Velar				-		+		
*hngw		+			-		+		
*m	Labial						+		
*n	(Coronal)						+		
*ng	Velar						+		
*ngw		+					+		
*hl					-				+
*l									+
*hr					-			+	
*r								+	
*ʔ			+ constricted glottis						
*ʔw		+	+ constricted glottis						

initial *w. Since there is no evidence that *w is not simply *u in a consonantal position, *w and *u have the same underlying features. An underspecification scheme for vowels is discussed below in section 3.3.1. Sagart's reconstruction has a gap, *j as an initial, so its underspecification is not considered, but would be specified the same way that *i is. *hw, the voiceless counterpart of *w, is treated as a labialised velar fricative, as in the IPA [ʍ]. According to this underspecification scheme, then, there is also another gap, a labial fricative.

Certain features were left unspecified because they could be filled in later by default rules. Since all of the affricates and liquids are redundantly coronal, they are unspecified for place of articulation. Likewise, all labialised sounds, except *ʔw, are velar, so that can be filled in later. The rule assigning velar place of articulation to all [+ round] consonants would be blocked by the feature [+ constricted glottis]. The rule assigning coronal place of articulation to all segments otherwise unspecified could either apply afterwards or not to segments labeled [+ round] or [+ constricted glottis].

The initial *ʔw is slightly out of place in that it is marked with [+ round], but does not have a velar place of articulation. This segment is not well attested as a phoneme in UPSID; only one language has [ʔ^{w}], whereas over two-fifth have [ʔ] (Maddieson, 1984: 215). Furthermore, it is not uncommon for *ʔw to be in xiéshēng series with *w.

60) 王 *wjang > hjwang > wáng, Cantonese wong4 'king'
61) 枉 *ʔ^{w}jangʔ > ʔjwangX > wǎng , Cantonese wong2 'bent, crooked; depraved, unjust' (Baxter, 1992: 217)
62) 員 *wjɨn > hjun > yún 'member; person engaged in some activity'
63) 隕 *ʔ^{w}jɨnʔ > ʔjunX > yǔn and *wrjɨn(ʔ) > hwinX > yǔn 'to fall; to die' (Baxter, 1992: 806)

The Cantonese pronunciations of 60) (with 陽平聲 ‘lower píngshēng’ tone) and 61) (with 陰上聲 ‘upper shǎngshēng’ tone) reflect the regular tonal reflexes of voiced and voiceless initials, respectively, after the Great Tone Split. After the devoicing of the earlier voiced initials, each tone split into two groups, the 陰 ‘upper’ tone, composed of the historically voiceless initials and the 陽 ‘lower’ tone, composed of the earlier voiced initials (Bauer and Benedict, 1997: 155-157). Based on the Modern Chinese reflexes of words such as 61), Baxter reconstructs ʔ as a Middle Chinese initial. In this way, all words with voiceless initials, including ʔ, have upper tone pronunciations (Baxter, 1992: 59).

An argument can be made that the supralaryngeal features are the same, allowing for xiéshēng contact. However, a simpler explanation is that Sagart’s *q-, which may have been *ʔ or had debuccalized to it, was prefixed to 60) and 62) to derive 61) and 63). Though Sagart allows for the reconstruction of *ʔw, he does not reconstruct any words with this initial. It may be possible to account for all reconstructions of *ʔw in terms of Sagart’s *q- prefix attached to a root beginning with *w.

The underspecification scheme presented in the table marks the underlyingly uncommon features of the segments, such as nasality or aspiration. To this end, voicing should be viewed as an equipollent feature (one that can take two different values, [+ voice] and [- voice]) in the reconstructions presented by Baxter and Sagart. This is in contrast to all the other features, which are all privative (only taking on one value, i.e. [+ nasal]), where the other state is taken as the common, unmarked case. A vast majority of nasals, liquids, and approximants are voiced in Old Chinese, as in all the languages of the

world. The situation is the reverse for stops, affricates, and fricatives. Obstruents that are [+ voice] are more rare, making them the marked case for these sounds. These facts and the principle that underspecification marks the less common cases necessitate the interpretation of the voicing feature as above. Treating voicing this way means that voiceless nasals need three features to identify them as opposed to the two features necessary for every other segment. This should cause us to re-evaluate the evidence for the voiceless nasals.

3.2 Applications of the Underspecification to the Licensing of Old Chinese Affixes

Morphemes also license features for the skeletal positions that they govern (Goldsmith, 1990: 127). Since the underlying specification of the features as stated above are the distinctive features of the segments, they will not change if the segment is licensed by the syllable or an affix, we can use it to analyze aspects of Baxter and Sagart's Old Chinese reconstructions.

3.2.1 Sagart's *N- Prefix

Many linguists reconstructing Old Chinese have noticed an alternation between voiced and voiceless initials in semantically related Middle Chinese words that have xiéshēng contact. Specifically, the word with a voiceless initial is a transitive verb and the word with a voiced initial is intransitive, as in the following two pairs:

64) 見 Sagart *[a]ken-s, Baxter *kens > kenH > jiàn 'to see'
65) 現 Sagart *[a]N-ken-s, Baxter *ɦikens > henH > xiàn 'to appear'
66) 折 Sagart *[b]tat, Baxter *tjat > tsyet > zhé 'to break, to bend (transitive)'
67) 折 Sagart *[b]N-tat, Baxter *ɦtjat > dzyet > shé 'to bend (intransitive)'

To account for these facts, Baxter reconstructs pre-initial *ɦ, whose phonological effect is to voice the following initial. Sagart proposes a nasal prefix *N- that is

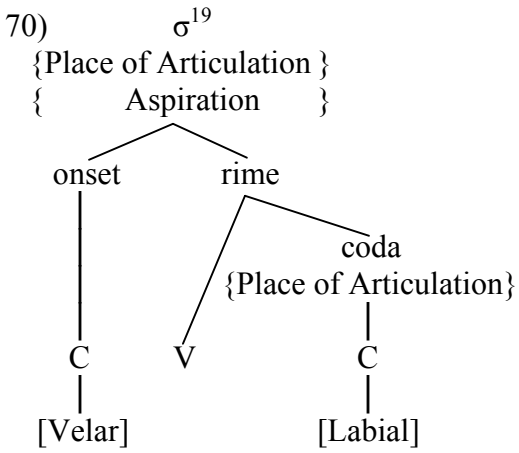
homorganic with the following stop, and can be interpreted as prenasalization. Unlike the other prefixes that Sagart reconstructs, *N- cannot appear in an iambic form and the infix *-r- appears after the root initial, not after *N-. This is unlike the other prefixes and, perhaps, reinforces the view of it as prenasalization. Also, *N- does not force the deletion of the Old Chinese root initial, as the other prefixes do under most circumstances. These facts, in addition to reconstructed prenasalized loans from Chinese into Miao-Yao languages, strongly favor the interpretation of *N- as prenasalization, rather than as a normal prefix like *s- or *t-.

The behavior of *N- can easily be accounted for by having a prefix that licenses a [+ nasal] segment but is not matched with an element on the skeletal tier. Thus, the [+ nasal] segment associates with the nearest consonant position on the skeletal tier, namely the root initial; the result is prenasalization of the root initial stop or affricate, given that [- nasal] is already assigned to the root consonant by a default rule. If the infix *-r- is inserted after the first C position on the skeletal tier, then it must be inserted after the root initial, as with unprefixed words, because the *N- is not attached to its own skeletal position.

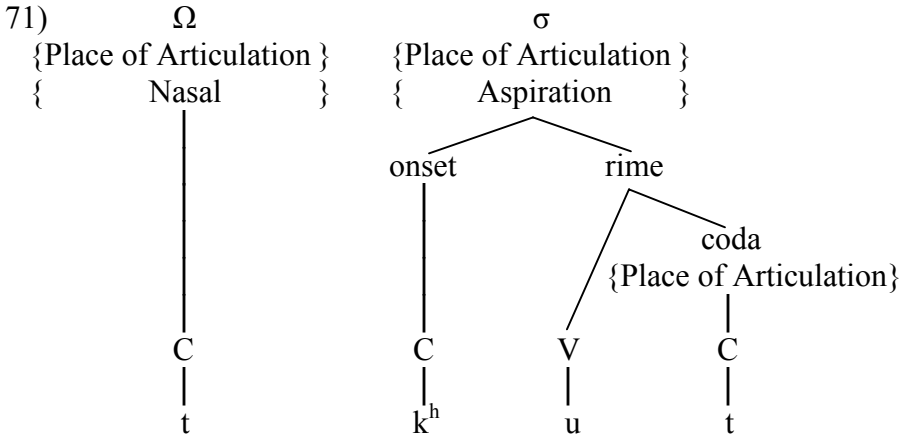
How affixation works for two different words, both with *-r- but with different prefixes is compared below.

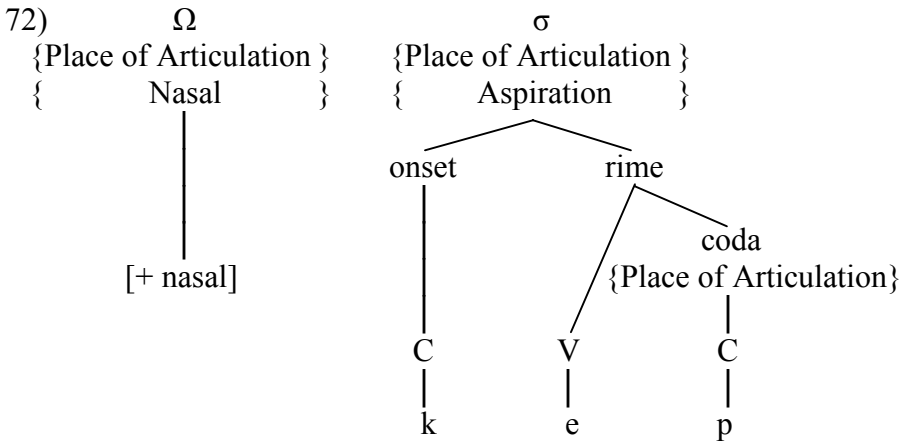
68) 黜 *^{b}tr-khut > trhwit > chù ‘to expel’ (Sagart, 1999: 111)
69) 狹 *^{a}N-krep > hɛp > xiá ‘narrow, pressed on both sides’ (Sagart, 1999: 75)

Before the affixes are added they each have the same structure. For example 69) above, syllable (marked with σ) licenses a velar place of articulation to the onset. The coda licenses labial place of articulation.



After this, the prefixes (marked with Ω) are added as in 71) and 72) below

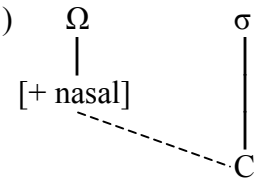




Since the [+ nasal] autosegment is unassociated, rule 73 below will associate it to the following consonant.

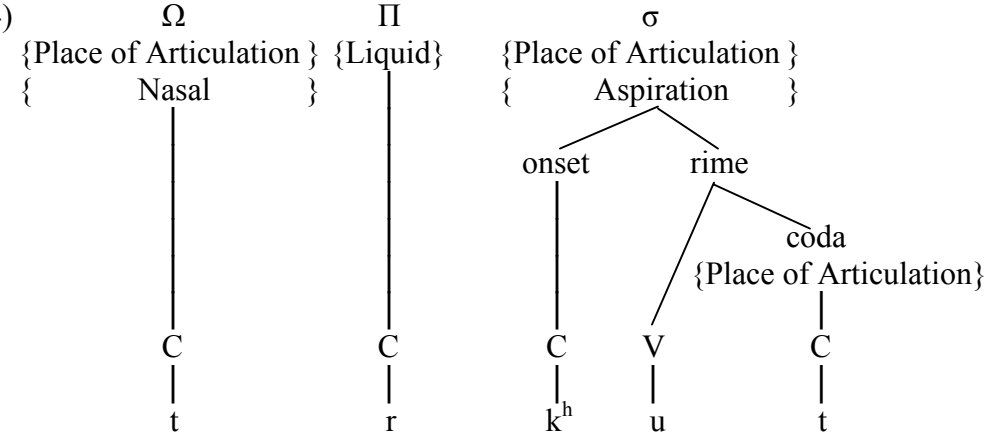
[19] I have place the relevant autosegments that a licenser licenses in curly braces, {}, under the licenser.

73)

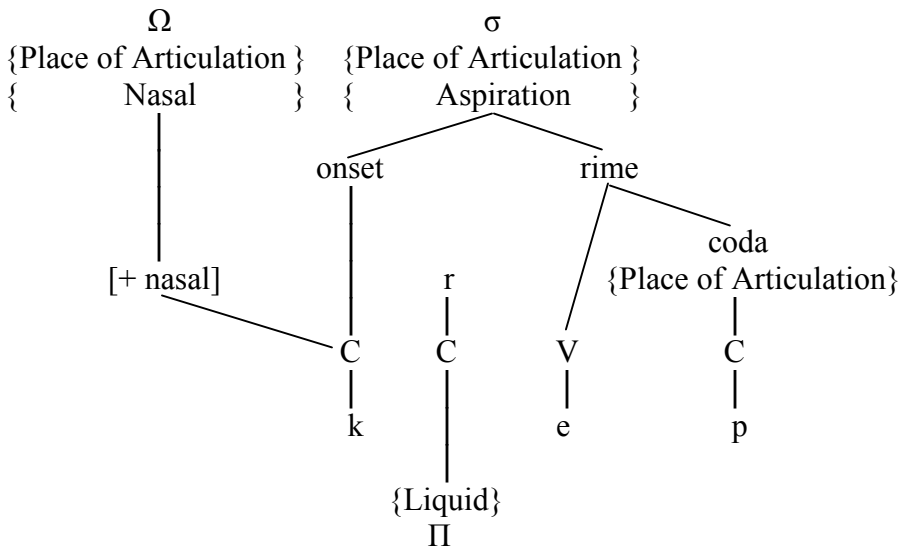


The associations after the application of rule 73 describe prenasalization of the initial stop, assuming [- nasal] is also attached to the timing slot by a default rule. Next, the infix (marked with Π) is added after the first skeletal position giving us the form

74)



75)



The placement of Π in example 75) is simply a graphic device so that the association lines of the syllable are not crossed. However, the autosegments of the infix would be found a different tier that the rest of the autosegments, though, so the display is artificial

only in the placement of the letter *r*. Thus we derive Sagart's reconstructions in 68) and 69).

While this is a nice explanation of the facts, there is a question of what would motivate such a situation in the first place. It is certainly not the case under Sagart's analysis that a prefix cannot license a nasal segment to its own position on the skeletal tier. For example, Sagart reconstructs a prefix *m-. This certainly has its own skeletal position and may be recorded in an iambic form in the *Shījīng* (Sagart, 1999: 82). Obviously, the [+ nasal] segment must accompany the [Labial] segment, but is it enough to say that a place of articulation segment is necessary for the prefix to command a skeletal position? The answer seems to be no if a coronal place of articulation is taken as unmarked. Sagart also reconstructs a prefix *t-, which remains completely without specification underlyingly, other than a C slot on the skeletal tier.

So, what is the explanation for the lack of a skeletal position for *N-? Let us begin by reviewing the facts that Sagart presents as evidence for this prefix. First, there are the Middle Chinese alternations between voiced and voiceless stops in pairs of intransitive and transitive verbs. These pairs of words have xiéshēng connections, so they are likely to be cognate. First, the Middle Chinese reflex is simply to contrast the pairs in terms of voicing of the initial consonants. Also, it remains unclear that any modern dialect of Chinese preserves the prenasalization that Sagart asserts was the phonetic realization of the prefix *N. According to Maddieson, standard Hakka (a dialect also called 客家 Kèjiā, based on the Mandarin pronunciation, as opposed to the Cantonese pronunciation) has prenasalized stops, but no simple nasal stops. In his discussion of prenasalized stops, Maddieson notes that prenasalized stop series pattern

more closely with the stop series found in the language than with the nasal series. Since Hakka lacks a voiced stop series, Maddieson places it in the category of languages for which the prenasalized stops are substituted for plain voiced stops (Maddieson, 1984: 67-68). Norman, in his survey of Chinese dialectology and using the same study of Kèjiā, concludes that Kèjiā has simple nasals but does not have either voiced or prenasalized stops (Norman, 1988: 224). The set of initials in Norman's analysis is essentially the same as most other dialect families and seems the most natural. In general, nasals in modern forms of Chinese do not show much variation, such as homorganic changes. Perhaps this is not surprising in light of the fact that Chinese is generally held up as the paragon of an isolating language, in which words are collocated with only a semantic connection between them. However, Baxter treats nasals as rather stable synchronically as well, allowing for some variation of *n based on the effects of medials. Prenasalization, as proposed, would be very different from the characterization of nasal segments in other parts of the syllable.

There are also similar alternations, generally also between voiced and voiceless initial obstruents, in recorded Miao-Yao languages. The linguists who reconstructed Proto-Miao and Proto-Yao believed that the alternations stemmed from a Miao/Yao prefix that caused prenasalization. Sagart believes that these words were borrowed from Chinese and that the prefix was productive in Old Chinese morphology. In the case of Proto-Yao, the modern reflexes of the words borrowed with prenasalization are all distinguished by a contrast in voicing, as in Middle Chinese. Below is the data Sagart provides from the Mien dialect of Yao (Sagart, 1999: 75-76).

76) tsheʔ$_7$ (< *tsh-) 'to pull down, pull apart' and dzeʔ$_7$ (< *ntsh-) 'to be cracked, as earth' from 坼 Middle Chinese trhæk > chè 'to split, be rent; fissure'

77) $khai_1$ (< *kh-) ‘to open’ (transitive: as in ‘open the door’) and
gai_1 (< *ŋkh-) ’to open’ (intransitive: as in ‘the heart opens: be happy’ and ‘to blossom’ from 開 Middle Chinese khoj > kāi ‘to open’
78) $gaat_7$ (< *ŋkh-) ‘thirsty’ from 渴 Middle Chinese khat > kě ‘thirsty’

For these three examples, there seems to be no record of a parallel alternation in the Chinese words that were loaned into Yao, as Sagart only provides corresponding Middle Chinese words with voiceless initials. This in turn raises questions about the designation of both words of the pair in 76) and 77) as Chinese loanwords since only one member of the pair survives to Middle Chinese and the other is unattested elsewhere. There is also a lack of correspondence between the Mien and Chinese words in 78) because the Mien form has a voiced onset but the Middle Chinese form is not voiced. Furthermore, 78) is certainly not an intransitive verb derived from a transitive one; it is an adjective.

The reflexes of two different pairs in White Hmong, a Miao language, are, however, distinguished by an initial nasal (Sagart, 1999: 76-77).

79) qe_4 (< *ga_B) ‘low’ and nqe_4 (< *ŋga_B or *NGa_B (Sagart, 1999: 208)) ‘go down, descend’ from 下 *^{a}(N-)graʔ > hæX > xià ‘go down, down, low’
80) $tshia_1$ (< *tsh-) ‘new, fresh, bright’ and $ntshia_1$ (< *ntsh-) ‘clear, fresh’ from 清 Middle Chinese tshjeng > qīng ‘clear, pure, bright’ and 星 Middle Chinese dzjeng > xīng ‘to clear during the night, of the weather’

In a footnote, Sagart says that the alternation in 79) may not, in fact, be related to *N-, but rather *m-, because the word classes do not match the other alternations. Also, the addition of *N- to voiced segments is something that Sagart generally does not consider. Given the Middle Chinese reflexes of voiceless stops, *N- would have had no effect on a voiced root initial (Sagart, 1999: 75-76). The pair in 80) is the only group with a full set of Chinese and Miao-Yao alternations. Further reconstructed loans take us farther afield, as with another reconstructed loan with a lack of correspondence between the Proto-Miao and Middle Chinese initials in 81).

81) Proto-Miao *$\text{N}qa_{\text{B}}$ ‘price’ from 價 Middle Chinese kæH > jià ‘price’

Notice that even the word classes in 80) and 81) do not match the alternation described above. Even if 星 were derived from 清, 星 is not an intransitive verb derived from a transitive verb, since 清 is an adjective. 價 is not even a verb; it is a noun. While originally an explanation for the alternation of the initials in pairs of transitive and intransitive verbs, *N- is generally reconstructed in words of any class to explain an alternation between voiced and voiceless initials in xiéshēng series.

The function of this prefix, most broadly, is to cause the root initial to be voiced. In Sagart’s brief discussion of the possibility that Old Chinese also prefixed the voiced counterparts of the prefixes he reconstructs, he asserts that there were two stages of prefixation. First, the original voiceless, fused prefix was attached. Then, after the prefixed initial replaced the root’s original initial consonant according to regular sound changes, *N- (or in certain cases *m-) was prefixed. The result, again by regular sound changes, was the voicing of the initial. At first glance, the reconstruction seems to point to voiced prefixes of limited scope that can, in fact, be explained by two steps without the complication of the set of prefixes. Two stages of prefixation may explain why *N- is not associated with its own skeletal position. It may have been the only element of a set of second stage prefixes that did not govern a skeletal position. Alternatively, *N- may also have been a full segment that was later cliticized, though the homorganic nasalization of *N- may still have to be explained.

In sum, the nasality of the prefix is unimportant in itself. The important feature is voicing. If no prefixes were underlyingly voiced (and there is no need to assume that any were if *m is taken as unmarked and *hm as marked) and *N- were simply a [+ voice]

autosegment, then *N- may not be allowed a skeletal position on account of its licensing a [+ voice] segment. This, in turn, could be explained by a constraint on underlyingly voiced prefixes. However, not allowing the [+ voice] segment its own skeletal position, forcing it to attach to the following consonant of the root, seems odd and perhaps capricious, not to mention the likely lack of parallels. There is no reason from a theoretical standpoint why there should not be floating [+ voice] autosegments, but prenasalization is much more common and always associated with voiced stops (Maddieson, 1984: 67).

3.2.2 Sagart's *-ŋ Suffix

Sagart mentions that there may be a vocative suffix *-ŋ (1999: 134). In support of this he offers the following data:

82) 我 *^{a}ŋajʔ > ngaX > wǒ 'we, I'
83) 卬 *^{a}ŋajʔ-ŋ > ngang > áng 'I, we, we on our part, I on my part'
84) 女 *^{b}nraʔ > nrjoX > nǔ 'woman'
85) 孃 *^{b}nraʔ-ŋ > nrjang > niáng 'mother, originally a term of address'

The other possible suffixes (*-s and *-n) are not licensed for a place of articulation, unlike *-ŋ. Since Sagart only adduces two pairs as evidence for the suffix, and given the relative simplicity of suffixes not being able to license a place of articulation, it may be preferable not to treat the alternation as a suffix.

3.3 Underspecification of Vowels

According to Goldsmith's analysis of underspecification account of vowel harmony and disharmony, the canonical five-vowel system can be reduced to only the use of two features, an equipollent value for rounding (which need not be assigned to all vowels) and a privative feature for height (which assigns a feature [+ low] to all vowels except [i] and [u]). Goldsmith mentions that this analysis can be extended to account for

a system of six vowels, the canonical five and a schwa, which is neither marked for rounding nor for height (1990: 302-303). While the six-vowel system reconstructed for Old Chinese does not include a schwa vowel, it is tempting to apply this analysis to the language. Table 3.2 shows how these features specify the six vowels of Old Chinese.

Table 3.2

	[- round]		[+ round]
	*i	*ɨ	*u
[+ low]	*e	*a	*o

3.3.1 Underspecification of Old Chinese Vowels

In favor of this analysis is a type of reduplicative compounding where the vowel alternates between *e and *o or between *i and *u

86) 邂逅 *[a]greʔ-groʔ > hɛɨX-huwX > xiè hòu ‘carefree and happy’
87) 蟋蟀 *[b]srit-srut > s(r)it-srwit > xī shuài ‘cricket’ (Sagart, 1999: 137)

Under this analysis, the only feature that would change under the reduplication is the specification for rounding. Since backness is later assigned based on roundness, a [- round] feature is simply changed to [+ round] in the following vocalic slot, or vice versa if the reduplication is based on the rightmost syllable. This also explains why the vowels *a and *ɨ do not take part in this sort of reduplication. These two vowels are not assigned a value for roundness, so it is impossible to change this feature, as the reduplication requires.

In Baxter’s final analysis, he separates the different rhymes that the Qīng dynasty philologists conflated. In each case, the rhymes that were conflated were all marked [+ low] or unmarked. In the views of the Qīng dynasty philologists, all of these rhymes appeared to constitute the same group (Baxter, 1992: 562-564). On the other hand, these two groups of vowels form natural classes under any system (under a traditional analysis,

the vowels are all either [+ high] or [- high]), and this should not be taken as an endorsement of this underspecification scheme.

An apparent flaw in this analysis is that it does not do a good job of explaining why *ɨ, *u, and *o do not occur before labialised velar coda consonants (Sagart, 1999: 61). According to the more traditional specification scheme used in Baxter, these vowels are all [+ back] and [- low]. These vowels do not form any class under a Goldsmith-like analysis, however. *ɨ is not specified for any vocalic features. *o is [+ round] and [+ low]. *u is simply [+ round]. If the [+ round] requirement is relaxed to a non-standard analysis of *not [- round]*, then the natural class also includes *a. To exclude *a by adding a *not [+ low]* requirement also eliminates *o from the natural class.

However, this may not be as much of a problem as it seems. For *u and *o, there is a restriction on their occurrence not only with labialised velar coda consonants, but also with labial coda consonants (Sagart, 1999: 56). The restriction with *ɨ is only for labialised velar consonants. Sagart does not discuss this gap at all. Baxter only states that he does not reconstruct sequences like *ɨw simply because he sees not need to do so (Baxter, 1992: 297). As for sequences like *ɨwk, Baxter once again states that the coda *wk is out of place since there is no corresponding nasal coda. Furthermore, *wk and *w alternate in several xiéshēng series, leading Baxter originally to believe that *wk may have been *wʔ and that the final element may have been a morpheme of Old Chinese (Baxter, 1992: 301-302). He leaves out any discussion about the possible co-occurrence of *ɨ and *wk. However, by comparing *ɨwk to *ɨw, it seems likely that *ɨwk is not reconstructed because there was no need to do so. The absence of codas like *ɨw and

*ɨwk is most likely accidental, unlike the sequences *uw and *owk, which would violate the Obligatory Contour Principle through licensing, although some languages, such as English, allow violations of the Obligatory Contour Principle, as in *woo*.

After the affixation process, the vowel *ə is also a member of the vowel matrix, since it optionally appears between the prefix and the root. There are several ways to specify each of the vowels in such a matrix. One is to maintain [+ low] as a privative feature and add a privative [+ high] feature to *i, *ɨ, and *u. At this point, the vowel features essentially reduplicate the traditional feature-bundle analysis to describe a seven-vowel system, as in Table 3.3.

Table 3.3

	[- round]		[+ round]
[+ high]	*i	*ɨ	*u
		*ə	
[+ low]	*e	*a	*o

This analysis has two apparent problems. First, the vowel grid using these features does not accord with either the traditional feature-bundle or the phonetic analysis. It assigns the same lowness to *e, *a, and *o. Also, the analysis misses the generalization that both of these feature specify a single phonetic quality, vowel height. Capturing this generalization and maintaining [+ low] underlyingly would necessitate making [low] an equipollent feature assigning [- low] to the vowels *i, *ɨ, and *u and maintaining the [+ low] feature only for to *a.

Starting with [+ low] marked underlyingly, requires this feature to be added to or changed for five of the six root vowels. A better way of explaining this is assigning [+ high] to the vowels *i, *ɨ, and *u underlyingly, as in Table 3.4, instead of marking *e, *a, and *o [+ low].

Table 3.4

	[- round]		[+ round]
[+ high]	*i	*ɨ	*u
	*e	*a	*o

After the addition of *ə, [+ high] would be changed to an equipollent feature assigning [- high] to *a, as shown in Table 3.5.

Table 3.5

	[- round]		[+ round]
[+ high]	*i	*ɨ	*u
	*e	*ə	*o
[- high]		*a	

The result is a simple derivation of the standard vowel space, in which the schwa is unmarked for any vowel features, instead of *a, which is unmarked at the deepest level and then specified with the feature [- high] during affixation. If the inserted vowel is a reduced root vowel, then the alternation may be between *ə and *a, as 88) below indicates. The feature [- high] may then be assigned to the vowel with primary stress as a default rule.

The vowel specifications in Table 3.5 would also explain the use of the following two words to represent the iambic prefixes Sagart describes.

88) 無 *^{b}ma > mju > wú ‘there is not, do not’ (Sagart, 1999: 82)
89) 不 *^{b}puʔ > Late OC *^{b}piʔ > pjuwX > bù ‘not’ (Sagart, 1999: 88)

These are reconstructed with the vowels *a and *ɨ at the time the poems are believed to have been written. Since *ə cannot appear as a root vowel and Chinese characters represent Old Chinese roots according to Sagart, there is not the option of using a character with a schwa vowel to represent the iambic prefixes. All of *a, *ɨ, and *ə are unmarked for roundness, making *a and *ɨ the most likely candidates to represent the

prefix. The only differences between *ɨ, *ə and *a is a marking for height under the analysis presented above.

3.3.2 Labialised Velar Initials and Their Reconstructed Vowels

Baxter does not generally reconstruct words with labialised velar onsets and round vowels; there are only three exceptions, two of which are discussed in Baxter's rule A.9: Rounding dissimilation (1992: 570). Specifically, the rule is designed to apply to the words

90) 頃 *g^{w}rju > gwij > kuí 'cheek bone, bones of the face' (Baxter, 1992: 216)
91) 軌 *k^{w}rjuʔ > Late OC *k^{w}rjɨʔ > kwijX > guǐ 'wheel-axle ends'
92) 逵 *g^{w}rju > Late OC *g^{w}rjɨ > gwij > kuí 'thoroughfare' (Baxter, 1992: 570)

The first and second examples are reconstructed based on xiéshēng contact with 九 jiǔ 'nine,' which is discussed below. The second and third words rhyme with others that must be reconstructed with *u as the main vowel in the *Guófēng* section of the *Shījīng* (Baxter, 1992: 355), which is the most diverse part of the collection. Their Middle Chinese reflexes, however, reflect words reconstructed with *ɨ, like

93) 龜 *k^{w}rjɨ > kwij > guī 'turtle, tortoise' (Baxter, 1992: 570)

Because of these facts, Baxter posits his rounding dissimilation rule.

It is important to remember that the language of the *Shījīng* may not, in fact, be the predecessor of Middle Chinese. There is evidence that changes that affected the rhyming patterns of the *Shījīng* are not also changes that affected the language described by the Middle Chinese lexicographers (Sagart, 1999: 6-7). Furthermore, 龜 rhymes in the *Dàyǎ*, which is one of the older sections of the *Shījīng*, as stated above. The other two words, 逵 and 軌, are found in poems that may not be as old. Hence, the situation may be

a reflection of different environments for the rule rounding assimilation. In the dialect of the *Shījīng*, assimilation may not have been blocked by *r, whereas, in the dialect that evolved into Middle Chinese, it was blocked by *r.

In fact, the only clear cases of the application of rounding dissimilation involve words with *r (Baxter, 1992: 570). Since Baxter does not have any other rule that requires the presence of *r, he attempts to show that the rule could be applied to all words with labialised initials and round vowels. To this end, he proposes an interesting history for the word 九:

94) 九 *k^{w}juʔ > *k^{w}jɨʔ > Late OC *k(^{w})juʔ > kjuwX > jiǔ 'nine' (Baxter, 1992: 570)

As shown above, Baxter posits that the word may have undergone rounding dissimilation and rounding assimilation, an independently motivated change, later to restore the original *u. In the absence of a historical record of the pronunciation of 九 prior to the Middle Chinese period, this is certainly possible, but it is not the simplest explanation. The simplest explanation is that 九 originally had the vowel *ɨ and then only underwent rounding assimilation and that there was no process of rounding dissimilation after labialised initials. Note that having an initial and a vowel both specified as [+ round] would violate restrictions on licensing. After rounding assimilation, Baxter does not reconstruct the Middle Chinese initial as a labialised velar, but rather reconstructs a simple velar, which does not violate licensing restrictions. Thus, rounding assimilation may have been blocked by *r in the dialect ancestral to Middle Chinese, but not blocked by *r in the dialect that was ancestral to the poems in the *Guófēng*.

It is interesting to note that this difference between the environment where rounding assimilation may have applied in the language of the *Qièyùn* and that of the *Shījīng* is simply a matter of when it is applied in the derivation of the word. Under Sagart's analysis, the labialised initial must be the root initial since there are no labialised affixes. After the speakers of Old Chinese revalued the constraint that a labialised initial must not precede a round vowel, rounding assimilation could apply before the affixes were added (specifically the infix *-r-) or after the affixes were added, using the theory of lexical phonology. In fact, it is not uncommon for the application of rules to be blocked by the presence of Baxter's medial *r, which generally corresponds to Sagart's infixed *-r-. One expects that such rules posited by Baxter in Appendix A apply after the addition of Sagart's morphemes. The other rules unaffected by the presence of medial *r would apply before affixation.

This analysis predicts that there would be no rhyming words in the earlier sections of the *Shījīng* which begin with a labialised velar initial would be reconstructed with the vowel *ɨ and that rhyming words from later sections of the *Shījīng* which begin with a labialised initial have the vowel *u. This is almost the case. In Baxter's discussion of rhyming practices, he mentions that certain pairs of words may be a stock rhyme, a pair of words that do not phonetically rhyme but are considered a rhyme by some broader community of literature creators. He says that such stock rhymes are useful in extemporaneous oral poetic creation or recreation. Baxter proposes that the following pair of words from the *Shījīng* is an example of a stock rhyme.

95) 懷 *gruj > hwɛj > huái 'to yearn'
96) 歸 *k^{w}jɨj > kjwɨj > guī 'return' (Baxter, 1992: 96)

歸 has a labialised initial and the vowel *ɨ. It rhymes with another word reconstructed with *u,

97) 罍 *C-ruj > lwoj > léi ‘urn-shaped wine vessel’ (Baxter, 1992: 772)

in the *Dàyǎ* (Ode 251). It is difficult to tell if this is a stock rhyme (which is only attested three times in the 306 poems of the *Shījīng*), or if they represent irregular rhyme sequences. In any event, 歸 will not undergo Baxter’s version of rounding assimilation and the rhyme is completely unexpected.

Sagart believes that the labialised velar initials freely occur before all vowels. To the complex facts above, Sagart adds several observations of his own. Sagart notes that rounding assimilation only occurs around Baxter’s *j and that it is strange that labial assimilation occur only around a non-labialised segment (Baxter, 1992: 575, Sagart, 1999: 57). In fact, he shows that Baxter reconstructs labialised velar initial and *ɨ when only *j is present, but reconstructs the vowel *u and simple velars when *j is not reconstructed at all or not by itself (Sagart, 1999: 57). This is clearly seen in the xiéshēng series of words with labialised velar initials followed by *ɨ and simple velar initials with *u.

98) 裘 *g^{w}jɨ >gjuw > qiú ‘furs, any garment of fur’
99) 求 *grju > gjuw > qiú ‘to solicit, to ask for, to beg for’
100) 丘 *k^{w}hjɨ > khjuw > qiū ‘hillock, mound’ (Baxter, 1992: 783)

Both 99) and 100) were used as phonetic elements, and, according to the principle of xiéshēng contact, 100) would be a better phonetic during the Old Chinese period than 99), which is found in 98). Sagart proposes only a dissimilation rule that affected type A and type B syllables unequally in different dialects of Late Old Chinese. All type A and

B syllables dissimilated from an earlier *K^{w}u and *Pu in the language of the *Shījīng*, but only type A syllables dissimilated in the dialect ancestral to Middle Chinese (Sagart, 1999: 58-59).

There are two caveats to this rule as stated above. The first deals with rule ordering. According to Sagart, type A syllables of the form *^{a}Pu(ʔ)(-s) diphthongized before dissimilation occurred. The second is that *^{b}K^{w}ru dissimilates by Middle Chinese in words like 逵 and軌 and the original round vowel is maintained in the *Shījīng*, but *^{b}K^{w}u does not. Sagart's explanation is that *-r- became rounded by the preceding labialised velar, causing it to favor dissimilation (Sagart, 1999: 59). The second part of the explanation seems to be unsatisfactory. Not only does Maddieson not mention labialised r sounds, but also, if the collocation of labialised sounds prompted dissimilation, then why would words like 九 not have dissimilated? However, Sagart's observations about Baxter's reconstructions in 98) through 100) are intriguing.

4. Theories on the Origin of Middle Chinese Type B Finals

4.1 Introduction to the Issue

The Old Chinese origin of Middle Chinese type B finals is the most controversial and confusing debate in Old Chinese reconstructions. These finals caused the palatalization of Old Chinese velars when followed by front vowels. They also caused some, but not all, dental consonants to become palatals. This can be stated confidently because of traditional Chinese phonological analyses like those represented in the *Yùnjìng*. These analyses adopted many aspects of Indian phonology (Branner, p.c., 2004), which was imported to China after Buddhism (Norman, 1994: 397). The Chinese philologists then identified 36 Middle Chinese initials and divided them into different

places and manners of articulation, with labels such as 牙音 yáyīn ‘back-tooth sounds’ for velars and 全清 quánqīng ‘fully clear’ to describe voiceless unaspirated sounds (Baxter, 1992: 45, 57-58).

For type B initials that are not palatal in Middle Chinese, the fǎnqiè spellings of the characters generally use type B words not only for the final, but also the initial. Likewise, the initials of type A words are spelled with type A words, though the type A word can come from any of the traditional divisions. Since a Middle Chinese /k/ in Division I can be spelled with any word that begins with /k/ in Division II or IV but would not be spelled with a /k/-initial word from Division III, it seems likely that the initial /k/ in Middle Chinese type B words was palatalized, unlike the initial /k/ in type A words (Baxter, 1992:69-70). Branner writes that these fǎnqiè create “two hazily distinct sets” (6). The matching of medials in the initial and final fǎnqiè spellers is likely part of the aesthetic of creating fǎnqiè spellings. Being more of an aesthetic principle, there are examples in which the medials of the two fǎnqiè spellers do not match (Branner: 5-6).

However, there is also xiéshēng contact across type A and type B words (Baxter, 1992: 287), as in

101) 良 *C-rjang > ljang > liáng ‘good, fine, desirable’
102) 狼 *C-rang > lang > láng ‘wolf’ (Baxter, 1992: 772)
103) 龍 *C-rjong > ljowng > lóng ‘dragon’
104) 龐 *b-rong > luwng > páng[20] ‘huge, enormous; disorderly, confused’ (Baxter, 1992: 774)

Furthermore, some characters have variant readings, one of which is a type A word, the other of which is a type B word, as in

[20] Baxter cites the Modern Mandarin pronunciation of this character as lóng, likely taken from the recitation tradition of the *Shījīng*. The common Modern Mandarin pronunciation is páng, which Sagart reconstructs as *^{A}b-rong > bæwng > páng. Sagart, Laurent. 1999. *The Roots of Old Chinese*. Philadelphia: John Benjamins Pub. Co. p. 238.

105) 熱 *^{b}t-nip > tsyip > zhé and *^{a}Cə-nip > nep (both) ‘scared stiff’ (Sagart, 1999: 48)

The relationship between the reconstructions of type A and B words and their attested Middle Chinese reflexes is complex and contradictory, allowing for no simple explanation.

Also, while everyone accepts that the Middle Chinese reflex of type B syllables is palatalization of the initial, there is also a question of markedness. According to statistics quoted in Baxter (Baxter, 1992: 288), approximately 52% of all finals in the *Qièyùn* belong to type B. Based on the data provided, there is a 95% probability that type B syllables compose more than 50.7% of the Middle Chinese lexicon, type B finals are statistically more common. Given this, the question becomes why mark the more common case, as is taken in the obvious solution of reconstructing *j or a palatalizing autosegment. Norman makes the intriguing observation that many of the commonly used grammatical words in Old Chinese belong to type B. Based on their frequency, it would be natural for them to have a simpler, unmarked phonology, as in Modern Chinese dialects where the functions words have the form CV(C) (Norman, 1994: 402).

Then again, type B words are only slightly more common and a great number of words will be marked either way, so a palatalizing autosegment may be equally appropriate from a markedness perspective. Furthermore, there seem to be almost twice as many type A words as type B words in Old Chinese (Sagart, 1999: 48). If this is true, then type B words should obviously be marked.

4.2 Previous Reconstructions

Previous reconstructions generally fall into one of two categories. In the first of these, type B finals have their origin in an Old Chinese palatal approximant segment. In

the other, the distinction between the words that developed into type A and B finals was some prosodic, or autosegmental, feature. The second group contains many theories, all quite different from each other, about the origin of the distinction between the palatalizing and non-palatalizing finals, including vowel length contrasts, stress on different moras, and a tense/lax contrast.

As Li writes, if such distinctions are phonemically important, we would not expect to find that they often rhyme (1980: 22). Tense vowels do not rhyme with lax vowels in English. In Classical Arabic, which has contrastive vowel length, short vowels do not rhyme with long vowels, though rhyming in classical odes is more complicated than considering only the rime of word final syllables (Kugle, p.c., 2003)[21]. Japanese poetry is based entirely on counting moras (haiku being the most well known) and rhymes are not considered. Also, Old English poetry is based only on alliteration and a stress system. Each line is composed of a fixed number of stressed syllables, interspersed with a certain number of unstressed syllables, according to poetic conventions (Crist, p.c., 2003). Rhyme is unimportant to Old English poetry.[22] All of these poetic composition practices are language and genre specific, but, if Old Chinese had such contrasts, then the *Shījīng* would be likely be quite a different text. This is only the first objection to these reconstructions, though there are other reasons to be suspicious of each theory (Sagart, 1999: 48-49).

---

[21] Scott Kugle is a professor of Religion at Swarthmore College. Professor Kugle is also a member of the Islamic Studies faculty.

[22] Scanning the first 100 line of a text of Beowulf online at http://www.georgetown.edu/labyrinth/library/oe/texts/a4.1.html, maintained by Irvine, Martin, and Everhart, Deborah. 2002. Beowulf: Georgetown University Medieval Studies Program., there is no rhyming either between the half-lines or the lines themselves.

The theories that explain the differences between type A and B syllables cover almost all of the suprasegmentals introduced in Ladefoged (2001: 231-240). In fact, none of them seem to be correct for Old Chinese, because Old Chinese is believed to be toneless. This belief is based on the parallel developments in Vietnamese. The proposition that syllable final *s became a final /h/ and developed into Middle Chinese qùshēng words is further supported by the pronunciations of these words in 孝義 Xiàoyì, a Jìn dialect. Qùshēng words read in isolation with the Xiàoyì pronunciation have tonal melodies [53^{h}] or [453^{h}]. Xiàoyì also preserves a glottal stop in shǎngshēng words, read with a [$31^{ʔ}12$] melody, pointing to that segment as the origin of the tone (Sagart, 1999: 132-133). Thus, these are no standard suprasegmentals to account for the phonetic distinction of the Old Chinese words. Since Sagart does not believe that any of the previous reconstructions are correct, he takes a neogrammarian stance, simply labeling his reconstructions with a superscript A or B based on the Middle Chinese reflexes (Sagart, 1999: 49).

The use of reconstructed *j is problematic. One of the most glaring issues is that foreign, specifically Sanskrit, words without a palatal approximant are transcribed with characters that are reconstructed with *j. (Examples 106) and 107) (Baxter, 1992: 287), examples 108) and 109) (Norman, 1994: 401)

106) 佛 Middle Chinese bjut > fó ‘Buddha’
107) 鳩摩羅什 Middle Chinese kjuw ma la dzyip > Jiū mó luó shí ‘Kumārajīva’
108) 優婆塞 Middle Chinese ʔjuw ba sok[23] > yōu pó sāi ‘upāsaka’
109) 群那 Middle Chinese gjun na > qún nuó ‘guṇa’

[23] The Middle Chinese transcriptions of examples 108) and 109) have been regularized according to Baxter’s system presented in Chapter 2 of his text. Baxter, William Hubbard. 1992. *A Handbook of Old Chinese Phonology*: Trends in Linguistics. Studies and Monographs: 64. Berlin; New York: Mouton de Gruyter. pp. 45-80.

Special characters with type B readings were created to transcribe the Sanskrit syllables /ka/, /k^ha/, and /ga/ in the Chinese Buddhist scriptures, as in the second syllable of 13).

13) 阿迦貳吒 Middle Chinese ʔa kja nyiH træ > Ē jiā èr zhā for Akaniṣṭha

Since these sounds were not palatalized in the source language and type A velar initials were pure velars under this hypothesis, then, even if there were neither type A nor type B words of the form *ka, *kha, *ga, there would certainly be no need to give them type B readings (Norman, 1994: 401).

Another major concern is that the reconstruction of *j is that it requires the syllable template to allow up to five onset consonants. Though Baxter does not reconstruct a syllable with that many consonants, it is not uncommon for him to reconstruct syllables with four initial consonants. This number seems like incredibly large, especially considering that the Middle Chinese reflexes are generally simple initials (Baxter, 1992: 45). The main reason that the syllable template is so long is because of Baxter's reconstructed medials, *j being the most common. This fact is not a criticism of Baxter's work; it merely suggests there may be another explanation with smaller initial clusters and needing fewer rules to derive Middle Chinese. By reconstructing initial clusters of this length, many of the rules Baxter proposes simply reduce the length of the onset.

Sagart at one time took *j to be an infix like *-r-, which would reduce the number of root initial segments. However, there are xiéshēng series that are composed entirely of type A or entirely of type B words. These are the segregating series that have a capital superscript letter. Based on the presence of these segregating series, which do not seem

to occur with *-r-, Sagart does not believe that *j was an infix but a lexical property (Sagart, 1999: 44-45).

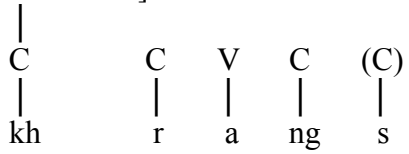
4.3 Palatalization Hypothesis

It would be possible to reduce the number of consonants in the onset by assigning a palatalizing autosegment to type B words. This would not resolve the problem with Sanskrit transcriptions, but would at least begin to reduce the complexity of Old Chinese onsets. Accepting the hypothesis that the presence of a palatalizing autosegment attached to the initial as a lexical property makes the explanation of some rules easier. Labial assimilation was mentioned above a rule that applied across *j. Sagart's objection that rounding assimilation occurred only around an unrounded segment would be removed. Also, *r affects a word's vowel, changing *a to Middle Chinese /æ/ and *e to Middle Chinese /ɛ/ (Baxter, 1992: 580). This change also occurs across *j.

110) 慶 *khrjang(s) > khjængH > qìng 'festivity; to celebrate' (Baxter, 1992: 783)

The result of rewriting *khrjang(s) in terms of a palatalizing autosegment is

111)



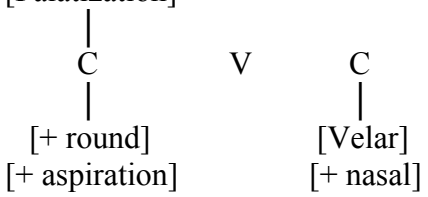
From this point of view, it is obvious that *khrjang(s) would undergo the same vowel changes that type A words would because the environment is the same.

Another phenomenon occurs in words with labialised velar initials that confirms the hypothesis, but raises new questions, as can be seen in the examples below

112) 原 *ng^{w}jan > ngjwon > yuán 'plain, highland' (Baxter, 1992: 567)
113) 葵 *g^{w}jij > gjwij > kuī 'sunflower'
114) 筐 *k^{w}hjang > khjwang > kuāng 'square basket; to regulate' (Baxter, 1992: 771)

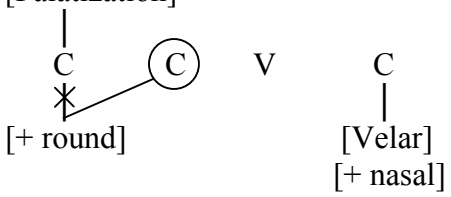
First, it is important to notice the Middle Chinese reflexes of these words. For each one, the labialisation of the initial consonant and the medial *j metathesize, according to a segmental account. If we accept the autosegmental hypothesis, then the autosegmental representation of 114) is

115) [Palatization]



This form was then reanalyzed as having another skeletal position filled by /w/ after the initial consonant (Baxter, 1992: 567). After the addition of the C slot, the reanalysis can be captured as follows

116) [Palatization]



This yields the correct Middle Chinese form /khjwang/ as long as [Velar] is associated with the first skeletal position. This seems to prompt the reanalysis, as explained below, so it is plausible.

The major concern is the phonetic realization of the first skeletal position in 115). When all of the features are filled in, the autosegment [Velar] also attaches to the skeletal position. In this case, the consonant is articulated at the three major places of articulation. In this environment, it is no wonder that the labialisation would be interpreted as a

separate segment. Perhaps, this situation caused the reanalysis, and not analogy with the /w/ after coronal initials after rounding diphthongization, as Baxter proposes (1992: 567).

4.4 Pharyngealization Hypothesis

Jerry Norman proposes a different hypothesis, that the distinction is between syllables with pharyngealization and those without this lexical feature (Norman, 1994: 403). This feature attached to the initial of type A words of divisions I and IV. Norman divides the finals of the *Qièyùn* into three types according the set of initials with which it co-occurs, the crucial test being co-occurrence with the various coronal initials. His type A words occur only with dental initials, type B only with retroflex initials, type C with dental and palatal initials (Norman, 1994: 399). His type C is generally equivalent to type B as used in this paper, and his types A and B together constitute type A. There is some discrepancy between the two system in the placement of words with retroflex initials and type B finals. Norman groups all words with retroflex initials in his type B. He cites other work that shows that *j is not distinctive after retroflex sibilants and writes that the same is true for the stop series (Norman, 1994: 404). This is logical since palatalized retroflex sounds would be incredibly difficult, if not impossible, to produce; there are no instances of these sounds in UPSID.

His theory is that all Chinese finals were palatalized before Middle Chinese unless pharyngealization or retroflexion prevented it (Norman, 1994: 403-404). This may have been the culmination of a long process where plain initials were weakly palatalized even as early as Old Chinese (Norman, 1994: 404). According to this theory, the source of retroflexion need not be a segment, simply the r-like property of retracting the tip of the tongue towards the post-alveolar region (Norman, 1994: 404). This would fit well with

preventing the attachment of a palatalizing autosegment. If pharyngealization, retroflexion, and palatalization were all found on the same tier, then the presence of one of the first two would prevent the addition of palatalization to that tier.

This analysis explains the development of the Middle Chinese sound transcribed as *h*, but was phonetically [ɣ] or [ɦ] from Old Chinese *g. The sound [g], written *g* in Middle Chinese transcriptions, is preserved only before palatalizing finals. Also, Chinese words with velar initials and non-palatalizing finals are borrowed by languages like Hmong, which has a set of uvular stops, as uvulars and not velars. This is likely why Sagart describes this analysis as "the most promising" of all explanations of the difference between type A and type B words (1999: 49). Also, it goes a long way towards explaining the xiéshēng connections between type A and type B words. Presumably, velar and uvular or pharyngealized velar stops would be close enough in place of articulation to fit the definition of xiéshēng contact.

It is interesting to consider how this analysis fits with Sagart's theory about the infix *-r- being responsible for most occurrences of *r in Old Chinese. It seems natural to assume that the initials with pharyngealization were marked throughout Old Chinese. We also know that Sagart's *-r- was reanalyzed as part of the onset before the Middle Chinese period, changing dental initials to retroflex ones. This change would introduce retroflexion as an autosegment, and palatalization would occur after this. Since pharyngealization and retroflexion cannot occur in the same word according to Norman's hypothesis, could words with pharyngealization have distributed object semantics, marked by *-r-?

All words with distributed object semantics would be placed in division II of the *Yùnjìng* based on the effects of *-r-. Pharyngealization is reconstructed only for finals from Divisions I and IV, so, by definition, these words did not have distributed object semantics. Pharyngealization could still be a lexical property of the root. The infix *-r- could be added to a root with a pharyngealizing autosegment attached to the initial as a full segment. When retroflexion became an autosegmental feature and not a separate segment (when dental initials become retroflex), the association between the consonant and pharyngealization could be broken and the initial consonant associated with retroflexion. This analysis accounts for xiéshēng contact across *-r-. Since retroflexion only affects vowels and coronals, I do not believe that this autosegment would attach to labial or velar initials, or necessarily break the association line with the pharyngealizing autosegment. This would preclude having segments specified for all three places of articulation as in

117) 瓜 *k^{w}ra > kwæ > guā ‘melon’
118) 化 *hng^{w}raj(s) > xwæH > huà ‘transform; reform; change’
119) 華 *hwra > xwæ > huā ‘flower’ (Baxter, 1992: 215-217)

Another issue is the fate of the pharyngealizing autosegment after Old Chinese. It does not appear that the Middle Chinese reflexes of any velar consonants, other than *g, in any environments changed their place of articulation to uvular or farther back. Also, *g becomes Middle Chinese /h/ in division II words as well (see Table 3 below). This is obviously not a function of pharyngealization since, according to Norman’s hypothesis, it and the retroflexion that defines division II words are mutually exclusive in Norman’s hypothesis. However, *g could have become [ɣ] before *-r- was reanalyzed as an

autosegment, or, as stated above, a retroflex autosegment may not have attached to grave initials.

A further point of interest is the paucity of uvular nasals in the languages of the world (Pulleyblank, 1977-1978: 199). There is not phonemic uvular nasal in UPSID (Maddieson, 1984: 60). While this nasal would not be phonemic, it would certainly be unlikely, as would any related sounds, such as *ng^{w} and *hng. *ng is a phoneme of Old and Middle Chinese and is still preserved in certain modern dialects. However, in Cantonese, the initial /ŋ/ does not occur before any finals that seem to be the reflexes of Middle Chinese type B syllables, except colloquially, but occurs only before the reflexes of type A finals (Bauer and Benedict, 1997: 486-487). This, as with the history of *g, seems to point to a commonality of velar initials in both of Norman's type A and type B finals. However, these are the only environments to escape palatalization under Norman's theory.

4.5 Comparison of the various reconstructions

Both theories account for the differences in the reflexes of Baxter's *i and *j because they no longer represent the same feature bundle. Also, the sound change lowering high vowels to mid vowels occurs in the environment in which *j is absent (Baxter, 1992: 290). The presence of a palatalizing feature would preserve vowel height, preventing the change. The presence of a pharyngealizing feature could condition the lowering of the following vowel, to accommodate the movement of the tongue root. There is nothing from this change that favors one theory over another.

However, Norman's hypothesis certainly provides a more acceptable answer to the problem of articulation of labialised velar initials. Though labialised velars occur

before non-palatalizing finals, these segments would be marked with pharyngealization in Norman's analysis. The phonetic realization of these sounds is much less complex and much more common than that of palatalized labialised velars. Also, the ratio of labialised uvular stops to simple uvular stops is similar to the ratio of labialised velar stops to plain velar stops (Maddieson, 1984: 213-215). Norman's hypothesis seems to do a better job of accounting for why *r blocks the changes induced by *j because the initial to which the palatalizing autosegment would attach is already linked to a retroflex autosegment. For some words, this would require very different reconstructions from Baxter's if palatalization occurred after *r were removed from the segmental tier.

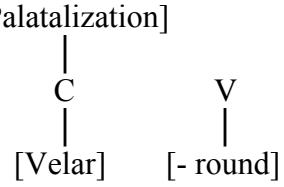
The change Velar palatalization, in which velar initials of type B words generally became palatal affricate before front vowels (there are many exceptions to this rule), is blocked by *r. The derivation of the changes according to the two theories as proposed above is applied to the words below.

120) 支 ~ 枝 *kje > tsye > zhī 'branch'
121) 技 *grjeʔ > gjeX > jì 'ability, talent' (Baxter, 1992: 569)

Norman does not reconstruct any initial velars that palatalize. In the set of words that Norman reconstructs (the first 100 characters of 孟子 Mèngzǐ *Mencius* (Norman, 1994: 405), there are also no characters that Baxter reconstructs with velar obstruents that undergo palatalization. Based solely on Norman's article, it is unclear how the pair 120) and 121) would be treated.

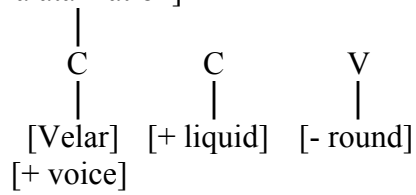
According to the palatalizing hypothesis, 120) and 121) are represented by

122)



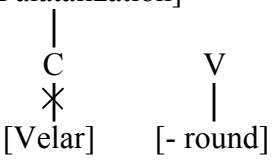
and

123) [Palatalization]



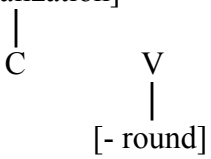
The rule Velar palatalization can be stated as

124) [Palatalization]



Since there are no palatal stops in Middle Chinese, it is safe to assume that when a C slot is attached to the palatalization autosegment and not to a place of articulation feature or the feature [+ continuant] the realization is a palatal affricate. Rule 124 will not apply to 123) because the structure of 123) is not the same as that of the rule. Rule 124 will apply to 122) and the resulting form will be

125) [Palatalization]



which is the correct Middle Chinese form, tsye.

Apart from concerns about the articulation of labialised velar initials attached to palatalizing autosegments and the strange transcriptions of Sanskrit velars, the palatalizing hypothesis seems to be better suited to the data. To be fair, the palatalization hypothesis fits Baxter's reconstructions and Baxter's rules derived from his reconstructions much better because it is most like his reconstruction. If Norman's

analysis results in more fully developed new reconstruction of Old Chinese, then there will be much less bias towards the palatalization hypothesis.

However, there may be a way to combine Baxter, Sagart, and Norman's theories. If Sagart's *-r- and the hypothesis that all type A finals, as defined in here, are marked with pharyngealization are assumed, the velar palatalization can be stated as in 124. The palatalization of initials would have to occur before *-r- was removed from the segmental tier. Once *-r- was reanalyzed as an autosegment, it would attach to the vowel and to the initial unless the initial were specified for a place of articulation or with the feature [+ round]. The retroflexion would change the vowel, creating the Division II finals. For initials with labial or velar place of articulation, the retroflex autosegment would not associate with them and the initials would remain unchanged. For coronal initials, the retroflex autosegment would attach to the initial skeletal position. Any other secondary articulation autosegments, palatalization or pharyngealization, previously associated with this position would no longer be associated with the initial. This would give rise to the retroflex initials of Division II, but all other Division II initials would remain unaffected. This analysis would account for the sporadic occurrence of retroflex initials and type B finals in the *Qièyùn* if one assumes that this choice of the fǎnqiè speller for the final was based on distinctions found in an earlier work when *-r- was still a segment if not an infix. The final would still be considered a type B palatalizing final at the time because it caused the palatalization of the initial segment. The choice of the fǎnqiè speller for the initial was based on Middle Chinese phonology with its increased inventory of initials. This combination would capture the finest distinctions possible, a goal of the authors of the *Qièyùn*.

5. Conclusions

Looking at the underspecification of vowels and consonants offers several insights into Old Chinese syllables and affixes. It predicts that the lack of words ending in *ɨw and *ɨwk is an accidental gap. It also explains why *a and *ɨ were used to transcribe *ə in words with iambic prefixes. It also indicates that *-ŋ should not be considered as a suffix in Old Chinese, so that suffixes (which are secondary in the morphology) do not license places of articulation.

Every hypothesis about the origin of the Middle Chinese palatalizing finals is fraught with problems. Both the pharyngealizing and palatalizing autosegmental hypotheses have merits. The palatalizing theory may be more natural, since the reflex is palatalization, but it leads to consonants articulated all over the mouth. The pharyngealizing theory, however, matches the data about the transcriptions of Sanskrit velar stops and about the common function words. The palatalization hypothesis matches the data, as they stand now, because it is quite similar to Baxter's work. However, the pharyngealizing counterpart of the palatalization hypothesis proposed in this paper seems to incorporate the best parts of both theories.

Table 1: Periodization of the Chinese Language

Dynasty/Historical Era		Dates[24]	Chinese Periodization	Chinese Dialectology[25]
夏	Xià	2070 - 1600 B.C.	Proto-Chinese	
商	Shāng	1600 - 1046 B.C.		
西周	Western Zhōu	1046 - 771 B.C.	Old Chinese	
東周 春秋	Spring and Autumn period	770 - 476 B.C.		
東周 戰國	Warring States period	475 - 221 B.C.		
秦	Qín	221 - 206 B.C.		
西漢	Western Hàn	206 B.C. - A.D. 25	Most Old Chinese sound changes complete (Baxter, 1992: 576)	
東漢	Eastern Hàn	A.D. 25 - 220		
三國	Three Kingdoms period	A.D. 220 - 280		
西晉	Western Jìn	A.D. 265 - 317		
東晉	Eastern Jìn	A.D. 317 - 420		
南北朝	Northern and Southern Dynasties	A.D. 420 - 589	(Early) Middle Chinese (Baxter, 1992: 14)	閩 Mǐn dialects differentiated from Middle Chinese
隋	Suí	A.D. 581 - 618		
唐	Táng	A.D. 618 - 907		粵 Yuè, 晉 Jìn, 客家 Kèjiā dialects differentiated from Late Middle Chinese
五代	The Five Dynasties	A.D. 907 - 960	Late Middle Chinese (Baxter, 1992: 14)	
北宋	Northern Sòng	A.D. 960 - 1127		
南宋	Southern Sòng	A.D. 1127 - 1279		
元	Yuán	A.D. 1206 - 1368		吳 Wú, 贛 Gàn, 湘 Xiāng dialects differentiated from Mandarin Chinese
明	Míng	A.D. 1368 - 1644		
清	Qīng	A.D. 1644 - 1911		
中華民國	Republic of China	A.D. 1912 -		
中華人民共和國	People's Republic of China	A.D. 1949 -		

[24] These dates are taken from the 我国历代纪元表 'My Country's Dynastic History and Historical Era Tables' appendix of Zhōngguó shè huì kē xué yuàn yǔ yán yán jiū suǒ cí diǎn biān jí shì. 中国社会科学院语言研究所词典编辑室. 2002. *Xiàn dài hàn yǔ cí diǎn* 〈现代汉语词典〉. Beijing: Shāng wù yìn shū guǎn 商务印书馆. pp. 1738-1752.

[25] Fuller, Michael A. 1999. *An Introduction to Literary Chinese*. Cambridge, Mass.: Harvard University Asia Center, Harvard University Press. p. 17.

Tables of Type A, Non-Palatalizing Finals

Table 2: Division I Syllables

Character	Gloss[26]	Mandarin	Phonetic[27]	Cantonese[28]	Phonetic[29]	Middle Chinese	Phonetic[30]	Baxter 1992		Sagart 1999	
奔	to run	bēn	pən	ban1	pɐn	pwon	pwʌn	*pun	252		
袍	robe	páo	p^haw	po4	p^how	baw	baw	*bu	779		
冬	winter	dōng	tʊŋ	dung1	to^wŋ	towng	tʌwŋ	*tung	257		
端	end, tip, point	duān	twan	duen1	ty:n	twan	twan	*ton	250		
敦	solid, thick	dūn	tun	dun1	tɵn	twon	twʌn	*tun	250		
投	to throw	tóu	t^hɤw	tau4	t^hɐw	duw	duw	*do	258		
作	to do, to make	zuò	tswɤ	jok3	tsɔ:k	tsak	tsak	*tsak	812		
餐	to eat	cān	ts^han	chaan1	ts^ha:n	tshan	ts^han	*tshan	748	*[a]tshan	197
酢	to toast the host	zuò	tswɤ	jok6	tsɔ:k	dzak	dzak	*dzak	812		
魯	stupid; vulgar; place name	lǔ	lu	lo5	low	luX	lu	*C-raʔ	774		
根	root	gēn	kən	gan1	kɐn	kon	kʌn	*kɨn	251		
廣	wide, broad	guǎng	kwaŋ	gwong2	k^wɔ:ŋ	kwangX	kwaŋ	*k^wangʔ	257		
鼜	big drum	gāo	kaw	go1	kow	kaw	kaw	*ku	257		
害	to injure, to damage; injury harm	haì	χaj	hoi6	hɔ:j	hajH	ɣaj	*ɦikat(s)	761		
逅	to meet unexpectedly	hòu	χɤw	hau6	hɐw	huwH	ɣuw	*gros	763		

[26] The glosses are taken from Baxter and Sagart except for words cited from Baxter 745-811 and Sagart 235-242, which are unglossed by them. For these unglossed words, glosses are from Liang, Shiqiu, and Zhu, Liangzhen. 2000. *Far East English-Chinese Dictionary*. Taipei: Far East Book Co.

[27] The phonetic transcriptions of Mandarin Chinese are based on the author's judgments after his study of Mandarin Chinese.

[28] The Cantonese transcriptions are done according to the Sidney Lau romanization system as presented in Lau, Kam Y. 1999. *Cantonese Phrasebook*. China: Lonely Planet Publications Pty Ltd. pp. 18-20.

[29] The phonetic transcriptions of the Cantonese words are based on Chapter 1 of Bauer, Robert S., and Benedict, Paul K. 1997. *Modern Cantonese Phonology*: Trends in Linguistics. Studies and Monographs: 102. New York: Mouton de Gruyter. pp. 17, 49-92. and compared with audio files on-line at http://www.arts.cuhk.edu.hk/Lexis/Canton/. A Chinese-Cantonese Syllabary: Research Centre for Humanities Computing of the Research Institute for the Humanities, Faculty of Arts, The Chinese University of Hong Kong.

[30] The phonetic transcriptions of Middle Chinese are based on Baxter's presentation in Chapter 2. Baxter, William Hubbard. 1992. *A Handbook of Old Chinese Phonology*: Trends in Linguistics. Studies and Monographs: 64. Berlin; New York: Mouton de Gruyter. pp. 45-80.

Table 3: Division II Syllables

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
白	white	bái	paj	baak6	paːk	bæk	bæk	*brak	746		
卯	cyclical sign (4th earthly branch)	mǎo	maw	maau5	maːw	mæwX	mæw	*mruʔ	260		
謫	to blame, to punish	zhé	ʈʂə	jaak6	tsaːk	trɛk	ʈɛk	*trek	195	*[a]trek	241
濁	muddy	zhuó	ʈʂwɤ	juk6	tso[w]k	dræwk	ɖæwk	*drok	195		
惙	confusion; confused, wild	náo	naw	naau4	naːw	nræw	ɳæw	*nru (?)	778		
捉	grasp, hold in the hand	zhuō	ʈʂwɤ	juk1	tso[w]k	tsræwk	ʈʂæwk	*tsr(j)ok	204		
窗	window	chuāng	ʈʂ[h]waŋ	cheung1	tʃ[h]œːŋ	tsrhæwng	ʈʂ[h]æwŋ	*tshr(j)ong	204		
柴	firewood	chái	ʈʂ[h]aj	chaai4	ts[h]aːj	dzrɛi	ɖʐɛi	*dzr(j)e	204		
山	mountain, hill	shān	ʂan	saan1	saːn	srɛn	ʂɛn	*srjan	785		
關	barrier	guān	kwan	gwaan1	k[w]aːn	kwæn	kwæn	*kron	260	*[a]kron	236
姦	adultery	jiān	cɕjɛn	gaan1	kaːn	kæn	kæn	*kran	261		
衡	to weigh; a beam	héng	χəŋ	hang4	hɐŋ	hæng	ɣæŋ	*grang	762		
瓦	pottery, a tile	wǎ	wa	nga5	ŋaː	ngwæX	ŋwæ	*ng[w]rajʔ	793		

Table 4: Division IV Syllables

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
迷	to bewitch, to enchant	mí	mi	mai4	mɐj	mej	mej	*mij	777		
天	heaven, sky	tiān	t^{h}jɛn	tin1	t^{h}iːn	then	t^{h}en	*hlin	249	*^{a}hlin	240
調	to mediate, to adjust	tiáo	t^{h}jaw	tiu4	t^{h}iːw	dew	dew	*diw	792		
年	year	nián	njɛn	nin4	niːn	nen	nen	*nin	779		
躋	to rise, to ascend	jī	cɕi	jai1	tsɐj	tsej	tsej	*tsɨj	765		
妻	consort, wife	qī	cɕ^{h}i	chai1	ts^{h}ɐj	tshej	ts^{h}ej	*tshɨj	219	*^{a}s-hlɨj	238
齊	uniform, equal; to be equal with	qí	cɕ^{h}i	chai4	ts^{h}ɐj	dzej	dzej	*ɦts(h)ɨj	219	*^{a}s-lɨj	238
先	first	xiān	ɕjɛn	sin1	siːn	sen	sen	*sɨn	205		
黎	numerous; dark	lí	li	lai4	lɐj	lej	lej	*C-rɨj	772		
見	to see	jiàn	cɕjɛn	gin3	kiːn	kenH	ken	*kens	218	*^{a}ken-s	237
褧	unlined hemp garment	jiǒng	cɕjʊŋ	gwing2	k^{w}e^{j}ŋ	khwengX	k^{h}weŋ	*k^{w}hengʔ	257		
現	to appear	xiàn	ɕjɛn	yin6	jiːn	henH	ɣen	*ɦkens	218	*^{a}N-ken-s	240
宴	feast	yàn	jɛn	yin3	jiːn	ʔenH	ʔen	*ʔens	249		

Tables of Type B, Palatalizing Finals (mainly from Division III)

Table 5: Type B Finals after Coronal Initials

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
力	strength, power	lì	li	lik6	leʲk	lik	lik	*C-rjɨk	773	*ᵇCə-rɨk	237
足	foot	zú	tsu	juk1	tsoʷk	tsjowk	tsʲʌwk	*tsjok	203		
雌	female, feminine	cī	tsʰɨ	chi1	tʃʰiː	tshje	tsʰʲe	*tshje	752		
字	Chinese character	zì	tsɨ	ji6	tʃiː	dziH	dzi	*ɦtsjɨ(ʔ)s	811	*ᵇN-tsɨ(ʔ)-s	242
誶	to reproach, to scold	suì	swej	sui6	søy	swijH	swij	*sjuts	790		
象	elephant	xiàng	ɕjaŋ	jeung6	tʃœːŋ	zjangX	zʲaŋ	*zjangʔ	206		
豬	pig	zhū	ʈʂu	jue1	tʃyː	trjo	ʈʌ	*trja	195	*ᵇtra	242
超	to excel	chāo	ʈʂʰaw	chiu1	tʃʰiːw	trhjew	ʈʰew	*thrjaw	195		
箸	chopsticks	zhù	ʈʂu	jue6	tʃyː	drjoH	ɖʌ	*drjaks	195		
女	woman	nǚ	ɲy	nui5	nøy	nrjoX	ɳʌ	*nrjaʔ	195	*ᵇnraʔ	238
莊	dignified, grave	zhuāng	ʈʂwaŋ	jong1	tsɔːŋ	tsrjang	ʈʂaŋ	*tsrjang	204		
初	begin, beginning	chū	ʈʂʰu	choh1	tsʰɔː	tsrhjo	ʈʂʰʌ	*tshrja	204		
崇	to honor, worship, venerate	chóng	ʈʂʰʊŋ	sung4	soʷŋ	dzrjuwng	ɖʐuwŋ	*dzrjung	751		
色	color, countenance	sè	sə	sik1	siːk	srik	ʂik	*srjɨk	204	*ᵇs-rɨk	239
周	circle, place and dynasty name	zhōu	ʈʂɤw	jau1	tsɐw	tsyuw	cçuw	*tjiw	193		
制	to cut out (clothes, etc.); institution; to regulate	zhì	ʈʂʐ	jai3	tsɐj	tsyejH	cçej	*kjets	211		
出	to go out	chū	ʈʂʰu	chut1	tʃʰɵt	tsyhwit	cçʰwit	*thjut	751	*ᵇt-khut	235
摯	to trail, drag	chè	ʈʂʰə			tsyhet	cçʰet	*khjet	211		
成	to achieve, complete	chéng	ʈʂʰəŋ	sing4	seʲŋ	dzyenH	ʥjen	*djans	193		
臣	slave, servant	chén	ʈʂʰən	san4	sɐn	dzyin	ʥjin	*gjin	212		
人	person	rén	ʐən	jan4	tsɐn	nyin	ɲin	*njin	194	*ᵇnin	238
熱	hot	rè	ʐə	jit6	tsiːt	nyet	ɲet	*ngjet	212		
燒	burn	shāo	ʂaw	siu1	siːw	syew	çew	*hngjew	212		

Table 6: Type B Finals after Labial Initials

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Late Middle Chinese	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
悲	sad	bēi	pɐj	bei1	pɐj		pij	pɨj	*prjɨj	190		
非	is not	fēi	fɐj	fei1	fɐj	fji	pjɨj	pʲɨj	*pjɨj	190		
漂	to drift, to float	piāo	pʰjaw	piu1	pʰiːw		phjiew	pʰʲew	*phjew	780		
翻	overturn	fān	fan	faan1	faːn	faan	phjon	pʰʲʌn	*phjan	190		
瓢	gourd	piáo	pʰjaw	piu4	pʰiːw	pɦjiaw	bjiew	bʲew	*bjew	190		
吠	to bark	fèi	fɐj	fai6	fɐj	fɦjiajˋ	bjojH	bʲʌj	*bjots	191		
明	bright	míng	miŋ	ming4	meʲŋ	miajŋ	mjæng	mʲæŋ	*mrjang	191	*ᵇmraŋ	238
亡	be gone	wáng	waŋ	mong4	mɔːŋ	ʋaăng	mjang	mʲaŋ	*mjang	191	*ᵇmaŋ	240

Table 7: Type B Finals after Velar Initials

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
君	a lord, a sovereign	jūn	cɕyn	gwan1	kʷɐn	kjun	kʲun	*kjun	771		
慶	festivity; to celebrate	qìng	cɕʰiŋ	hing3	heʲŋ	khjængH	kʰʲæŋ	*khrjang(s)	783		
群	a group, multitude, flock	qún	cɕʰyn	kwan4	kʰʷɐn	gjun	gʲun	*gjun	784		
玉	jade	yù	y	yuk6	jyːk	ngjowk	ŋʲʌwk	*ng(r)jok	208		
許	approve, allow	xǔ	ɕy	hui2	hɵy	xjoX	xʲʌ	*hng(r)jaʔ	208		
龜	turtle, tortoise	guī	kwej	gwai1	kʷɐj	kwij	kʲwij	*kʷrjɨ	215		
筐	square basket; to regulate	kuāng	kʰwaŋ	kwaang1	kʰʷaːŋ	khjwang	kʰʲwang	*kʷhjang	771	*ᵇk-phaŋ (?)	99
葵	sunflower	kuí	kʰwej	kwai4	kʰʷɐj	gjwij	gʲwij	*gʷjij	771		
原	plain, highland	yuán	ɥɛn	yun4	jyːn	ngjwon	ŋʲwʌn	*ngʷjan	567		

Table 8: Data concerning Sagart's *N- prefix

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
見	to see	jiàn	cɕjɛn	gin3	kiːn	kenH	ken	*kens	218	*[a]ken-s	75
現	to appear	xiàn	ɕjɛn	yin6	jiːn	henH	ɣen	*ɦikens	218	*[a]N-ken-s	75
敗	to defeat	bài	paj	baai6	paːj	pæjH	pæj	*prats	218	*[a]prat-s	75
敗	to be defeated	bài	paj	baai6	paːj	bæjH	bæj	*ɦprats	218	*[a]N-prat-s	75
折	to break, to bend (transitive)	zhé	ʈʂə	jit3	tʃiːt	tsyet	cɕet	*tjat	219	*[b]tat	75
折	to bend (intransitive)	shé	ʂə	jit3	tʃiːt	dzyet	ʥjet	*ɦtjat	219	*[b]N-tat	75
夾	to press between	jiā	cɕja	gaap3	kaːp	kɛp	kɛp	*krep	219	*[a]krep	75
狹	narrow, pressed on both sides	xiá	ɕja	haap6	haːp	hɛp	ɣɛp	*ɦkrep	219	*[a]N-krep	75

Table 9: Data concerning Sagart's *-ŋ suffix

Character	Gloss	Mandarin	Phonetic	Cantonese	Phonetic	Middle Chinese	Phonetic	Baxter 1992		Sagart 1999	
我	we, I	wǒ	wɤ	ngoh5	ŋɔː	ngaX	ŋa	*ngajʔ	795	*[a]ŋajʔ	135
卬	I, we, we on our part, I on my part	áng	aŋ	ngong4	ŋɔːŋ	ngang	ŋaŋ	*ngang	745	*[a]ŋajʔ-ŋ	135
女	woman	nǚ	ɲy	nui5	nɵy	nrjoX	ɳʲʌ	*nrjaʔ	195	*[b]nraʔ	135
孃	mother, originally a term of address	niáng	njaŋ	neung4	nœːŋ	nrjang	ɳʲaŋ			*[b]nraʔ-ŋ	135

Table 10: A Comparison of Norman, Baxter, and Sagart's reconstructions

Character	Norman 1994		Baxter 1992		Sagart 1999		Middle Chinese	Modern Mandarin	Gloss[31]
多	’tay	406	*taj	755	*^{a}t-laj	236	ta	duō	many, much
國	’kwik	406	*k^{w}ɨk	761	*^{a}k-wɨk	236	kwok	guó	country, nation
見	’ken	405	*kens	767	*^{a}ken-s	237	kenH	jiàn	to see
來	’li	405	*C-rɨ(k)	772	*^{a}mə-rɨ	237	loj	lái	to come
力	lik	407	*C-rjɨk	773	*^{b}Cə-rɨk	237	lik	lì	strength; force
鹿	’luk	407	*C-rok	775	*^{a}rok	237	luwk	lù	deer
民	min	407	*mjin	777	*^{b}miŋ	238	mjin	mín	populace, the people
鳥	’tiw	407	*tiwʔ	779	*^{a}tewʔ	238	tewX	niǎo	bird
其	gi	406	*k(r)jɨ	765	*^{b}gɨ	238	ki	qí	3rd person possessive pronoun
人	nin	406	*njin	784	*^{b}nin	238	nyin	rén	a person
身	hyin	406	*hljin	786	*^{b}hlin	239	syin	shēn	body
詩	hyi	407	*stjɨ	787	*^{b}s-tɨ	239	syi	shī	poetry, poem
庶	hya	406	*stjaks	789	*^{b}s-ta(k)-s	239	syoH	shù	numerous, various; general, common; almost, nearly
下	gra	406	*graʔ	796	*^{a}(N-)graʔ	240	hæX	xià	below, under; to descend
以	yi	406	*ljɨʔ	802	*^{b}lɨʔ	240	yiX	yǐ	by means of
之	ti	406	*tjɨ	809	*^{b}tɨ	241	tsyi	zhī	to go to; third person pronoun; particle marking subordination
濯	drakw	407	*lrewk	811	*^{a}lrewk	242	dræwk	zhuó	to wash
子	tsi	405	*tsjɨʔ	811	*^{b}tsɨʔ	242	tsiX	zǐ	son; master; 2nd person pronoun
乘	ding	406	*Ljɨngs	787	*^{b}m-liŋ	235	zyingH	chéng	to ride, to mount
								shèng	ancient carriage, team of four horses

[31] All of these glosses are taken from Liang, Shiqiu, and Zhu, Liangzhen. 2000. *Far East English-Chinese Dictionary*. Taipei: Far East Book Co.

Character	Norman 1994		Baxter 1992		Sagart 1999		Middle Chinese	Modern Mandarin	Gloss
白	brak	407	*brak	746			bæk	bái	white, bright, clear
成	deng	407	*djeng	750			dzyeng	chéng	to achieve; to become
大	’dath	406	*lats	752			dajH	dà	big, large
對	’tuy	406	*k-lups	755			twojH	duì	direct at, to; to respond
奪	’dot	406	*lot	755			dwat	duó	to snatch, to grasp; to rob
夫	pa	406	*p(r)ja	757			pju	fū	man; master; husband
伏	bik	407	*bjɨk	757			bjuwk	fú	to prostrate, to yield; to hide
攻	’kung	407	*kong	759			kuwng	gōng	to attack, to raid; to accuse
顧	’ka	406	*kaʔ(s)	760			kuH	gù	to look at, to gaze
何	’gay	406	*gaj	762			ha	hé	what, how, where, why
後	’gu	406	*ɦ(r)os	763			huwH	hòu	behind, at the back of
乎	’ga	406	*ɦa	763			hu	hū	interrogative particle; exclamatory particle
惠	’gwiy	405	*wets	764			hwejH	huì	to benefit; benefit; benevolent
亟	kik	407	*k(r)jɨk	765			kik	jí	urgently, pressingly
家	kra	406	*kra	766			kæ	jiā	home, household, family
將	tsang	405	*tsjang	767			tsjang	jiāng	verb expressing future action
經	’keng	407	*keng	768			keng	jīng	to pass through; classic book
君	kun	406	*kjun	771			kjun	jūn	sovereign, lord; 2nd person pronoun
樂	’lak	407	*g-rawk	772			lak	lè	happy, glad, joyful; pleasant, agreeable
里	li	405	*C-rjɨʔ	773			liX	lǐ	measure of distance (about ⅓ mile)
利	liy	406	*C-rjits	773			lijH	lì	profit, benefit, gain; to benefit, to serve; sharp
梁	lang	405	*C-rjang	773			ljang	liáng	bridge; beams of a house; place name
靈	’ling	407	*C-reng	774			leng	líng	spirit, soul; mysterious, supernatural, divine

Character	Norman 1994		Baxter 1992		Sagart 1999		Middle Chinese	Modern Mandarin	Gloss
麋	mi	407	*mrjij	777			mij	mí	a kind of deer
千	’tshin	405	*snin	782			tshen	qiān	one thousand; many, numerous
親	tshin	406	*tshjin	783			tshin	qīn	parents, relatives
取	tsho	406	*tshjoʔ	784			tshjuX	qǔ	to select; to summon; to take, to receive
仁	nin	406	*njin	784			nyin	rén	benevolence, mercy, kindness
日	nit	407	*njit	785			nyit	rì	the sun; a day; daytime
上	dang	406	*djangs	786			dzyangH	shàng	above; to ascend
始	hyi	407	*hljɨʔ	787			syiX	shǐ	beginning; to begin, to start
士	dzri	406	*ɦsrjɨʔ	787			dzriX	shì	scholar; officer; a person
叟	’su	405	*srju	790			srjuw	sǒu	an elder, a senior, an old man; venerable sir
臺	’di	407	*lɨ	791			doj	tái	tower, observatory; terrace, elevated platform, stage
王	wang	405	*wjang	794			hjwang	wáng	king, ruler
爲	way	406	*w(r)jaj	794			hjwe	wéi	to do, to act; to serve as; to become; to be
謂	wuy	407	*wjɨts	794			hjwɨjH	wèi	to tell, to say; to name, to call
先	’sin	406	*sɨns	796			senH	xiān	first, foremost
賢	’gin	407	*gin	796			hen	xián	good, virtuous, worthy
焉	wan	406	*ɦ(r)jan	800			hjen	yān	here, there (adverb); interrogative particle; final particle
遺	wuy	406	*ljuj	802			ywij	yí	to lose; lost; to miss; to abandon
已	yi	406	*ljɨʔ	802			yiX	yǐ	already; to cease, to stop; final particle
矣	wi	406	*ɦjɨʔ	802			hiX	yǐ	final particle indicating perfective aspect
營	weng	407	*wjeng	803			yweng	yíng	military barracks; a battalion; to manage, to administer
有	wi	405	*wjɨʔ	804			hjuwX	yǒu	to have, to be present, to exist, there is
囿	wi	407	*wjɨ(k)s	804			hjuwH	yòu	an enclosure for keeping animals; to confine

Character	Norman 1994		Baxter 1992		Sagart 1999		Middle Chinese	Modern Mandarin	Gloss
魚	nga	407	*ng(r)ja	804			ngjo	yú	fish
遠	won	405	*wjanʔ	806			hjwonX	yuǎn	far, distant
云	wun	407	*wjɨn	807			hjun	yún	to say, to speak
在	ʼdzi	407	*dzɨʔ	807			dzojX	zài	at, in, on, up to
沼	taw	406	*tjawʔ	808			tsyewX	zhǎo	lake, pond, pool, marsh
者	ta	406	*tjAʔ	808			tsyæX	zhě	nominalizing particle
征	teng	406	*tjeng	809			tsyeng	zhēng	to conquer, to attack; to levy taxes
不	pi	405			*^{b}puʔ	235	pjuwX	bù	verbal negative
孟	mrang	405			*^{a}mraŋ-s	238	mængH	mèng	family name; eldest brother
文	mun	407			*^{B}mɨn	240	mjun	wén	language; literature, culture
吾	ʼnga	406			*^{a}ŋa	240	ngu	wú	1st person pronoun
亦	yak	405			*^{b}lak	241	yek	yì	also, too
曰	wot	405			*^{b}wat	241	hjwot	yuē	to say; to call, to name

Bibliography

. A Chinese-Cantonese Syllabary: Research Centre for Humanities Computing of the Research Institute for the Humanities, Faculty of Arts, The Chinese University of Hong Kong.

Bauer, Robert S., and Benedict, Paul K. 1997. *Modern Cantonese Phonology*: Trends in Linguistics. Studies and Monographs: 102. New York: Mouton de Gruyter.

Baxter, William Hubbard. 1992. *A Handbook of Old Chinese Phonology*: Trends in Linguistics. Studies and Monographs: 64. New York: Mouton de Gruyter.

Branner, David Prager. The Suí-Táng Tradition of *Fǎnqiè* Phonology. In *History of the Language Sciences: An International Handbook on the Evolution of the Study of Language from the Beginnings to the Present*, eds. Sylvain Auroux, Konrad Koerner, Hans-Josef Niederehe and Kees Versteegh. New York: Walter de Gruyter.

Buckley, Eugene. 1994. *Theoretical Aspects of Kashaya Phonology and Morphology*: Dissertations in Linguistics. Stanford, Calif.: CSLI Publications.

Falla, P. S., Wheeler, Marcus, Unbegaun, Boris Ottokar, and Howlett, Colin. 1993. *The Oxford Russian Dictionary*. New York: Oxford University Press.

Fuller, Michael A. 1999. *An Introduction to Literary Chinese*. Cambridge, Mass.: Harvard University Asia Center, Harvard University Press.

Goldsmith, John A. 1990. *Autosegmental and Metrical Phonology*. Cambridge, Mass.: Blackwell.

Harbaugh, Rick. 1998. *Chinese Characters: A Genealogy and Dictionary*. Taipei: Zhongwen.com.

Irvine, Martin, and Everhart, Deborah. 2002. Beowulf: Georgetown University Medieval Studies Program.

Karlgren, Bernhard. 1949. *The Chinese Language*. New York: The Ronald Press.

Kenstowicz, Michael. 1994. *Phonology in Generative Grammar*. Cambridge, Mass.: Blackwell.

Ladefoged, Peter. 2001. *A Course in Phonetics*. Philadelphia: Harcourt College Publishers.

Lau, Kam Y. 1999. *Cantonese Phrasebook*. China: Lonely Planet Publications Pty Ltd.

Li Fang-kuei 李方桂 (Lǐ Fāngguì). 1980. *Shàng gǔ yīn yán jiū* 〈上古音研究〉. Beijing: Shāng wù yìn shū guǎn 商務印書館.

Liang, Shiqiu, and Zhu, Liangzhen. 2000. *Far East English-Chinese Dictionary*. Taipei: Far East Book Co.

Maddieson, Ian. 1984. *Patterns of Sounds*. New York: Cambridge University Press.

Norman, Jerry. 1988. *Chinese*. New York: Cambridge University Press.

Norman, Jerry. 1994. Pharyngealization in Early Chinese [July - September]. *Journal of the American Oriental Society* 114:397-408.

Pulleyblank, E. G. 1977-1978. The Final Consonants of Old Chinese. *Monumenta Serica* 33:180-206.

Rhew, Hyong. Shijing: Chinese Humanities at Reed College.

Sagart, Laurent. 1999. *The Roots of Old Chinese*. Philadelphia: John Benjamins Pub. Co.

Xǔ Shèn 許慎. 1963. *Shuōwén jiězì* 〈說文解字〉. Beijing: Zhōng huá shū jú 中華書局.

Yao, Tao-chung, Liu, Yuehua, Ge, Liangyan, Chen, Yea-fen, Bi, Nyan-ping, and Wang, Xiaojun. 1997. *Integrated Chinese Level 1, Part 1*. Boston: Cheng & Tsui Co.

Zhōngguó shè huì kē xué yuàn yǔ yán yán jiū suǒ cí diǎn biān jí shì. 中国社会科学院语言研究所词典编辑室. 2002. *Xiàn dài hàn yǔ cí diǎn* <现代汉语词典>. Beijing: Shāng wù yìn shū guǎn 商务印书馆.