Recent discoveries on Old Chinese and pre-Qín documents

Columbia Early China Seminar, 13 November 2015

William H. Baxter and Laurent Sagart

University of Michigan (Ann Arbor) and CNRS/ INaLCO (Paris)

[To the reader: This is an unfinished draft, containing some gaps, which I am sending as is so that people will hopefully have time to read it before the seminar on Friday. Please keep this in mind, and do not cite this version without our permission.—WHB]

The recently published *Old Chinese: a new reconstruction* (Baxter & Sagart 2014b) offers an updated linguistic reconstruction of Old Chinese (approximately, the language of the earliest Chinese classical texts) that takes into account several kinds of evidence that have been given little or no attention in previous reconstructions. In this paper I describe the rationale for the new reconstruction and its main features, and illustrate it with examples. In particular, I will show how the newly discovered texts allow us to improve our reconstruction; and on the other hand, I will argue that our new reconstruction is a sharper tool for analyzing early texts than the traditional frameworks now generally used for this purpose.

1. Background

We use the term "Old Chinese" in a broad sense to refer to any variety of Chinese dating from before the Qín unification of Chinese in 221 BCE. But we also use the term in a narrower sense to refer to the common ancestor of all attested varieties of Chinese, which is the object of our reconstruction. This includes varieties attested in written documents, modern Chinese dialects, and the varieties of Chinese from which words have been borrowed into other languages. In principle, a reconstruction of Old Chinese in this narrow sense should be able to account for the available evidence about all these varieties of Chinese. It appears that this common ancestor was fairly close to the actual language of the earliest Chinese classical texts, from about 1000 BCE.

It was the Swedish scholar Bernhard Karlgren (1889–1978) who made the first systematic attempts to reconstruct Old Chinese ("Archaic Chinese" in his terminology),

including representations of Old Chinese pronunciations using alphabetic phonetic symbols (culminating in his *Grammata serica recensa* of 1957). Other scholars attempted to improve on Karlgren's reconstruction (e.g. Dǒng Tónghé 1948, Li Fang-kuei 1971, Zhèngzhāng Shàngfāng 1987, Starostin 1989, Baxter 1992), but all these reconstructions relied for the most part on three main kinds of evidence:

1. the Middle Chinese phonological system ("Ancient Chinese" in Karlgren's terminology) represented in the *Qièyùn* 《切韻》 rhyme dictionary of 601 CE (and other contemporary sources), which was assumed to be descended from Old Chinese;

2. the rhymes of the *Shījīng* 《詩經》 and (in principle) other pre-Qín texts, and

3. the phonetic elements of the Chinese script (in practice, usually the script that became standardized in Qín and Hàn, rather than the script of the pre-Qín period).

This traditional approach had a number of limitations.

First, for the most part, it ignored evidence from modern varieties of Chinese, and took Middle Chinese as a surrogate for them, under the assumption that Middle Chinese was their ancestor. Actually, Karlgren himself recognized that the dialects of the Mǐn 閩 group had split off from the others before the time of the *Qièyùn* and could not be derived from it;[1] but he never attempted to take them into account in his Old Chinese reconstruction. Starostin (1989) included in his Old Chinese reconstruction some features of Proto-Mǐn as reconstructed by Jerry Norman (1973, 1974, 1981), but not all of them (not, for example, the "softened" initials reconstructed to account for the Northern Mǐn dialects).

Another problem was that many reconstructions were based, not on the *Shījīng* rhymes themselves, but rather on Qīng-dynasty scholars' analysis of them, that is, the system of *yùnbù* 韻部 (Old Chinese rhyme groups) based on the work of Wáng Niànsūn 王念孫 (1744–1832) and Jiāng Yǒugào 江有誥 (d. 1851). Karlgren did depart from the Qīng scholars' analysis in some ways (for example, reconstructing both *-â and *-âr in the traditional 歌部 Gē bù, and both *-u

---

1   "By »Ancient Chinese» ... we designate the language around 600 A. D. codified in the dictionary Ts'ie yün, essentially the dialect of Ch'ang-an in Shensi; during the lapse of the T'ang era it became a kind of Koine, the language spoken by the educated circles in the leading cities and centres all over the country, except the coastal province of Fukien.… [note 2:] … the Koine was sufficiently wide-spread and accepted by a sufficiently large proportion of the population, from the highest officials down to the lower middle class, to have become the ancestor of nearly all the present dialects (except the Min dialects in Fukien and adjacent regions)" (Karlgren 1954:212).

and *-ug in the traditional 侯部 Hóu bù), but Dǒng Tónghé and Li Fang-kuei rejected these innovations and stayed close to the traditional analysis. Starostin (1989) and Baxter (1992) argued that the traditional rhyme groups were not sufficiently fine-grained, and that they overlooked many rhyming distinctions that were present in the data but not recognized in the Qīng scholars' analysis of the data.

In terms of the analysis of the writing system, most work on Old Chinese has relied primarily on the phonetic elements of the script that has been in use since Qín and Hàn times, or on the *Shuōwén jiězì* 說文解字 of 100 CE, although Baxter (1992) occasionally brought brought the pre-Qín script into the argument. This was perhaps understandable at a time when the available corpus of documents in pre-Qín script (mostly oracle bones and bronze inscriptions) was small and restricted in content, but it was clearly anachronistic: the standard script of Qín and Hàn includes a number of rather late characters that do not reflect Old Chinese phonology.[2]

Because of new research and discoveries in recent decades, it is now possible to go beyond the three traditional kinds of evidence and make greater use of three additional kinds of evidence, as we have done in Baxter & Sagart (2014b):

1. pronunciations in modern dialects (especially the Mǐn 閩 dialects) that preserve distinctions lost in the Middle Chinese system;

2. very early Chinese loanwords into neighboring languages (especially of the Vietic, Tai-Kadai = Kra-Dai, and Hmong-Mien = Miáo-Yáo families); and

3. phonetic elements of the pre-Qín script as found in recently discovered pre-Qín documents.

1.1 Notation

A few words are in order about notation.

Although Karlgren believed that Middle Chinese represented in the *Qièyùn* was simply the dialect of Cháng'ān 長安 (modern Xī'ān 西安), subsequent scholarship has shown that this is

---

2 An example is the word {聞} wén < MC *mjun* < OC **mu[n]** 'to hear', which in pre-Qín times was normally written either with a pictogram or with the phonetic element 昏 hūn < MC *xwon* < OC **m̥ˤu[n]** 'dusk, dark'. In Old Chinese, 昏 **m̥ˤu[n]** was pronounced enough like {聞} **mu[n]** to serve as the phonetic element used to write it, but it became less appropriate after OC **m̥ˤ-** was denasalized to **xˤ-**, and 門 mén < MC *mwon* < OC **mˤə[r]** was used instead; by that time, **ə** and **u** had probably merged after labial initials in this context, and **-r** may have merged with **-n** in some dialects. (see Baxter 1992:352–353, Baxter & Sagart 2014b:63). See below for our notation for Middle Chinese and Old Chinese.

most unlikely: Zhōu Zǔmó (1966) argued persuasively that the pronunciations indicated in the *Qièyùn* did not represent the dialect of any single place, but were probably a conflation of two main prestigious dialects or styles of pronunciation: those of Luòyáng and Nánjīng. For that reason, we do not attempt to *reconstruct* a single pronunciation of Middle Chinese based on the early Middle Chinese written sources; instead, our Middle Chinese notation is designed to be a convenient alphabetic representation of the information on pronunciation given in those sources —especially the *Qièyùn* and related rhyme dictionaries, and the *Jīngdiǎn shìwén* 經典釋文 of Lù Démíng (556–627). It is thus a transcription rather than a reconstruction, and for convenience we avoid phonetic symbols and restrict the transcription to ASCII symbols. Thus the plus sign is used instead of the IPA barred-*i* [ɨ]), and *-ae- is used as a mnemonically convenient way to represent a vowel plausibly reconstructible as [æ]. A final -*X* is an arbitrary mark for the shǎngshēng 上聲 tonal category of Middle Chinese, and a final -*H* is the mark for qùshēng 去聲. Syllables with a final -*p, -t,* or -*k* are in the rùshēng 入聲 category; syllables written without any of these marks are in the píngshēng 平聲 category. We generally put Middle Chinese forms in italic type. (For a more detailed description of our Middle Chinese notation, see Baxter & Sagart 2014b:12–20.)

Even though Middle Chinese does not consistently represent the pronunciation of any single dialect, it is far from artificial: essentially all the distinctions it includes would have been found in some variety of Chinese in the early Middle Chinese period. In spite of the importance of other kinds of evidence, Middle Chinese is still perhaps the single most important source of evidence for reconstructing Old Chinese pronunciation, and it is widely used in discussions of pronunciation in traditional commentarial and philological literature, as well as recent literature on newly discovered texts.

Quite apart from the reconstruction of Old Chinese and the Chinese commentarial literature, a familiarity with Middle Chinese also makes it easier to identify and remember Chinese loan words in Japanese, Korean, and Vietnamese. Here are some examples of book titles from these languages, largely composed of vocabulary borrowed during the Middle Chinese period. With Japanese it helps to know that Old Japanese *p* has changed to *h*, and that Middle Chinese final -*ng* is usually reflected by a long vowel:

Japanese:	太平	天国	革命	の	歴史	と	思想
	Taihei	*Tengoku*	*Kakumei*	*no*	*rekishi*	*to*	*shisō*
Middle Chinese:	*thajH-bjaeng*	*then-kwok*	*keak-mjaengH*		*lek-sriX*		*si-sjangX*
	Tàipíng	Tiānguó	gémìng		lìshǐ		sīxiǎng
	'History and thought of the Tàipíng Tiānguó revolution'						

Or consider this Vietnamese book title, consisting entirely of Sino-Vietnamese vocabulary (modifiers follow the noun in Vietnamese):

Vietnamese:	*Lịch*	*sử*	*Triết*	*học*	*phương*	*đông*
Middle Chinese:	*lek*	*sriX*	*trjet*	*haewk*	*pjang*	*tuwng*
Chinese:	lìshǐ		zhéxué		fāng	dōng
	歷史		哲學		方	東
translation:	東方哲學歷史					
	'History of Eastern philosophy'					

And here is a Korean book title, consisting entirely of Sino-Korean vocabulary:

Korean:	한국	현대	문학	연구
	Han'guk	*hyŏndae*	*munhak*	*yŏn'gu*
Middle Chinese:	*han-kwok*	*henH-dojH*	*mjun-haewk*	*ngen-kjuwH*
Chinese:	Hánguó	xiàndài	wénxué	yánjiū
	韓國	現代	文學	研究
	'A study of modern Korean literature'			

For all these reasons, we strongly recommend that any student of premodern Chinese should become familiar with Middle Chinese pronunciations; and the Baxter-Sagart transcription is a convenient way to do so.[3]

---

3   Accordingly, we applaud the decision to include Middle Chinese pronunciations in *A student's dictionary of Classical and Medieval Chinese* (Kroll et al. 2014). The initial edition, however, contains many errors in Middle Chinese transcription. To address this problem, Baxter prepared an extensive list of corrigenda for the Middle Chinese readings, which is now available on the Brill website (http://www.brill.com/products/reference-work/students-dictionary-classical-and-medieval-chinese-0): to download it, on that page, near the bottom under "More information", click on "Downloads - Corrigenda A Students Dictionary of Classical and Medieval Chinese.pdf". The website also says that the corrections "will soon be reflected in the online version of the dictionary, available through chinesereferenceshelf.brillonline.com/chinese-english."

Unlike Middle Chinese, our notation for Old Chinese does represent a linguistic reconstruction, so Old Chinese reconstructed forms are given in the International Phonetic Alphabet, preceded by an asterisk (a convention for indicating that the form is not directly attested). In addition, we use the following conventions:

1. If an element is in parentheses, it means that we do not have enough information to tell whether the element was present or not. For example, we reconstruct 宜 yí < MC *ngje* < ***ŋ(r)aj** 'proper; should': the "**(r)**" just means that, for all we know, there may have been an ***-r-** before the main vowel. It does **not** mean that there is any particular reason to reconstruct an ***-r-**.

2. If an element is in square brackets, "***[X]**", it means that the sound was either ***X** or something else that would have had the same Middle Chinese reflex as ***X**. We use this notation in situations where more than one reconstruction is possible. For example, we reconstruct

(1)   頻 pín < MC *bjin* < ***[b]i[n]** 'climb on all fours'

Middle Chinese *bjin* could reflect OC ***bin**, but the square brackets remind us that other reconstructions are possible for both the syllable onset and the syllable coda (e.g., ***biŋ**, ***m.pin**, ***m.piŋ**), and these possibilities can't be excluded on the basis of current evidence.

3. A hyphen "**-**" represents a morpheme boundary, and a period "**.**" represents a syllable boundary. If we are not confident that elements are separate morphemes, we write a period between them rather than a hyphen. For example, we reconstruct

(2)   秫 shú < *zywit* < ***m.lut** 'glutinous millet'

The ***m** in the presyllable is supported by the Proto-Hmong-Mien form *mblut 'glutinous/sticky' (Ratliff 2010:255), which we believe is an early loan from Chinese. But we write a period after ***m** because we have no evidence that the ***m** is a prefix; it could be part of the root. On the other hand, we reconstruct

(3)   背 bèi < *pwojH* < ***pˤək-s** 'back (n.)'

        背 bèi < *bwojH* < ***m-pˤək-s** 'turn the back on'

Here, we write ***m-** in the verb, with a hyphen, because we can identify the ***m-** as a prefix that derives volitional verbs from nouns (Baxter & Sagart 2014b:55).

4. Angle brackets around ***-r-** indicate that the ***-r-** is an infix (see Baxter & Sagart 2014b:57–58). For example, we reconstruct

(4)   腫 zhǒng < *tsyowngX* < ***toŋʔ** 'swell, swollen'

重 zhòng < *drjowngX* < ***N-t\<r\>oŋʔ** 'heavy (adj.)

重 zhòng < *drjowngH* < ***N-t\<r\>oŋʔ-s** 'weight (n.)'

We believe that all three words come from the root 腫 ***toŋʔ** 'swell, swollen'. Here ***N-** is a prefix that derives intransitive verbs (including adjectives); ***\<r\>** is an infix marking intensiveness; and ***-s** is a suffix whose most common function is to derive nouns from verbs (and adjectives). This example illustrates the fact that considerable derivational morphology can be reconstructed for Old Chinese, though the patterns are often obscured by later sound changes. For example, MC *drjowngX* and *drjowngH* have merged in Mandarin as *zhòng*; similarly, the two readings for 背 bèi, MC *pwojH* and *bwojH*, have also merged.

2. Main features of the Baxter-Sagart Old Chinese reconstruction

The Old Chinese reconstruction in Baxter & Sagart (2014b) includes a number of features inherited from earlier reconstructions, as well as some new ones. The main ones are summarized below.

2.1 *-ʔ and *-s as the source of shǎngshēng and qùshēng (respectively)

As in Baxter (1992) (as well as Starostin 1989 and Zhèngzhāng 2003), we adopt the proposal by Haudricourt (1954a, 1954b) that tones in Chinese developed from lost consonants, parallel to the process of tonal development in Vietnamese. Shǎngshēng words originally had a final glottal stop ***-ʔ**, and qùshēng words had a final ***-s**. The final glottal stop was probably accompanied by a rise in pitch, which became phonologically distinctive when the glottal stop was lost. In qùshēng, a final ***-s** first changed to ***-h**, which was accompanied by a relaxation of the vocal folds, leading to a lower pitch. When the final ***-h** was lost, the lowered pitch became phonologically distinctive.

Final ***-s** was clearly a suffix in many cases, and for the time being, we treat it as a suffix in every case. Its most common function was to derive nouns from verbal roots. Middle Chinese has many pairs of related words where one is in qùshēng (the one that originally had an ***-s** suffix), while the other is in some other tone category. (Sometimes, but not always, the different words are written with the same character or the same phonetic element.) Final voiceless stops were lost before *-s, creating pairs like the following:

	Our reconstructions:	**Karlgren's:**
(5)	度 ***[d]ˤak** > *dak* > duó 'measure (v.)'	"*d'âk"
	度 ***[d]ˤak-s** > *duH* > dù 'measure (n.)'	"*d'âg"
(6)	契 ***kʰˤet** > *khet* > qiè 'cut'	"*k'iat"
	契 ***[kʰ]ˤet-s** > *khejH* > qì 'script notches'	"*k'iad"
(7)	設 ***ŋ̊et** > *syet* > shè 'set up'	"*śi̯at"
	勢 ***ŋ̊et-s** > *syejH* > shì 'circumstances, setting'	"*śi̯ad"
(8)	接 ***[ts][a]p** > tsjep > jiē 'connect'	"*tsi̯ap"
	際 ***[ts][a]p-s** > ***[ts][a]t-s** > *tsjejH* > jì 'connection'	"*tsi̯ad"
(9)	合 ***m-kˤop** > *hop* > hé 'come together; bring together'	"*g'əp"
	會 ***m-kˤop-s** > ***m-kˤot-s** > ***m-kˤwat-s >** *hwajH* > huì 'meeting; have a meeting'	"*g'wâd"

Other changes affecting the words above are (1) denasalization of ***ŋ̊-**, resulting in MC *sy-*, as in (7); (2) assimilation of ***-p-s** to ***-t-s**, as in (8) and (9); and (3) the diphthongization of rounded vowels before dental consonants (including ***-t-s** from earlier ***-p-s**), as in (9). Notice that together with the six-vowel system, recognizing these changes makes it possible to identify morphological patterns in Old Chinese that were less transparent in earlier reconstructions. It also removes the rationale for Karlgren's distinction between voiced and voiceless final stops.

The reconstruction of ***-s** in qùshēng words is supported by many kinds of evidence, especially from Chinese transcriptions of Indic words in Buddhist texts (for examples see Baxter & Sagart 2014b:196–197). Additional support (not mentioned in Baxter & Sagart 2014b) comes from bamboo strips containing fragments of the *Shījīng,* found in 1977 in Fùyáng 阜陽, Ānhuī province, dating from early Western Hàn (Hú Píngshēng & Hán Zìqiáng 1988). Among the fragments is an interesting passage that supports the reconstruction of final ***-s** in qùshēng. The relevant passage is from of Ode 57 (Wèi fēng: Shuò rén 衛風・碩人). It appears that in the expression 瓠犀 *hùxī* 'melon seeds', where a qùshēng word 瓠 hù < *huH* (reconstructed with final ***-s**) is followed by a word with initial ***s-** (犀 xī < *sej*), the final ***-s** of 瓠 *huH* has absorbed the initial ***s-** of 犀 xī, leaving in its place a glottal stop (plain ***ʔ-** or pharyngealized ***ʔˤ-**).

The Máo 毛 version of the poem has the following (stanza 2):

(10)

	齒	如	瓠	犀[4]
	chǐ	rú	hù	xī
MC:	*tsyhiX*	*nyo*	*huH*	*sej*
Hàn:	***t-kʰəʔ** (?)	***na**	***ɦˤwa-s**	***sˤij**
OC:	***t-[k]ʰə(ŋ)ʔ**	***na**	***gʷˤa-s**	***s.lˤəj**
	(or ***t.ŋ̊əʔ**)		or ***ɢʷˤa-s**	
Karlgren:			'Her teeth are like melon seeds'	

By Hàn times, 瓠 ***gʷˤa-s** or ***ɢʷˤa-s** would probably have changed to something like ***ɦˤwa-s**, the ***s.lˤ-** of 犀 xī would probably have changed to ***sˤ-**, and the rhyme ***-əj** would probably have been fronted to ***-ij**. Although 犀 xī normally means 'rhinoceros', the expression 瓠犀 hùxī is traditionally interpreted as 'melon seeds' (as Karlgren translated it), a beautiful woman's teeth being compared to the orderly arrangement of seeds seen when a melon like a canteloupe is cut open.

Where the Máo *Shī* has "瓠犀", the fragment from Fùyáng (S069) has

(11)

	會	[言伊]
MC:	*'jij* or *'ej* (?)	
MC:	*hwajH*	*'jij* or *'ej* (?)
Hàn:	***ɦˤwaj-s**	***ʔij** or ***ʔˤij**

Now for Old Chinese, we reconstruct

(12)   會 huì < *hwajH* < ***m-kˤwat-s** < ***m-kˤot-s** < ***m-kˤop-s** 'meeting; have a meeting', (as in (9) above), but by Hàn times this would probably have undergone a number of changes in pronunciation, resulting in something like ***ɦˤwaj-s**. The character [言伊] is not found in dictionaries as far as I know, but we reconstruct 伊 yī < *'jij* as ***ʔij** 'this'; [言伊] could have been ***ʔij** also, or the corresponding pharyngealized syllable **ʔˤij** (which would regularly become MC *'ej*).

The point here is that it appears that in the Fùyáng version of the text, the sequence 瓠犀 *hù xī*, pronounced something like ***ɦˤwa-s sˤij**, was replaced by ***ɦˤwaj-s ʔij** or ***ɦˤwaj-s ʔˤij**: While the original text (we suppose) had a final ***-s** in 瓠 *hù* and an initial ***s-** in 犀 *xī*, in the Hàn

---

4   In another version of the text, the last character 犀 xī is written as the homonym 棲 xī < *sej.* Xiàng Xī (1986:168) says that the version with 棲 xī is quoted in the commentary on the *Ěryǎ* 爾雅 by Guō Pú 郭璞 (276–324).

version the ***-s** at the end of 會 ***ɦˤwaj-s** appears to have absorbed the ***s-** at the beginning of the next syllable, leaving ***ʔ-** or ***ʔˤ-** behind.[5] Without the hypothesis that qùshēng came from ***-s**, it would be difficult to explain this textual variant.

2.2 Voiceless resonants

As in Li (1971) and Baxter (1992) (and anticipated to some extent by Karlgren and Dǒng Tónghé), we reconstruct voiceless resonants ***m̥-**, ***n̥-**, ***ŋ̊-**, ***r̥-**, ***l̥-** (and their pharyngealized counterparts ***m̥ˤ-**, ***n̥ˤ-**, ***ŋ̊ˤ-**, ***r̥ˤ-**, ***l̥ˤ-**) for Old Chinese; these had disappeared by the time of Middle Chinese. This accounts for xiéshēng relationships like the following:

(13)   芴 ***m̥ˤut** > *xwot* > hū 'careless; confused'; cf.

       勿 ***mut** > *mjut* > wù 'don't'

(14)   灘 ***n̥ˤar** > *than* > tān 'foreshore'.

       難 ***nˤar** > *nan* > nán 'difficult'

(15)   饟 ***n̥aŋ** > *syang* > xiǎng 'bring food to'

       壤 ***naŋʔ** > *nyangX* > rǎng 'cultivated soil'

(16)   犧 ***ŋ̊(r)a[j]** > *xje* > xī 'sacrificial animal'

       我 ***ŋˤajʔ** > *ngaX* > wǒ 'we, I'

(17)   湯 ***l̥ˤaŋ** > *thang* > tāng 'hot liquid'

       陽 ***laŋ** > *yang* > yáng 'bright'

2.3 Lateral initials

We believe that it is Pulleyblank who was responsible for the important discovery that Old Chinese had a set of lateral initials (originally reconstructed in 1962–1963 as "*θ-" and "*δ-", modified in 1973 to "*lh-", "*l-"). The Middle Chinese *l*- does not come from Old Chinese laterals, but from ***r-** or ***rˤ-**. The discovery is important because it made it possible to recognize previously overlooked distinctions in initial consonants. For example, the following pair are homonyms in Middle Chinese:

---

5   It is possible that the **ʔ-** or ***ʔˤ-** at the beginning of the next syllable represents the pharyngealization in the onset of the second syllable, after the ***s-** was removed: phonetically, *ʔˤ- was very likely a pharyngeal fricative [ʕ]. (Roman Jakobson's analysis of the Arabic ʿayin [ʕ] treats it as the pharyngealized counterpart to ʾalif [ʔ]; see Jakobson [1957] 1971.)

(18)　鐸 ***lˤak** > *dak* > duó 'a kind of bell'

　　　度 ***[d]ˤak** > *dak* > duó 'measure (v.)'

However, ***lˤ-** and ***dˤ-** show different patterns of xiéshēng connections: ***lˤ-** has connections with MC *y-* < ***l-**, and MC *sy-* < ***l̥-**, but OC has connections with MC *dzy-* < ***d-** and MC *tsy-* < ***t-**:[6]

(19)　鐸 ***lˤak** > *dak* > duó 'a kind of bell'

　　　譯 ***lAk** > *yek* > yì 'interpret'

　　　釋 ***l̥Ak** > *syek* > shì 'release; dissolve'

(20)　石 ***dAk** > *dzyek* > shí 'stone'

　　　妬 ***tˤak-s** > *tuH* > dù 'jealous'

Traditional phonology recognizes cases where MC **y-** and MC **d-** had a similar origin in Old Chinese; but the statement "喻四古歸定 Yù sì gǔ guī Dìng" '[The initial] 喻四 Yù sì [= MC *y-*] in ancient times goes back to [the initial] 定 Dìng [= MC *d-*]' is an overgeneralization: it is not ALL cases of MC *y-* that are related to all cases of MC *d-*, but only those that come from OC ***l-** and ***lˤ-** respectively. (MC *y-* can also come from OC ***ɢ-**, for example, and MC *d-* can also come from ***dˤ-**).

2.4 The six-vowel system

The reconstruction of Baxter & Sagart (2014b) follows Baxter (1977), Zhèngzhāng Shàngfāng (1987), Starostin (1989), and Baxter (1992) in reconstructing a system of six main vowels for Old Chinese. In our current notation, they are:

(21)　　　　　　　　*i　　　　　　*ə　　　　　　*u

　　　　　　　　　　*e　　　　　　　　　　　　*o

　　　　　　　　　　　　　　　　*a

This system was arrived at independently by Baxter (1977), Starostin (1989), and Zhèngzhāng Shàngfāng (1987),[7] largely based on analyzing the distribution of initials and finals in Middle Chinese (a line of reasoning that began with Jaxontov 1960). The reasoning behind this

---

6　The "*-A-" in these reconstructions is capitalized to call attention to an unsolved problem: OC ***Cak** becomes MC *Cjak* under some conditions and *Cjek* under others. The development is probably conditioned by something in the syllable onset, but for the time being we write ***CAk** for those syllables that become MC *Cjek*. The ***A** is NOT to be understood as a seventh vowel; it means "a case of ***a** that for poorly understood reasons is fronted in Middle Chinese".

7　Although as far as I know, Zhèngzhāng did not publish his ideas until 1987, I understand from Pān Wùyún (p. c.) that he had worked out the main ideas already during the Cultural Revolution (1966–1976).

reconstruction is explained in Baxter (1992:236–257) and Baxter & Sagart (2014b:198–211), and we will not repeat it here. For analyzing pre-Qín and other early Chinese texts, what is important is that the six-vowel system suggests that the set of Old Chinese rhyme groups now generally in use is not sufficiently fine-grained: quite a few of the traditional groups actually contain two or more different rhymes. In other words, many rhyming distinctions were overlooked in the traditional analysis (see below). As a consequence, analyses using the traditional categories are prone to overgeneralization.

We will mention two cases in which our reconstruction can be shown to have greater discriminatory power in analyzing early texts than the traditional rhyme groups. The first is the following passage from the received text of the *Lǎozǐ* (from Baxter & Sagart 2014b:210):

TABLE 5.20 A rhyme sequence from *Lǎozǐ* 39[18]

	from *Lǎozǐ* 39:		MC	OC
1	天無以清將恐裂	tiān wú yǐ QĪNG jiāng kǒng LIÈ	*ljet*	*[r]at
2	地無以寧將恐發	dì wú yǐ NÍNG jiāng kǒng FĀ	*pjot*	*Cə.pat
3	神無以靈將恐歇	shén wú yǐ LÍNG jiāng kǒng XIĒ	*xjot*	*qʰat
4	谷無以盈將恐竭	gǔ wú yǐ YÍNG jiāng kǒng JIÉ	*gjet*	*N-[k](r)at
5	萬物無以生將恐滅	wàn wù wú yǐ SHĒNG jiāng kǒng MIÈ	*mjiet*	*[m]et
6	侯王無以貴高將恐蹶	hóu wáng wú yǐ guì gāo jiāng kǒng JUÉ	*kjwot*	*kʷat

In reconstructions which do not depart from the traditional rhyme groups, all six lines appear to rhyme properly; the rhyme words are all in the traditional 月 Yuè group:

TABLE 5.21 The rhyme words of *Lǎozǐ* 39, in several reconstructions

	MC	rhyme group	Karlgren (1957)	Li (1971)	Guō Xīliáng (1986)	Baxter-Sagart
1	裂 *ljet*	月 Yuè	*li̯at	*ljat	*lĭăt	*[r]at
2	發 *pjot*	月 Yuè	*pi̯wăt	*pjat	*pĭwăt	*Cə.pat
3	歇 *xjot*	月 Yuè	*χi̯ăt	*xjat	*xĭăt	*qʰat
4	竭 *gjet*	月 Yuè	*g'i̯ăt	*gjat	*gĭăt	*N-[k](r)at
5	滅 *mjiet*	月 Yuè	*mi̯at	*mjiat	*mĭăt	*[m]et
6	蹶 *kjwot*	月 Yuè	*ki̯wăt	*kwjat	*kĭwăt	*kʷat

But in a six-vowel reconstruction, line 5 stands out as irregular, because while the others are to

be reconstructed with the rhyme ***-at**, 滅 miè < MC *mjiet* in the fifth line can only be reconstructed with the rhyme ***-et**. And in fact, there are ample independent reasons for believing that the fifth line is a late addition. It is missing in both of the Mǎwángduī silk manuscripts of the *Lǎozǐ*, as also in the Peking University manuscript of *Lǎozǐ* from Western Hàn.[8] Boltz (1985) also argued, on the basis of other versions of the text, and on grounds unrelated to phonology, that the line is a late addition.

The other case is Ode 106, stanza 3 of the Shījīng. The Máo version has this text:

(22) 猗嗟孌兮   孌 *ljwenX*   ***[r]onʔ**

      清揚婉兮   婉 *'jwonX*   ***[ʔ]o[n]ʔ**

      舞則選兮   選 *sjwenX / sjwenH*   ***[s]o[n]ʔ(-s)**

      射則貫兮   貫 *kwanH*   ***kˤon-s**

      四矢反兮   反 *pjonX*   ***Cə.panʔ**

      以禦亂兮   亂 *lwanH*   ***[r]ˤo[n]-s**

Baxter (1992:364) pointed out that all lines except the fifth are to be reconstructed with ***-onʔ** or ***-on(ʔ)-s**, but 反 fǎn < *pjonX* in the fifth line must be reconstructed with ***-anʔ**. However, the *Jīngdiǎn shìwén* says that the Hán 韓 version has 變 biàn (< MC *pjenH* < ***pro[n]-s**) instead. The distinction between ***-a[n]** and ***-[o]n** was probably lost in the Qín and Hàn period, for there are no traces of it in Middle Chinese; it seems likely that 變 biàn < ***pro[n]-s** was the earlier reading, and that the version with 反 fǎn < ***Cə.panʔ** dates from a time when the ***-a[n]** / ***-o[n]** distinction had been lost.

As it happens, this line is quoted in the text "Kǒngzǐ Shī lùn 孔子詩論" in volume 1 of the Shànghǎi Museum bamboo strips (Mǎ Chéngyuán 2001–), which has

(23)

---
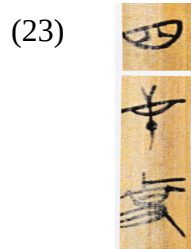
8   We thank Wolfgang Behr for pointing this out.

i.e. "四矢變"; the character represents

(24)　　{弁} biàn < *bjenH* < **C.[b]ro[n]-s** 'cap',

which is often used in pre-Qín documents as a loan character for {變} biàn < **pro[n]-s** 'change' (see Lǐ Jiāhào 1979; see also the discussion below on 卞 *biàn*, a variant of 弁 *biàn* 'cap').

In this case, as in the *Lǎozǐ* passage discussed above, the traditional rhyme categories are not sufficiently fine-grained: both 反 fǎn < **Cə.panʔ** 'return' and 變 biàn < **pro[n]-s** 'change' are in the same traditional rhyme group, 元 Yuán. Nothing in the traditional rhyme groups, or in reconstructions which adhere to the traditional rhyme groups, would indicate that there is anything odd about the Máo text as it stands. But our reconstruction predicts that earlier versions of the text should have 變 biàn < **pro[n]-s**, as confirmed by the "Kǒngzǐ Shī lùn 孔子詩論".

2.5 The role of medial *-r-

Sergei Jaxontov (1960a, 1963) proposed that the distinctive vocalism of division-II finals in Middle Chinese (which are written with either "-*ae*-" or "-*ea*" in our MC notation) was due to an "*-l-" before the vowel in Old Chinese; in more recent versions of the hypothesis, "*-l-" has been replaced by ***-r-**. Pulleyblank (1962–1963) extended the hypothesis to account for certain distinctions in division-III syllables also. Both parts of the hypothesis were adopted in Baxter (1977, 1992) and in other six-vowel reconstructions.

In most cases, we assume that the ***-r-** colors the following vowel and is then lost; the distinctions in vowel color which had once been a predictable concomitant of the preceding ***-r-** then became phonologically distinctive, producing a more complex vowel system in Middle Chinese. For example, we assume that ***-r-** usually changed a following ***-a-** to something like [æ] (for which the "-*ae*-" of our MC notation is mnemonic); when the ***-r-** was lost, then [a] and [æ] became distinct phonemes.[9]

It is the recognition of the role of medial ***-r-** in modifying the vowel system that allows us to reconstruct a simple six-vowel system for Old Chinese, and to assume that Old Chinese rhyming normally required the identity (not just the similarity) of the main vowel and final

---

9　In some cases the ***-r-** remained as a feature of retroflexion in the initial, and did not have the same effect on the following vowel; for example, 張 **C.traŋ** > *trjang* > zhāng 'draw a bow', 莊 **[ts]raŋ** > *tsrjang* > zhuāng 'dignified, grave'.

consonant. For example, all four of the words in (25) are assigned to the traditional 陽 Yáng rhyme group, and rhyme freely with each other in Old Chinese. But in Karlgren's "Archaic Chinese", they are reconstructed with three different vowels: "*â", "a", and "ă". When the effect of medial **-r-** is taken into account, they can all be reconstructed with the same vowel ***a**.

(25)	Our reconstructions:	Karlgren:
	剛 ***kˤaŋ** > *kang* > gāng 'strong; hard' (division I)	"*kâng"
	更 ***kˤraŋ** > *kaeng* > gēng 'change (v.)' (division II)	"*kăng"
	疆 ***kaŋ** > *kjang* > jiāng 'boundary' (division III)	"*ki̯ang"
	京 ***[k]raŋ** > *kjaeng* > jīng 'hill; capital city' (division III)	"*kli̯ăng"

(Karlgren reconstructed "*-l-" in 京 jīng < *kjaeng* < ***[k]raŋ** to account for the fact that it is phonetic in 涼 liáng < *ljang* < ***C.raŋ** 'cool' (Karlgren's "*gli̯ang"), but he did not understand that the medial—his "*-l-", our ***-r-** —also had an effect on the main vowel.)

3. What's new?

Features that are new in Baxter & Sagart (2014b), by comparison with Baxter (1992), include (1) reconstructing pharyngealized initials in type-A syllables (roughly, syllables categorized as division I, II, and IV in Middle Chinese terms), inspired by Norman (1994); (2) reconstructing a syllable coda ***-r**, contrasting with both ***-j** and ***-n**, following Starostin (1989); (3) reconstructing a more complex word structure for Old Chinese, including minor syllables before the main syllable, inspired by Sagart (1999); (4) reconstructing a set of uvular stops contrasting with the velars and labiovelars (following Pān Wùyún 1997, with some modifications); and (5) a greater attention to Old Chinese morphology (inspired by Sagart 1993, 1999).

3.1 Pharyngealization of the syllable onset in type-A syllables

Syllables of Middle Chinese (and by extension, Old Chinese) can be divided into two types, which Pulleyblank (1973) called type A and type B. Type-A syllables are those with division-I, division-II, and division-IV finals; type-B syllables are those with division-III finals. Many initial consonants undergo palatalization in type-B syllables, so Karlgren reconstructed

them with a high front glide "*-i̯-" before the vowel. (Li (1971) and Baxter (1992) adopted the same solution, but substituted "*-j-" for Karlgren's "*-i̯-"). But this solution is unsatisfactory for a number of reasons. For example, it treats type A as the unmarked category, and type B as the marked category. We would expect small function words like 之 zhī < MC *tsyi,* 而 ér < MC *nyi,* and 於 yú < MC *'jo* to be in the unmarked category, but they are all type B.

Various alternatives to Karlgren's "*-i̯-" have been proposed (see Baxter & Sagart 2014b:68–76), but we were ultimately convinced that the solution that had the most explanatory power was that proposed by Norman (1994:403): that type-A syllables were characterized by a "syllabic feature of pharyngealization{. Pharyngealization is a secondary articulation in which the pharynx is constricted by retracting the root of the tongue; the "emphatic" consonants of many varieties of Arabic are pharyngealized. Pharyngealized consonants tend to be resistant to palatalization, and they tend to cause adjacent vowels to become lower (or to stay low if they are already low). This is a good match to the different developments that type-A and type-B syllables underwent in (approximately) the Hàn dynasty. We treat the pharyngealization of type-A syllables as a feature in the syllable onset (since it does not appear to affect rhyme in Old Chinese), and indicate pharyngealization in Old Chinese reconstructions with the IPA symbol for pharyngealization [ˤ] placed after the initial consonant.

3.2 The syllable coda *-r

Starostin (1989) reconstructed syllable codas ***-j** and ***-n**, corresponding to our ***-j** and ***-n**; in Middle Chinese, ***-j** develops to either MC *-j* or *-∅* [zero, as in 歌 ***[k]ˤaj** > *ka* > gē 'sing; song'], while ***-n** remains as MC *-n*. But Starostin proposed that there was a third coda ***-r** whose development differed in different dialects: in most dialects, it merged with ***-n**, but in other dialects it merged with ***-j**. This difference in dialect development led what look like cases where ***-j** and ***-n** seem to be written with the same phonetic element. Sometimes the same word will have two Middle Chinese readings, one as if from ***-n** and the other as if from ***-j**. We accept this proposal:

(26)   鼉 ***[d]ˤar** > (dial.) ***dˤaj** > *da* > tuó 'alligator', also read

      *[d]ˤar > ***dˤan** > *dan*

(27)   單 ***Cə.tˤar** > *tan* > dān 'single, simple'

單于 [OC ***dar** + ***ɢʷa**] > Hàn: ***dar-ɣʷa** > *dzyen-hju* > chányú 'Xiōngnú ruler' (cf. Written Mongolian *daruɣa*'governor', loaned into Persian as *dārūġa* 'governor', see Doerfer 1963–1975, 1.319–1.323).

觶 ***tar** > (dial.) *tsye* > zhī 'ritual vessel'

Starostin did not speculate about the geographical location of the dialects he assumes, but from Hàn-dynasty commentaries and other evidence we have been able to locate the dialect where *-r > *-j to the region in and near the Shāndōng peninsula (Baxter & Sagart 2014b:252–268).

3.3 Uvular stop initials

Pān Wùyún (1997) proposed to reconstruct a series of uvular initials for Old Chinese. We accept his proposals, with some modifications, and reconstruct initial uvular and labiouvular stops *q-, *qʰ-, *ɢ-, *qʷ-, *qʷʰ-, *ɢʷ- and their pharyngealized counterparts. The general development was as follows:

(28)   ***q(ʷ)(ˤ)-** > MC *'-*

     *q(ʷ)ʰ- > MC *x-* (or *sy-* before front vowels)

     *q(ʷ)ʰˤ- > MC *x-*

     *ɢ- > MC *y-*[10]

     *ɢʷ- > MC *hj(w)-* (*yw-* before front vowels)

     *ɢ(ʷ)ˤ- > MC *h(w)-*

Pān assumed that uvulars and velars were phonetically similar enough to be written with the same phonetic elements, but we propose that uvulars changed to velars if there was a preceding consonant (see the next section), accounting for cases like

(29)   影 ***qraŋʔ** > *'jaengX* > yǐng 'shadow (n.)'

     景 ***C.qraŋʔ** > *kjaengX* > jǐng 'bright; image'

which we suspect may be from the same root.

3.4 Presyllables in Old Chinese word structure

In our reconstruction it is possible for the main syllable to be preceded by one or more "preinitial" minor syllables, consisting of a consonant (from a restricted set) with or without the

---

10 Pān Wùyún's proposal was that ***ɢ-** in type-B syllables became MC *hj-*.

vowel *ə. In part this is suggested by early Chinese loanwords into other languages, which appear to have retained the preinitial (Proto-Vietic and Vietic forms were generously supplied by Michel Ferlus):

(30)　牀 **k.dzraŋ** > *dzrjang* > chuáng 'bed', Norman's Proto-Mǐn *dzhoŋ A; cf. Proto-Vietic *k-ɟəːŋ 'bed', Chứt [Sách, Rục] /kəcɨːŋ²/, Maleng [Brô] /kacɨ əŋ/; Maleng [Kha Pong] kəcɨːŋ², Vietnamese *giường* [zɯʌŋ A2]

(31)　賊 *k.dzˤək > dzok > zéi 'injure; murderer, bandit', Norman's Proto-Mǐn *dzhət D; cf. Rục /kəcʌk/ 'bandit, rebel';

Although the **k.** presyllable is not directly attested in Chinese itself, we believe it is responsible for the fact that Mǐn dialects have aspirated initials in (30) and (31): Norman's Proto-Mǐn *dzhoŋ A 'bed' and *dzhət D 'bandit' respectively. The **k.** of 賊 **k.dzˤək** also appears to be preserved in the Lakkia form /kjak 8/ (Baxter & Sagart 2014b:36–37).

　　Reconstructing presyllables also makes it possible to solve a number of puzzles involving phonetic series. Usually, words written with the same phonetic element have initial consonants with the same or a similar position of articulation. There are contacts of velar initials K- with Tsy-type initials that can be ascribed to a palatalization of (nonpharyngealized) velar initials before front vowels:

(32)　稽, 䭫 *kʰˤijʔ > khejX > qǐ 'bow the head to the ground'

　　　旨 *kijʔ > tsyijX > zhǐ 'fine-tasting'

But there are also velar/palatal contacts before nonfront vowels, which are more difficult to explain. In Baxter (1992:213–214), these were written with velar initials in capital letters—as a notation to indicate an unsolved problem: For example, 赤 chì < tsyhek 'red' was reconstructed as "*KHjAk".

　　But the problem is not limited to unpredictable palatalization of velars. For example, it is now generally agreed that the character 九 jiǔ < *kjuwX* 'nine' was the earlier graph for {肘} zhǒu < *trjuwX* 'elbow': a curved arm, sometimes with a mark at the curve to indicate the elbow, as in (33):

(33)

But 'elbow' has the MC initial *tr-*, while 'nine' has *k-*; how can we reconcile these two initials in an Old Chinese reconstruction? In Sagart (1999), it was proposed to handle such cases with a *t- in the presyllable. If Old Chinese words could begin with minor syllables, then we can reconstruct

(34)　　肘 ***t-[k]<r>uʔ** > *trjuwX* > zhǒu 'elbow'[11]

　　　　九 **[k]uʔ* > *kjuwX* > jiǔ 'nine'

We assume that initial ***t-kr-** would have simplified to ***tr-** > MC *tr-*. We write the initial of 'nine' as *[k]- because although 'nine' could have been simply ***kuʔ**, it could also have been something more complex, like ***tə.kuʔ**. With 肘 zhǒu < ***t-[k]<r>uʔ** 'elbow' we can compare Written Tibetan *khru* and Gyarong /təkru/, both 'elbow'; and with 九 jiǔ < **[k]uʔ* (or ? ***tə.kuʔ**) 'nine', we can compare Written Tibetan *dgu* (see Huáng Bùfán 1992, #106, #805). Similarly, to account for the palatal initial in 出 chū < MC *tsyhwit* 'go or come out' vs. 屈 qū < MC *khjut* 'bend, subdue', we reconstruct ***t-kʰut** and ***[kʰ]ut** respectively (following Sagart 1999).

[To add:

　　　　萬 ***t.ma[n]-s** > *mjonH* > wàn '10,000'

(We released the reconstruction "***C.ma[n]-s**" online, but we now believe that Tocharian A *tmaṁ* '10,000' and Turkic *tümän* '10,000', Mongolian *tümen* 'a military unit of 10,000 troops' are ultimately from Chinese, possibly first by way of Tocharian. This also helps explain the initial in

　　　　蠆 ***t.m̥ra[t]-s** > *trhaejH* > chài 'scorpion'

The character for '10,000' is a depiction of a scorpion. We also released a reconstruction "*mə-r̥ˤa[t]-s" online for 'scorpion', but now believe it should be corrected to ***t.m̥ra[t]-s.** See Doerfer (1963–1975: 2.632–642, 4.449). [In 4.449, Doerfer says "Pulleyblank leitet das tü. Wort von chin. **tman* '10,000' ab (heute *wan*), Mitteilung von H. Franke." So it appears that Pulleyblank

3.5 The curious incident of the dog in the nighttime

11　The 寸 cùn 'thumb' on the right side of 肘 is now understood to have been originally a form of "九". The early characers for {寸} cùn 'thumb' and {肘} zhǒu 'elbow' were similar and easily confused; see Jì Xùshēng 2010:348–349. The reconstruction in Sagart (1999) was "*tr-kuʔ.

"Is there any point to which you would wish to draw my attention?"

"To the curious incident of the dog in the night-time."

"The dog did nothing in the night-time."

"That was the curious incident," remarked Sherlock Holmes.

(Doyle 1894:22)

We cite this passage to call attention to the fact that one of the persistent problems in reconstructing Old Chinese is the failure to notice things that do NOT happen. For example, it is easy to notice that certain words which do not rhyme in modern pronunciation did rhyme in the *Shījīng.* But it is much more difficult to notice cases where words which do rhyme now did not rhyme in the *Shījīng.* The words we reconstruct with ***-on** and ***-an** mostly rhyme with each other now; so it is easy to overlook the fact that (for the most part) they do not rhyme in the *Shījīng.* A single example is enough to alert us to the fact that (for example) 有 yǒu < *[ɢ]ʷəʔ and 采 cǎi < *s.r̥ˤəʔ rhyme in the *Shījīng*; but to check that two groups of words do not rhyme in the *Shījīng,* we must check the whole corpus.

A similar situation arises in analyzing phonetic series. For example, in all previous reconstructions of which we are aware, the words 工 gōng < MC *kuwng* and 公 gōng < MC *kuwng* have been reconstructed the same way. But the fact that 公 gōng was originally the phonetic element in 容 róng (see Baxter & Sagart 2014b:28–29, 66, 383n7) led us to reconstruct 公 gōng with a uvular initial, contrasting with 工 gōng:

(35)   容 ***[ɢ](r)oŋ** > *yowng* > róng 'contain'; the phonetic was originally

公 ***C.qˤoŋ** > *kuwng* > gōng 'father; prince'. Cf.

工 ***kˤoŋ** > *kuwng* > gōng 'work'

It was only then that it occurred to us to check whether 公 gōng ***C.qˤoŋ** and 工 gōng ***kˤoŋ** were used differently as phonetic elements. If 公 gōng and 工 gōng really had been homonyms in Old Chinese, we would expect that they could be used interchangeably in the pre-Qín script. But they are not: according to Bái Yúlán (2008:254–257), the words written with 公 gōng and 工 gōng in pre-Qín documents do not overlap at all, a fact that has remained unnoticed until now.

Similarly, in Middle Chinese, although the finals *-ang, -wang, -eng, -weng, -en,* and *-wen*

all occur after *k-* and other velar initials, it was noticing the absence of Middle Chinese syllables like "*twang*", "*tweng*", and "*twen*" that led to the hypotheses of the six-vowel system. Our experience suggests that looking for negative evidence is as important as looking for positive evidence.

4. Old Chinese reconstruction and pre-Qín texts

4.1 The role of the pre-Qín script in Old Chinese reconstruction

There are many cases where the script of recently discovered texts makes it possible to correct or improve previous reconstructions of particular words. For example, before the present corpus of pre-Qín documents was discovered, we had few clues as to how to reconstruct 身 shēn < *syin* 'body'. Karlgren simply projected MC *sy-* (ś-, in his notation) back to Archaic Chinese as "*ś-*". In our reconstruction, MC *sy-* has many different OC sources (including ***l̥-**, ***n̥-**, ***s.t-**, ***s.tʰ-**; ***ŋ̊-** and ***qʰ-** before front vowels), but it was unclear which to reconstruct for 身 shēn. There was a similar problem with 設 shè < *syet*. But we now have examples of {仁} rén 'kind' written with 身 shēn as phonetic; this leads us to reconstruct

(36)   仁 ***niŋ** > *nyin* > rén 'kind'

        身 ***n̥i[ŋ]** > *syin* > shēn 'body; self'

Similarly, because of Qiú Xīguī's paper showing that the graph ancestral to 埶 yì 'plant' was used to write {設} shè 'set up', we can now reconstruct

(37)   埶 ***ŋet-s** > *ngjiejH* > yì 'to plant'

        設 ***ŋ̊et** > *syet* > shè 'set up (v.)'

        勢 ***ŋ̊et-s** > *syejH* > shì 'circumstances, setting'

Moreover, this reconstruction considerably clarifies the etymology and the meaning of the important term 勢 shì < ***ŋ̊et-s**, whose multiple possible translations have long puzzled English-speaking scholars (at least). It is a noun derived by the ***-s** suffix from 設 shè < ***ŋ̊et** 'set up (v.)'. Its various uses can be understood to derive from a meaning like 'the way things are set up (either by Heaven or by human rulers)', hence 'the way things happen naturally, the natural course of events' (if the agent is Heaven), 'circumstances' (which humans cannot control but can learn to manipulate to their advantage); 'power' (deriving from the way human rulers have set

things up, through institutions and laws), and doubtless many other senses as well. For example, we believe that our reconstruction allows us to make better sense of this sentence from *Hán Fēi zi: Nán shì* 《韓非子・難勢》:

吾所為言勢 (***ŋ̊et-s**) 者，言人之所設 (***ŋ̊et**) 也

'The 'set-up' (***ŋ̊et-s**) of which I am speaking refers to what is set up (***ŋ̊et**) by men.'

To a contemporary reader or listener, it would have been clear that two different froms of the same root ***ŋ̊et** 'to set up' were being used.[12]

4.2 A critique of traditional Chinese phonology: Middle Chinese

The examples in 4.1 have shown how evidence from pre-Qín documents is helpful in reconstructing Old Chinese. But it is also the case that a good reconstruction of Old Chinese should be useful in reading the pre-Qín documents, because judgments about what word a character represents are often based on assumptions about similarities in pronunciation in Old Chinese times.

Most Chinese scholars analyzing newly discovered texts do not use any alphabetic notation for either Middle Chinese or Old Chinese: instead they a set of named categories adapted from traditional Chinese phonology. To specify a Middle Chinese pronunciation completely in the traditional way, one gives a set of categories to which the syllable in question belonged. One convenient reference supplying such information is Dīng Shēngshù & Lǐ Róng (1981). Words are arranged according to Mandarin pronunciation, and Middle Chinese pronunciations are specified by the fǎnqiè spellings of the *Guǎngyùn*, and in terms of traditional categories. For example:

(38)	example	fǎnqiè	traditional specification
	新 xīn < MC *sin*	息林切	臻開三平真心
		(*sik + lin = sin*)	
	全 quán < MC *dzjwen*	息鄰切	山合三平仙從
		(*dzit + ywen = dzjwen*)	

But it is not easy to become proficient in the use of these traditional descriptions. Here is what

---

12 We think this gives a better understanding of the text than A. C. Graham's translation of 勢 shì as 'power base': "When I speak of the power-base it is of something instituted by man" (1989:280).

the traditional terms in these two cases mean:

for 新 xīn < MC *sin:*	
臻 zhēn	= 臻攝 Zhēn shè. The *shè* 攝 are a set of 16 broad categories for classifying syllables according to their Middle Chinese rhymes. The 臻 Zhēn shè includes MC syllables with high vowels and a final *-n* or *-t*.
開 kāi	= 開口 kāikǒu 'open-mouth' means that in Middle Chinese there is no *-w-* before the final *-in*.
三 sān	= 三等 sānděng. This means that the syllable belongs to the 三等 sānděng ('grade 3' or 'division 3') category; i.e. its final *-in* is one that occurs in the third row of rhyme tables like the *Yùnjìng* 韻鏡. The děng are a way of categorizing Middle Chinese finals according to the way they are treated in the rhyme tables. (The term is somewhat confusing in this case, because 新 xīn < *sin* is actually in the fourth row of the rhyme tables, since initial *s-* can only appear in the first and fourth rows of the table; but most words with the final *-in* are indeed placed in the third row.)
平 píng	= 平聲 píngshēng, the MC tonal category to which 新 xīn < MC *sin* belongs. The four tones of Middle Chinese (平 píng, 上 shǎng, 去 qù, and 入 rù) should not be confused with the four tones of Mandarin, to which they correspond only indirectly. MC píngshēng words regularly go to tone 1 or tone 2 in Mandarin, depending on whether the MC initial consonant is voiceless or voiced.
真 zhēn	= 真 zhēn < MC *tsyin*, the *Guǎngyùn* rhyme (韻 yùn) to which 新 xīn < MC *sin* belongs.
心 xīn	= 心母 xīn mǔ, the traditional name (心 MC sim) for the MC initial *s-*.

for 全 quán < MC *dzjwen:*	
山 shān	= 山攝 Shān shè, the shè to which 全 MC *dzjwen* belongs; it includes syllables with nonhigh vowels and final *-n* or *-t*.
合 hé	= 合口 hékǒu means that there is a *-w-* before the main vowel in 全 MC *dzjwen*.
三 sān	= 三等 sānděng, as above
平 píng	= 平聲 píngshēng, as above
仙 xiān	= 仙 xiān < MC *sjen*, the Guǎngyùn rhyme to which 全 MC *dzjwen* belongs.

| 從 cóng | = 從母 cóng mǔ, the traditional name (從 MC *dzjowng*) for the MC initial *dz-*. |

The practice of specifying Middle Chinese pronunciations in this way has the weight of tradition behind it, but it is rather difficult to master, and inconvenient to use even when mastered. Our transcription "*sin*" for 新 xīn gives exactly the same information as "臻開三平真心", but it is much easier to learn and remember, and easier to connect to modern pronunciations and Sinitic vocabulary in other languages:

Mandarin *xīn* ([ɕin 55], where [ɕ] comes from earlier [s] before [i] or [y])

Japanese *shin* (where *sh-* is from earlier *s-* before *-i-*)

Korean *sin*

Vietnamese *tân* (derivable from MC *sin* by regular rules)

The difficulty and inconvenience of the system is reflected in the fact that even prestigious scholars sometimes make mistakes when giving Middle Chinese pronunciations in traditional terms. (We omit examples, but be assured that they exist.) Notice also that the traditional terminology does not reveal that both 新 MC *sin* and 全 MC *dzjwen* end with the same consonant -*n*; that is a fact about the 臻 Zhēn shè and the 山 Shān shè that has to be learned separately.

In practice, Middle Chinese pronunciations are rarely specified fully in phonological arguments about pre-Qín texts; the usual practice is to give two terms: the traditional name for the Middle Chinese initial consonant and the traditional name for the Old Chinese rhyme group. The traditional names for MC initial consonants are given in the table below, labeled with our MC transcription; the last column gives their typical reflexes in standard Mandarin.

Initials (syllable onsets) of Middle Chinese in the Baxter/Sagart notation

Píng 平 = píngshēng 平生 (tone A); zè 仄 = shǎngshēng 上聲, qùshēng 去聲, or rùshēng 入聲 (tones B, C, and D), Karlgren's "Ancient Chinese" reconstruction is given (under "K") for reference. "L" in the "Mandarin reflex" column indicates the reflex in literary readings (wéndú 文讀).

	B&S	K	labial	labiodental	Mandarin reflex
labials: P-[13]					
1	*p-*	*p-*	幫 Bāng (MC **pang**)	非 Fēi (MC **pj+j**)	**b-** or **f-**
2	*ph-*	*p'-*	滂 Pāng (MC **phang**)	敷 Fū (MC **phju**)	**p-** or **f-**
3	*b-*	*b'-*	並 Bìng (MC **bengX**)	奉 Fèng (MC **bjowngX**)	平: **p-** or **f-** (tone 2); 仄: **b-** or **f-**
4	*m-*	*m-*	明 Míng (MC **mjaeng**)	微 Wēi (MC **mj+j**)	**m-** or **w-**
dentals: T-					
5	*t-*	*t-*	端 Duān (MC **twan**)		**d-**
6	*th-*	*t'-*	透 Tòu (MC **thuwH**)		**t-**
7	*d-*	*d'-*	定 Dìng (MC **dengH**)		平: **t-** (tone 2); 仄: **d-**
8	*n-*	*n-*	泥 Ní (MC **nej**)		**n-**
9	*l-*	*l-*	來 Lái (MC **loj**)		**l-**
retroflex stops: Tr-					
10	*tr-*	*t-*	知 Zhī (MC **trje**)		**zh-**
11	*trh-*	*t'-*	徹 Chè (MC **trhjet**)		**ch-**
12	*dr-*	*d*	澄 Chéng (MC **dring**)		平: **ch-** (tone 2); 仄: **zh-**
13	*nr-*[14]	*ń-*	娘 Niáng (MC **nrjang**)		**n-**
dental sibilants: Ts-					

13 The *fǎnqiè* of the *Qièyùn* and *Guǎngyùn* show no evidence of the process which changed the labial initials of some syllables to labiodentals in most dialects. (In such cases, MC **p-**, **ph-**, and **b-** become [f] in Mandarin, and **m-** becomes [w].) However, the distinction is made in the traditional thirty-six initials (sānshíliù zìmǔ 三十六字母), so it is conventional to use the names 非 Fēi, 敷 Fū, 奉 Fèng, and 微 Wēi for those cases of MC **p-**, **ph-**, **b-**, and **m-** which eventually became labiodentals.

14 Y. R. Chao (1941:XXX) pointed out that MC **n-** and **nr-**, though usually distinguished in fǎnqiè spellings, are in complementary distribution in the *Guǎngyùn*, and therefore would not need to be distinguished in a phonemic analysis. Following this, Dīng and Lǐ (1981a, 1981b) use the name 泥 Ní (MC **nej**) for both. However, they are distinguished in more traditional treatments, and in our notation, so we following tradition in using 泥 Ní (MC **nej**) for MC **n-** and 娘 Niáng for MC **nr-**.

No.			conventional name	traditional name	
14	**ts-**	ts-	精 Jīng (MC *tsjeng*)		**z-** (**j-** before -i- or -ü-)
15	**tsh-**	ts'-	清 Qīng (MC *tshjeng*)		**c-** (**q-** before -i- or -ü-)
16	**dz-**	dz'-	從 Cóng (MC *dzjowng*)		平: **c-** / **q-** (tone 2); 仄: **z-** / **j-**
17	**s-**	s-	心 Xīn (MC *sim*)		**s-** / **x-**
18	**z-**	z-	邪 Xié (MC *zjae*)		平: **s-** ~ **c-** / **x-** ~ **q-** (tone 2); 仄: **s-** / **x-**
retroflex sibilants: *Tsr-* [15]					
19	**tsr-**	tṣ-	莊 Zhuāng (MC *tsrjang*)	照二 Zhào èr	**zh-** (L **z-**)
20	**tsrh-**	tṣ'-	初 Chū (MC *tsrhjo*)	穿二 Chuān èr	**ch-** (L **c-**)
21	**dzr-**	dẓ'-	崇 Chóng (MC *dzrjuwng*)	床二 Chuáng èr	平: **ch-** ~ L **c-** (?) (tone 2); 仄: **zh-** ~ **sh-** (L **z-**)
22	**sr-**	ṣ-	生 Shēng (MC *sraeng* < *srjaeng*)	審二 Shěn èr	**sh-** (L **s-**)
23	**zr-**	—	[俟 Sì (MC *zriX*)][16]	[禪二 Shàn èr]	**ch-** (tone 2; one word only: 漦 MC *zri* > chí 'dragon spittle'); 仄: **s-** (L)
palatals: *Tsy-*					
24	**tsy-**	tś-	章 Zhāng (MC *tsyang*)	照三 Zhào sān	**zh-**
25	**tsyh-**	tś'-	昌 Chāng (MC *tsyhang*)	穿三 Chuān sān	**ch-**
26	**dzy-**[17]	ź-	禪 Shàn (MC *dzyenH*)	禪三 Shàn sān	平: **ch-** ~ **sh-** (tone 2); 仄: **sh-** ~ **zh-**

15 The fǎnqiè of the *Qièyùn* and *Guǎngyùn* clearly distinguish the *Tsr-* initials (retroflex sibilant initials **tsr-**, **tsrh-**, **dzr-**, **sr-**, and **zr-**) from the *Tsy-* initials (the palatals **tsy-**, **tsyh-**, **dzy-**, **sy-**, and **zy-**); this distinction was discovered by Chén Lǐ 陳澧 (18XX). But in Late Middle Chinese, the two sets were apparently in complementary distribution and had merged phonemically: the traditional thirty-six initials treat them as a single set, with the names 照 Zhào, 穿 Chuān, 床 Chuáng, 審 Shěn, and 禪 Shàn. Since in the rhyme tables the *Tsr-* set were put in division II (èrděng 二等) and the *Tsy-* set in division III (sānděng 三等), one way to distinguish the two sets terminologically is to use the traditional initial names followed by 二 er for *Tsr-* or 三 sān for *Tsy-*. However, it has become more common to use the separate names used by Dīng and Lǐ (1981b), given in the left-hand column of this section.

16 MC *dzr-* and *zr-* are confused in the fǎnqiè spellings of the *Guǎngyùn*, and Dīng Shēngxù and Lǐ Róng (1981a) use the same name 崇 Chóng (MC *dzrjuwng*) for both. But the distinction is implied in the rhyme tables such as the *Yùnjìng* 韻鏡, where 俟 MC *zriX* and 士 MC *dzriX* are put in different positions (俟 MC *zriX* in the same column as MC *z-*, and 士 MC *dzriX* in the same column as MC *dz-*). The distinction is also supported (as Dīng and Lǐ themselves point out, 1981a:37) by the fǎnqiè spellings of the *Kānmiù bǔquē Qièyùn* 刊謬補缺切韻 by Wáng Rénxù 王仁昫, an almost complete manuscript of which was discovered in the late 1940s. We propose the name 俟 Sì for MC *zr-* (since this is the only common word which has this initial).

#					
27	sy-	s-	書 Shū (MC syo)	審三 Shěn sān	sh-
28	zy-[5]	dz'-	船 Chuán (MC zywen)	床三 Chuáng sān	平: sh- ~ ch-; 仄: sh-
29	ny-	ńź-	日 Rì (MC nyit)		r-
30	y-[18]	ï	以 Yǐ (MC yiX)	喻四 Yù sì	y- (~ w-) (NB: earlier Mandarin "yóng" > róng, "yuì" > ruì: 容 MC yowng > yóng > róng, 銳 MC ywejH > yuì >ruì)

velars and laryngeals: K-

#					
31	k-	k-	見 Jiàn (MC kenH)		g- (j- before -i-, -ü-)
32	kh-	k'-	溪 Xī (MC khej)		k- (q- before -i-, -ü-)
33	g-	g'-	群 Qún (MC gjun)		平: k- / q- (tone 2); 仄: g- / j- (sometimes irregularly h- / x-)
34	ng-	ng-	疑 Yí (MC ngi)		Ø- (usually), n- (occasionally): 牛 MC ngjuw > niú, 擬 MC ngiX > nǐ)
35	'-	-	影 Yǐng (MC 'jaengX)		Ø-
36	x-	x-	曉 Xiǎo (MC xewX)		h- (x- before -i-, -ü-)
37	h-	ɣ-	匣 Xiá (MC haep) (NB: MC h- is voiced, IPA [ɦ] or [ɣ])		h- (x- before -i-, -ü-)
38	hj-[6]	ji-	云 Yún (MC hjun)	喻三 Yù sān	Ø-

---

17 The positions of MC *dzy-* and *zy-* are reversed in the rhyme tables: MC *dzy-* is put in the same column as MC *z-* and *zr-*, while MC *zy-* is put in the same column as MC *dz-* and *dzr-*. This probably reflects the fact that MC *dzy-* and *zy-* were not clearly distinguished in most dialects (including Mandarin). Karlgren did not understand this, so in his Ancient Chinese MC *dzy-* is written as "ź-", and MC *zy-* is written as MC "dź'-".

18 Although MC *hj-* and MC *y-* are clearly distinguished in the rhyme tables, MC *hj-* and MC *y-* are in the same column, with MC *hj-* in division III (sāndĕng 三等) and MC *y-* in division IV (sìdĕng 四等). As with the *Tsr-* and *Tsy-* initials, one way to distinguish them is to call them 喻三 Yù sān and 喻四 Yù sì respectively; but it is probably more common to follow Dīng and Lǐ (1981b) and use 云 Yún for MC *hj-* and 以 Yǐ for MC *y-*. [MC *h-* and *hj-* are (nearly) in complementary distribution, so our notation uses the same symbol *h-* for both; but they are distinguished in the traditional terminology, so we distinguish them in lists like this as "*h-*" and "*hj-*", even though the latter is not always written as "*hj-*" in our transcription. Our "*h-*" is the traditional initial 云 Yún "*hj-*" in syllables where *h-* is either immediately followed by "-*j*-", or followed by "-*i*-" later in the syllable. Thus in traditional terms, 位 wèi < MC *hwijH* is treated as the initial 云 Yún *hj-* followed by the final -*wij*, in qùshēng (-H), even though we do not write it as "*hj-*".

Although the philological tradition that led to the terminology above is justly considered a major intellectual achievement, the fact remains that as a notation, the traditional categories are inconvenient and confusing: using them to describe Middle Chinese pronunciation is like trying to do arithmetic with roman numerals. At the same time, since the traditional terminology is still widely used in Chinese studies of early texts, the student of early texts cannot simply ignore it. But we believe that approaching the traditional terminology through our Middle Chinese transcription is a good way to become familiar with it.

4.3 A critique of traditional Chinese phonology: Old Chinese

As noted above, it is common in discussions of pre-Qín texts to pay attention to the Middle Chinese initials and the Old Chinese rhyme groups of the words being discussed. The use of the names of Middle Chinese initials in discussing Old Chinese syllable onsets is a serious anachronism, and leads to many overgeneralizations and spurious suggestions. In effect, the traditional practice presupposes that the same terms can be used for initial consonants in both Middle and Old Chinese; and it is usually assumed that the initial consonants of Old Chinese were fewer than those of Middle Chinese. It is true that Middle Chinese had initial consonants that Old Chinese did not; but the traditional procedure in effect assumes that if words had the same initial in Middle Chinese, they also had the same initial in Old Chinese. As we have seen with MC *sy-*, this is a serious misunderstanding: MC *sy-* has many different sources, as can be seen by analyzing the phonetic elements of characters; to treat them all as if they were a single initial in Old Chinese ignores a great deal of the available evidence. And slogans like "喻四古歸定 Yù sì gǔ guī Dìng" 'MC *y-* in ancient times went back to MC *d-*' overlook the fact that both MC *y-* and MC *d-* have more than one Old Chinese source. The contacts between MC *y-* and *d-* are not due to the fact that *y-* and *d-* are "音近 yīn jìn" 'similar in sound', or that they have a "密切的關係 mìqiè de guānxi" 'a close relationship'. It is OC **l-*, only one of the sources of MC *y-*, and OC **lˤ-*, only one of the sources of MC *d-*, that are similar in sound. Using the same terms for Middle Chinese consonants and Old Chinese consonants inevitably leads to confusion and overgeneralization.

As for the analysis of rhymes, the traditional approach is to use names for Old Chinese rhyme groups that were identified in traditional phonology. This system has some of the same

disadvantages as the use of a non-alphabetic notation for Middle Chinese: its abstract nature makes it difficult to get clear what the specific similarities and differences among the rhymes may have been. But our main criticism of the traditional rhyme categories is that they overlook a great many rhyme distinctions that would be helpful in making decisions about unfamiliar loan characters and phonetic compounds.

There are several competing terminologies for the Old Chinese rhyme groups, but the most commonly used version is probably that given in Wáng Lì (1999:677–688), summarized in the table below. The table includes Wáng Lì's own reconstructions for the rhymes (in quotation marks, to avoid confusing them with ours).

	陰聲 yīnshēng (zero or vocalic coda)		入聲 rùshēng (voiceless stop coda)		陽聲 yángshēng (nasal coda)	
甲類 jiǎ lèi	1. 之 Zhī	"*-ə"	2. 職 Zhí	"*-ək"	3. 蒸 Zhēng	"*-əng"
	4. 幽 Yōu	"*-u"	5. 覺 Jué	"*-uk"	6. 冬 Dōng	"[*-ung]"
	7. 宵 Xiāo	"*-ô"	8. 藥 Yào	"*-ôk"		
	9. 侯 Hóu	"*-o"	10. 屋 Wū	"*-ok"	11. 東 Dōng	"*-ong"
	12. 魚 Yú	"*-a"	13. 鐸 Duó	"*-ak"	14. 陽 Yáng	"*-ang"
	15. 支 Zhī	"*-e"	16. 錫 Xī	"*-ek"	17. 耕 Gēng	"*-eng"
乙類 yǐ lèi	18. 脂 Zhī	"*-ei"	19. 質 Zhì	"*-et"	20. 真 Zhēn	"*-en"
	21. 微 Wēi	"*-əi"	22. 物 Wù	"*-ət"	23. 文 Wén	"*-ən"
	24. 歌 Gē	"*-ai"	25. 月 Yuè	"*-at"	26. 元 Yuán	"*-an"
丙類 bǐng lèi			27. 緝 Qī	"*-əp"	28. 侵 Qīn	"*-əm
			29. 葉 Yè	"*-ap"	30. 談 Tán	"*-am"

Rhyming conventions vary considerably from one literary tradition to another (on this see Baxter 1992:87–97), but the most straightforward definition of rhyming is probably that rhyming syllables must have the same main vowel and coda. If Old Chinese rhyming was based on that principle, then it should be clear that if the six-vowel reconstruction is correct, the traditional rhyme groups overlooked a large number of rhyming distinctions. Baxter (1992) was devoted in large part to testing the hypothesis that the rhyming distinctions predicted by the six-vowel

hypothesis actually are present in the *Shījīng*; and were simply overlooked in the traditional analysis. Of course, it is possible to imagine a literary tradition that would allow ***-on** and ***-an** to rhyme freely with each other. But the probabilistic analysis in Baxter (1992) showed that the predicted rhyming distinctions do indeed exist.[19] To take one of the clearest cases, the traditional 文 Wén rhyme group can be divided very cleanly into at least two rhymes in the Shījīng: one with the unrounded vowel ***ə** (including ***-ən** and ***-ər**) and the other with the rounded vowel ***u** (including ***-un** and ***-ur**); there are very few rhymes that mix these two groups.[20]

Table XX summarizes the correspondences between the traditional rhyme groups and the rhymes reconstructed according to the six-vowel system. In this table we include the reconstructions of Wáng Lì and Li Fang-kuei in quotation marks (so that they will not be confused with ours).

	rhyme group	Wáng Lì	Li Fang-kuei	Baxter-Sagart
1.	之 Zhī	"*-ə"	"*-əg"	*-ə
2.	職 Zhí	"*-ək"	"*-ək"	*-ək
3.	蒸 Zhēng	"*-əng"	"*-əng"	*-əŋ
4.	幽 Yōu	"*-u"	"*-əgw"	*-u, *-iw
5.	覺 Jué	"*-uk"	"*-əkw"	*-uk(s), *-iwk(s)
6.	冬 Dōng	"[*-ung]"	"*-əngw"	*-uŋ
7.	宵 Xiāo	"*-ô"	"*-agw"	*-aw, *-ew
8.	藥 Yào	"*-ôk"	"*-akw"	*-awk(s), *-ewk(s)
9.	侯 Hóu	"*-o"	"*-ug"	*-o
10.	屋 Wū	"*-ok"	"*-uk"	*-ok(s)
11.	東 Dōng	"*-ong"	"*-ung"	*-oŋ

---

19 The testing procedure involved what is known as a Monte Carlo method. The null hypothesis was words assigned to the same traditional rhyme group rhymed freely with each other (i.e., without regard to whether they were reconstructed with different main vowels). Those tokens of *Shījīng* rhyme words assigned to a single traditional group whose main vowel could be reconstructed unambiguously (according to the six-vowel hypothesis) were repeatedly scrambled randomly to estimate how often the observed degree of separation in rhyming could be expected to happen by chance. If the observed degree of separation would have occurred with a probability less than 0.05, then the null hypothesis (that the words within a traditional rhyme group rhymed with each other freely) was rejected. The method is described in Baxter (1992:97–137). In some cases (such as the rhymes ending in labial codas *-m and *-p) the rhymes were too few to yield a meaningful result; but on the whole, the analysis showed that the rhyming distinctions predicted by the six-vowel hypothesis did indeed exist, and had simply been overlooked in the traditional analysis.

20 This was established by the probabilistic analysis just described, in Baxter (1992:425–434) (where the vowel we now write as ***ə** was written as barred 'i', ***i**). Impressionistically is seems clear that at least some parts of the *Shījīng* also distinguish in rhyming between ***-ər** and ***-ən** and between ***-ur** and ***-un**, but it is difficult to test this hypothesis with the same technique, because it is difficult to avoid circular reasoning: the reconstruction of the distinction between *-n and *-r is partly based on rhyme evidence in the first place.

12.	魚 Yú	"*-a"	"*-ag"	*-a
13.	鐸 Duó	"*-ak"	"*-ak"	*-ak(s)
14.	陽 Yáng	"*-ang"	"*-ang"	*-aŋ
15.	支 Zhī	"*-e"	"*-ig"	*-e
16.	錫 Xī	"*-ek"	"*-ik"	*-ek(s)
17.	耕 Gēng	"*-eng"	"*-ing"	*-eŋ
18.	脂 Zhī	"*-ei"	"*-id"	*-ij
19.	質 Zhì	"*-et"	"*-it"	*-it(s), *-ik(s)
20.	真 Zhēn	"*-en"	"*-in"	*-in, *-iŋ, *-ir
21.	微 Wēi	"*-əi"	"*-əd"	*-əj, *-uj
22.	物 Wù	"*-ət"	"*-ət"	*-ət(s), *-ut(s)
23.	文 Wén	"*-ən"	"*-ən"	*-ən, *-un, *-ər, *-ur
24.	歌 Gē	"*-ai"	"*-ar"	*-aj, *-oj
25.	月 Yuè	"*-at"	"*-at"	*-at(s), *-et(s), *-ot(s)
26.	元 Yuán	"*-an"	"*-an"	*-an, *-en, *-on, *-ar, *-er, *-or
27.	緝 Qī	"*-əp"	"*-əp"	*-əp(s), *-ip(s), *-up(s)
28.	侵 Qīn	"*-əm"	"*-əm"	*-əm, *-im, *-um
29.	葉 Yè	"*-ap"	"*-ap"	*-ap(s), *-ep(s), *-op(s)
30.	談 Tán	"*-am"	"*-am"	*-am, *-em, *-om

The reconstructions of Wáng Lì and Li Fang-kuei were designed so as to be consistent with the traditional rhyme groups. In order to accomplish this, it is necessary to include in the reconstruction elements that we believe are spurious. For example, we reconstruct three different rhymes for these three words:

(39)   爛 làn < *lanH* < **[r]ˤan-s** 'cooked until soft'

　　　見 jiàn < *kenH* < **[k]ˤen-s** 'see (v.)'

　　　亂 luàn < *lwanH* < **[r]ˤo[n]-s** 'disorder, rebellion'

In Li Fang-kuei's reconstruction, they are "*lanh", "*kianh", and "*luanH" respectively, all ending in "*-anh", since they are all three assigned to the traditional 元 Yuán rhyme group (Li's "*-h" is an arbitrary notation for qùshēng). But we believe that Li's 'vocalic clusters' "*ia" in 見 < MC *kenH* and "*ua" in 亂 luàn < *lwanH* are spurious: their function is to reconcile the Middle Chinese forms with the traditional rhyme-group analysis. In fact, the three words do not have the same rhyme: 爛 làn < *lanH* < **[r]ˤan-s** rhymes as **-a[n]-s** (Odes 82.1, 124.3); 見 jiàn < *kenH* < **[k]ˤen-s** rhymes as **-en-s** (Ode 217.3), and 亂 luàn < *lwanH* < **[r]ˤo[n]-s** rhymes as **-on-s** (Odes 106.3 and 250.6).[21]

---

21 These are all the rhymes of the relevant words. The word 見 jiàn might appear to rhyme in Ode 102.3: "婉

Example 1: a passage from the Guōdiàn "Zī yī"

As an example of the kinds of arguments we find in the literature on recently discovered texts, let us consider a passage from the "Zī yī" 緇衣 that appears in both the Guōdiàn version and the Shànghǎi Museum version. In the Guōdiàn version, the passage appears at the end of strip 2 and the beginning of strip 3;

Guōdiàn "Zī yī", strips 2 and 3	Shànghǎi Museum "Zī yī", strip 2

The Guōdiàn editors transcribed these four characters as "民青（情）不 [糸弋]（弋）" Qiú Xīguī comments:

> 裘按：今本此句末字作「貳」，《釋文》所據本則作「忒」，簡本「[糸弋]」
> 字亦當讀為「忒」。

In the current version, the character at the end of the sentence is "貳"; the version on

---

[*[ʔ]o[n]ʔ] 夗變 [*[r]onʔ(-s)] 夗 / 總角丱 [*[k]ˤron-s] 夗 / 未幾見 [*[k]ˤen-s] 夗 / 突而弁 [*C.[b]ro[n]-s] 夗", but we believe that the third line was not intended as a rhyme. The *Jīngdiǎn shìwén* (p. 66) says that one version of the text has "未幾見之" where the Máo Shī has "未幾見夗", which would be consistent with this analysis.

which the [*Jīngdiǎn*] *shìwén* is based has "忒"; the character "[糸弋]" in the [Guōdiàn] bamboo strip version should also be read as "忒".

The words and characters involved are as follows:

(40) 弋 yì < *yik* < **lək** 'shoot an arrow with string attached'

  貳 = {二} èr < *nyijH* < **ni[j]-s** 'two'

  忒 tè < *thok* < **l̥ˤək** 'err; change; deceitful"

Qiú Xīguī's point is that 貳 èr in the current version of the "Zī yī" is just a graphic error for 忒 tè, and the character [糸弋] on the strip, which (like 忒 tè) has 弋 yì as phonetic element, should also be understood as {忒}.

  In Jì Xùshēng et al. (2004:96–97), the commentary by Zōu Jùnzhì says further:

虞萬里先生《上博簡、郭店簡《緇衣》與傳本合校補證（上）》則引王引之《經義述聞・毛詩下》謂今本 "貳" 當為 "貸（他得切）" 之誤，即 "忒" 之借字。

Professor Yú Wànlǐ (2002) quotes the "Máo Shī xià" section of *Jīng yì shù wén* by Wáng Yǐnzhī [1766–1834], saying that the character 貳 in the current version should be an error for 貸 (他得切 [i.e. MC *tha* + *tok* = *thok*]), a loan character for 忒 [MC *thok*].

Zōu Jùnzhì adds:

濬智案：“弋” 釋 “忒” 可從，弋（喻紐職部）、忒（透紐職部），聲近韻同，可以通假。忒，差錯也。

Note (by [Zōu] Jùnzhì): interpreting 弋 [yì < MC *yik*] as 忒 [tè < MC *thok*] is acceptable. 弋 yì has the [MC] initial [紐 niǔ] 喻 Yù [= MC *y-* or *hj-*, but in this case *y-*; see below] and is in the [OC] 職 Zhí rhyme group [Wáng Lì's *-ək, our *-ək(s)]; 忒 has the [MC] initial 透 Tòu [= MC *th-*] and is [also] in the [OC] 職 Zhí rhyme group; their initial consonants [聲 shēng] are close, and their rhymes [i.e., OC rhyme groups, 韻部 yùnbù] are the same, so they can be used interchangeably. 忒 tè means 'error'.

Zōu Jùnzhì argues that 弋 yì is an acceptable loan character to write 忒 tè, using the common practice of comparing their Middle Chinese initials and their Old Chinese rhyme groups. In "喻紐 Yù niǔ", 喻 Yù is an initial name from the traditional list of "thirty-six initials" (三十六字母) originating in the rhyme-table tradition of the Sòng dynasty. Chén Lǐ's analysis of Middle Chinese initials on the basis of the fǎnqiè spellings of the Guǎngyùn (18XX) showed that the traditional 喻 Yù initial actually includes two different initials, of which one ("喻三 Yù sān", also called 云 Yún, our *hj-*) is put in division III of the rhyme tables, and the other ("喻四 Yù sì, also called 以 Yǐ, our *y-*) is put in division IV (see Table XX above). Zōu Jùnzhì does not explicitly distinguish these two initials, using the term 喻 Yù, which includes both. But in this case, the initial *y-* of 弋 yì < MC *yik* is 喻四 Yù sì = 以 Yǐ, and that initial does indeed frequently show graphic connections with the 透 Tòu initial (MC *th-*), because OC *l- is one of the several sources of MC *y-*, and OC *l̥ˤ- is one of the several sources of MC *th-*; and OC *l- and *l̥ˤ- were evidently phonetically similar enough to be written with the same phonetic element, as many examples attest. So this phonological argument is sound. (Actually, when the characters involved have the same phonetic element, as here, that fact alone shows that they are phonetically close enough to be interchangeable in the pre-Qín script, so an explicit phonological argument is hardly necessary.)

But Jì Xùshēng, the editor and Zōu Jùnzhì's teacher, adds this note:

旭昇案：貳（日紐脂部），與忒聲均屬舌頭，韻為旁對轉，可以通假，不必視為 "貣" 之誤。

Note by [Jì] Xùshēng: 貳 èr has the [MC] initial 日 Rì [= MC *ny-*] and is in the [OC] 脂 rhyme group (our **-ij**); its initial and the initial of 忒 tè [MC *th-*] are both dentals [舌頭 shétóu], and their [OC] rhymes are in the relationship of 旁對轉 páng duìzhuǎn [see below], so they can be interchanged; it is not necessary to see 貳 èr as an error for 貣 tè.

We find this argument to be unconvincing.

Jì Xùshēng's argument about the initials of 貳 èr < MC *nyijH* and 忒 tè < MC *thok* is not entirely without basis, because there are cases of MC *ny-* and OC *th-* being written with the same

phonetic element. It is not because MC *ny-* and MC *th-* are both dentals (which is true but not sufficient evidence that they are interchangeable), but because the main source for MC *ny-* is OC ***n-**, and one of the sources of MC *th-* (apart from the ***l̥ˤ-**, mentioned above) is OC ***n̥ˤ-**, as in this example:

(41)   綏 ruí < *nywij* < ***nuj** 'tassle'

      妥 tuǒ < *thwaX* < ***n̥ˤojʔ** 'at ease'

(It is not uncommon for OC ***-uj** and ***-oj** to be written with the same phonetic element in spite of the different vowels). But his argument is based on an overgeneralization: it treats all cases of MC *th-* as the same, not recognizing that MC *th-* has several different OC origins, written with different sets of phonetic elements. We reconstruct four different Old Chinese origins for MC *th-* (even without taking possible preinitial elements into account): ***tʰˤ-**, ***l̥ˤ-**, ***n̥ˤ-**, and ***r̥ˤ-**, as in these examples:

(42)   汀 tīng < *theng* < ***tʰˤeŋ** 'level land beside water'

      丁、釘 dīng < *teng* < ***tˤeŋ** 'nail'

(43)   剔 tì < *thejH* < ***l̥ˤek-s** 'shave'

      易 yì < *yeH* < ***lek-s** 'easy'

      易 yì < *yek* < ***lek** 'change; exchange'

(44)   綏 ruí < *nywij* < ***nuj** 'tassle'

      妥 tuǒ < *thwaX* < ***n̥ˤojʔ** 'at ease'

(45)   體 tǐ < *thejX* < ***r̥ˤijʔ** 'body; limbs'

      禮 lǐ < *lejX* < ***rˤijʔ** 'propriety, ceremony'

The phonetic element 弋 yì < *yik* < ***lək** shows that the *th-* in 忒 tè < *thok* < ***l̥ˤək** comes from ***l̥ˤ-**, not ***n̥ˤ-**, so we would not expect it to show graphic connections to MC *ny-* < ***n-**.

    As for the (Old Chinese) rhymes of 貳 èr < *nyijH* < ***ni[j]-s** and 忒 tè < *thok* < ***l̥ˤək**, Jì Xùshēng says that they are in the relationship of "旁對轉 páng duìzhuǎn". The term *duìzhuǎn* refers to alternations among the three main rhyme categories shown in Table XX above: of 陰聲 yīnshēng rhymes (with vocalic codas), 入聲 rùshēng rhymes (with voiceless stop codas *-p, *-t, or *-k), and 陽聲 yángshēng rhymes (with nasal codas). Contacts between rhymes that are in different columns of Table XX, but on the same row, are called *duìzhuǎn*. The following is an example of "陰入對轉 yīn-rù duìzhuǎn":

(46)　　條 tiáo < *dew* < ***lˤiw** 'branch (n.), shoot (n.)'

　　　　滌 dí < *dek* < ***lˤiwk** 'wash, clean'

because 條 ***lˤiw** is in the traditional 幽 Yōu rhyme group (no. 4 in Table XX), which is in the yīnshēng column, while 滌 ***lˤiwk** is in the traditional 覺 Jué rhyme group (no. 5 in Table XX), which is in the same row as 幽 Yōu (no. 4), but in the rùshēng column: 覺 Jué (which includes both our ***-uk** and our ***-iwk**) is considered to be the rùshēng rhyme group corresponding to the yīnshēng group 幽 Yōu (which includes our ***-uk** and our ***-iwk**).

　　　　The term 旁轉 *pángzhuǎn* describes an alternation between rhymes in the same column of Table XX, but in different rows. Example () above (緌 ruí < *nywij* < ***nuj** 'tassle' and 妥 tuǒ < *thwaX* < ***n̥ˤojʔ** 'at ease') is an example of pángzhuǎn: 緌 ruí would normally be assigned to 微 Wēi (no. 21, our *-**əj** and ***-uj**), while 妥 tuǒ would be assigned to 歌 Gē (no. 24, our ***-aj** and ***-oj**), which are in the same column but different rows.

　　　　The term 旁對轉 *páng duìzhuǎn* is used to describe alternations where the rhyme groups involved are both in different columns and in different rows of Table XX. The first word 貳 èr is in the traditional 脂 Zhī rhyme group (no. 18 in Table XX, our ***-ij**), while 忒 tè is in 職 Zhí (no. 2, our ***-ək**). These are not only in different columns but also in different rows. Now there are some genuine alternations that could be described as "páng duìzhuǎn", but it should be clear that if such moves are allowed in an unconstrained way, then one can move from any place in the table to any other place. There may be cases of OC ***-ij** and OC *-**ək** written with the same phonetic, but none come to mind, and they are at least uncommon. In any case, terms like "duìzhuǎn" and "páng duìzhhuǎn" are simply descriptions of putative phonetic alternations, not explanations for them or justifications for assuming them. The fact that an alternation can be described in these terms does not mean that it is plausible without further explanation.[22]

---

22 Here is an example of páng duìzhuǎn for which an explanation can be found. The character 輠 has three MC readings: *hwaeX*, *hwanX*, and *hwojX*, all apparently meaning 'turn round (as a wheel)'. MC *hwaeX* would normally be assigned to the 歌 Gē rhyme group (no. 24 in the table); MC *hwanX* to 元 Yuán (no. 26), and MC *hwojX* to 微 Wēi (no. 21). The *hwaeX* reading in 歌 Gē (no. 24) and the *hwanX* reading in 元 Yuán (no. 26) can be described as a case of "yīn-yáng duìzhuǎn", since they are in different columns as well as different rows. MC *hwojX* (微 Wēi, no. 21) and MC *hwaeX* (歌 Gē, no. 24) are a case of "pángzhuǎn". MC *hwojX* (微 Wēi, no. 21) and MC *hwanX* (元 Yuán, no. 26) are in the "páng duìzhuǎn" relationship, since they are both in different columns and in different rows. But using these terms does not explain the alternations, nor does it generalize to all apparently parallel cases. Based on their MC forms, we reconstruct MC *hwaeX* < ***[g]ˤ<r>orʔ**, *hwojX* < ***[g]ˤurʔ**, and *hwanX* < ***[g]ˤorʔ**. Our explanation for the yīnshēng/yángshēng alternation is that OC final ***-r** merged with final ***-n** in most dialects, but with final ***-j** in others (in fact, that is the main criterion for identifying cases of final ***-r**; see Baxter & Sagart 2014b:252–268). The apparent alternation between ***-or** (歌

To sum up this somewhat complicated example, we would not say that 貳 *ni[j]s is close enough to 弑 *l̥ˤək (or its homonym 貣) are close enough in sound to be written with the same phonetic element. It is much more plausible that 貳 èr in the current version of the "Zī yī" is a graphic error for 弑 or 貣, as suggested by Wáng Yǐnzhī and Qiú Xīguī, and that phonological similarity is not involved.

Example 2: from "Tāng chǔ yú Tāng qiū 湯處於湯丘" (Tsinghua strips, vol. 5)

Here is a second argument to illustrate the traditional style of analysis and some of its weaknesses. On strip 2 of the text "Tāng chǔ yú Tāng qiū 湯處於湯丘", in volume 5 of the Tsinghua strips (p. 61), the following character occurs:

The character is transcribed by the editors as "惜（舒）" (p. 135), and the explanation is in note 8 (p. 137):

"「惜」字從昔聲，在心母鐸部，讀為書母魚部之「舒」。"

"惜" [MC *sjek*] has 昔 [also MC *sjek*] as its phonetic element; it belongs to the [MC] initial 心 Xīn [i.e., *s-*] and to the [OC] rhyme group 鐸 Duó [**-ak**]. It is to be read as "舒" [MC *syo*], which belongs to [MC] initial 書 Shū (i.e., *sy-*) and to the [OC] rhyme group 魚 Yú [**-a**].

We find this interpretation implausible on phonological grounds. There are cases of MC *s-* and *sy-* in the same phonetic series [find some], but would be easiest to explain if the initials were OC laterals (OC **s.l-** > MC *s-*, OC **l̥-** > MC *sy-*), but that is not the case here. The MC initials of words written with 昔 as phonetic (GSR 798) include those in Table XX below. (NB: the reconstructions of Old Chinese syllable onsets followed by "(?)" are all consistent with our reconstruction, but these reconstructions are provisional and need to be checked against

---

Gē) and **-ur** (微 Wēi) reflects the fact that the script did not always distinguish **o** from **u** before the front codas **-j**, **-n**, **-r**, and **-t**—as seen in example (); in this environment **o** and **u** may have actually merged in some dialects. But note that these explanations apply to only a small subset of the traditional 微 Wēi, 歌 Gē, and 元 Yuán groups (those syllables coming from **-or** and **-ur**), not the whole groups.

additional evidence. We have therefore not approved these forms for public release, and they are not for citation; they are given here only to illustrate the pattern of this phonetic series.)

昔 (798a)	*sjek*	***[s]Ak**	
惜 (798f)	*sjek*	***[s]Ak**	
腊 (798g)	*sjek*	***[s]- (?) + -Ak**	
耤 (798i)	*dzjek*	***[dz]- (?) + -Ak**	
踖 (798k)	*tsjek*	***[ts]- (?) + -Ak**	
	dzjek	***[dz]- (?) + -Ak**	
	tshjak	***[tsʰ]- (?) + -ak**	
趞 (798l)	*tshjak*	***[tsʰ]- + -ak**	
鵲 (798n)	*tshjak*	***[tsʰ]ak**	
斮 (798o)	*tsrjak*	***[ts]rak**	
	tsraewk	*< tsrjak < ***[ts]rak***	
厝 (798p)	*tshak, tshuH*	***[tsʰ]ˤ- (?) + -ak, *[tsʰ]ˤ- (?) + -ak-s**	
造 (798q)	*tshak*	***[tsʰ]ˤ- (?) + -ak**	
錯 (798s)	*tshak, tshuH*	***[tsʰ]ˤak, *[tsʰ]ˤak-s**	
醋 (798t)	*tshuH*	***[tsʰ]ˤ- (?) + -ak-s**	
借 (798u)	*tsjek, tsjaeH*	***[ts]Ak, *[ts]Ak-s**	
唶 (798v)	*tsjaeH*	***[ts]- (?) + -Ak-s**	
措 (798x)	*tshuH*	***[tsʰ]ˤak-s**	
蜡 (798y)	*tshjoH* *dzraeH*	***[tsʰ]- (?) + -ak-s** ***[dz](ˤ)- (?) + -rak-s}**	
耤 (798z)	*tsrheak* *dzreak*	***[tsʰ]- (?) + -rAk** ***[dz]- (?) + -rAk**	This example suggests that the MC change *Tsrj- > Tsr-* sometimes had the effect of changing *Tsrak > Tsrjaek*? to *Tsreak* instead of *Tsraek*.
籍 (798a')	*dzjek*	***[dz]Ak**	
藉 (798b')	*dzjek* *dzjaeH*	***[dz]Ak** ***[dz]Ak-s**	

籍 (798c')	dzjek	*[dz]- (?) + -Ak	
	tsrjaek	*[ts]- (?) + -rAk	
	dzraewk	*[dz]- (?) + -rak}	

Notice that in the whole list there are no examples of MC initials of the *Tsy*- series, nor are there any píngshēng words: only rùshēng < ***-ak** or ***-Ak** and qùshēng < ***-ak-s** or ***-Ak-s**.

What about 舒? The phonetic element is evidently 予 [MC *yo* < ***la**, *yoX* < ***laʔ**), Karlgren's GSR 83. The words Karlgren includes in this series are listed in Table XX. (As with the previous table, forms listed in curly brackets are consistent with our reconstruction but need to be checked further before we release them publicly; please do not cite them.)

予 (83a-b)	yo	***la**	
	yoX	***laʔ**	
迻 (83c-d)	—	—	K says: "The reading of this is unknown; it seems to mean 'to run' (inscr. 329). It is adduced here in order to show the late Chou form of a. above; d is Chou III/IV (inscr. 329)."
豫 (83e)	yoH	***la(ʔ)-s**	
杼 (83f)	drjoX	***lr- (?) + -aʔ**	
	zyoX	***[Cə].l- (?) + -aʔ**	
抒 (83g)	zjoX	***sə.laʔ**	
	zyoX	***Cə.laʔ**	
序 (83h)	zjoX	***sə.laʔ**	
芧 (83i)	drjoX	***lr- (?) + aʔ**	
紓 (83j)	syo	***l̥a**	
	zyoX	***Cə.laʔ**	
舒 (83k)	syo	***l̥a**	
野 (83l)	yaeX	***lAʔ**	
	dzyoX	?	
壄 (83m)	yaeX	***lAʔ**	
埜 (83n-o)	yaeX	***lAʔ**	

Notice that all the initials are consistent with a lateral initial (with the possible exception of 野 *dzyoX*; but MC *dzy*- and *zy*- are sometimes confused in the Middle Chinese sources). There are

also no rùshēng words in the group.

To sum up the reasons why we believe that 惜 xī < *sjek* < *[s]Ak is not a plausible loan character for {舒} shū < *syo* < ***l̥a**: although alternations between MC initial *s-* and *sy-* do occur, they occur mostly in words that can be reconstructed with an OC lateral: {舒} shū < *syo* < ***l̥a** is such a word, but 惜 xī < sjek < ***[s]Ak** is not. As for the rhymes, the phonetic 昔 *[s]Ak appears to be used to write only rùshēng words and qùshēng words derived from them (OC *-ak(s)), while there are no rùshēng words written with 予 as phonetic. Interchanges between rùshēng words (like 惜 xī < *sjek*) and píngshēng words (like 舒 shū < *syo*) are actually rather rare. At the moment we have no other interpretation to suggest for the character 惜 xī here, but interpreting it as {舒} is implausible on phonological grounds.

Example 3: {辯} biàn < *bjenX* < *[b]renʔ 'disputation' on strip 1 of the Guōdiàn Lǎozǐ (A)

The very first strip of the "A" (甲 jiǎ) version of the Guōdiàn Lǎozǐ begins as below (reading right to left):

The editors from the Jīngmén Museum transcribed this passage this way:

(47)　𢇍（絕）智（知）弃卞（辯），民利百伓（倍）。

"Renounce wisdom, abandon disputation, and the people will benefit a hundredfold."[23]

23　The received version is somewhat different: "絕聖棄智，民利百倍 jué shèng qì zhì, mín lì bǎi bèi" 'Renounce the sage, abandon wisdom, and the people will profit a hundredfold'.

The character we will focus on is the fourth, written on the strip as

(48)

Now there is no controversy about what word is represented; as far as I know, everyone agrees that it represents the word {辯} biàn 'disputation'. But Professor Qiú Xīguī 裘錫圭, whose comments are recorded in the commentary, says that the editors were mistaken in transcribing the character in (XX) above as "卞 biàn". Here is what he says (Jīngmén shì bówùguǎn 1998:113; we have separated the text into paragraphs for clarity):

“弃” 下一字當是 “鞭” 的古文，請看 《望山楚簡》…… “鞭””辯” 音近，故可通用。

後面《老子》丙第八號簡也有此字，讀為 “偏”。

本書《成之聞之》三二號簡、《尊德義》一四號簡也都有此字，分別讀為 “辨” 和 “辯”。

《五行》：三四號簡又有以此字為聲旁的从 “言” 之字，馬王堆帛書本《五行》與之相當之字為 “辯”。

The character after " 弃"qì ['abandon'] should be [interpeted as] an ancient form of "鞭"biān [MC *pjien* < ***pe[n]** 'whip (n.)']; see *Wàngshān Chǔ jiǎn* 《望山楚簡》…. "鞭" biān [MC pjien < ***pe[n]** 'whip'] and 辯" biàn [MC *bjenX* < ***[b]renʔ** 'disputation'] are close in pronunciation, so they could be used interchangeably.

Below [p. 9], strip 8 of *Lǎozǐ* version C [丙 bǐng] also has this character, and it is read as "偏" piān [MC *phjien* < ***pʰe[n]** 'oblique'].

Strip 32 of the text "Chéng zhī wén zhī" 《成之聞之》 in this book [p. 51], and strip 14 of "Zūn dé yì" 《尊德義》 [p. 56] also have this character; they are ro be read as "辨" biàn [MC *bjenX* < ***[b]renʔ** 'distinguish'] and 辯" biàn [MC *bjenX* < ***[b]renʔ**

'disputation'] respectively.[24]

Strip 34 of the text "Wǔ xíng" 《五行》 [p. 33] also has a character composed of this character and "言" yán; in the Mǎwángduī version of "Wǔ xíng", the character corresponding to it is "辯" biàn [MC *bjenX* < ***[b]renʔ** 'disputation'])."

The character erroneously used by the editors to transcribe the character in () above is 卞 biàn, a variant of 弁 biàn 'cap' (see Lǐ Jiāhào 1979). Our reconstruction is

(49)　　卞、弁、覍 biàn < *bjenH* < ***C.[b]ro[n]-s** 'cap'.

(The reason for reconstructing ***-o[n]** is that the word rhymes as ***-o[n]** in Ode 102.3, and 弁 is frequently used as a loan for 變 biàn < *pjenH* < ***pro[n]-s** 'change', as we saw above, which must also be reconstructed with ***-o[n]**.)

The point of this example is that both 鞭 biān 'whip' and 卞 ~ 弁 biàn 'cap' are both in the traditional 元 Yuán rhyme group, and all the words listed above begin with bilabial stops, so according to the traditional categories, they should all be interchangeable in the script; traditional phonology gives no reason for the editors to suspect that {辯} 'disputation' could not be written using 卞 biàn 'cap' as a phonetic element. But in the six-vowel reconstruction, {卞} 'cap' must be reconstructed with ***-o[n]**, and both {鞭} 'whip' and {辯} 'disputation' must be reconstructed with ***-e[n]**. We would therefore predict that 卞 'cap' should not be used to write {辯} 'disputation'. Qiú Xīguī's examples show that, as predicted by the six-vowel reconstruction, the words written with the 'whip' character as phonetic are all (for independent reasons) to be reconstructed with ***-e[n]**. As a guide to which words should be written with which phonetic element, the six-vowel reconstruction is a more reliable guide than the traditional rhyme groups.

Example 4: 芋 for {華} in Kǒngzǐ Shī lùn

[Problem of notation: Ma Chengyuan didn't recognize 芋 as 華]
[Other problems: not all words assigned correctly: 美; 脂 vs. 微]

5. Conclusion

24　In all these cases, the editors also transcribed the characters in question as 卞 biàn or as [卞/又], i.e. 卞 biàn over 又 yòu

Our purpose has been to summarize some of the major features of our new reconstruction, and to suggest ways in which the phonological analysis of early texts can be improved. Just the use of a conventional alphabetic notation for Middle Chinese, we believe, would make arguments about phonology easier to formulate and understand—even though the more traditional notation will certainly continue to be used. For Old Chinese, our approach not only has the advantages of an explicit alphabetic notation, but also recognizes many distinctions in both initial consonants and rhymes that are overlooked in the traditional approach to Old Chinese phonology. We believe it can serve as a useful tool in understanding newly discovered early documents. At the same time, the work of reconstruction is not finished, and we hope it will be possible to improve the reconstruction as more evidence from early texts becomes available.

References

Bái Yīpíng 白一平 [William H. Baxter]. 2010. "Yì", "shì", "shè" děng zì de gòunǐ hé zhōnggǔ *sy-* (shūmǔ = shěn sān) de láiyuán "埶"、"勢"、"設"等字的構擬和中古 sy-（書母＝審三）的來源. *Jiǎnbó* 簡帛 5. 161–178.

Baxter, William H. 1977. Old Chinese origins of the Middle Chinese chóngniǔ doublets: a study using multiple character readings. Cornell University Ph.D. dissertation.

Baxter, William H. 1992. *A handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.

Baxter, William H. 2015. LIST OF CORRIGENDA for *A Student's Dictionary of Classical and Medieval Chinese*. http://www.brill.com/sites/default/files/ftp/downloads/69824_Corrigenda-A_Students_Dictionary_of_Classical_and_Medieval_Chinese.pdf (31 October 2015).

Baxter, William H. & Laurent Sagart. 2014a. The Baxter-Sagart reconstruction of Old Chinese (version 1.1, 20 September 2014). [The site provides downloadable .pdf files, sorted in various ways, with reconstructions of almost 5000 items, many more than are included in *Old Chinese: a new reconstruction*.]

Baxter, William H. & Laurent Sagart. 2014b. *Old Chinese: a new reconstruction*. New York: Oxford University Press.

Boltz, William G. 1985. The Lao tzu text that Wang Pi and Ho-shang Kung never saw. *Bulletin of the School of Oriental and African Studies* 48(3). 493–501.

Dīng Shēngshù 丁聲樹 & Lǐ Róng 李榮. 1981. *Gǔ-jīn zìyīn duìzhào shǒucè* 古今字音對照手冊. Běijīng 北京: Zhōnghuá shūjú 中華書局.

Doerfer, Gerhard. 1963–1975. *Türkische und mongolische Elemente im Neupersischen; unter besonderer Berücksichtigung älterer neupersischer Geschichtsquellen, vor allem der Mongolen- und Timuridenzeit*. 4 vols. Wiesbaden: F. Steiner.

Dǒng Tónghé 董同龢. 1948. Shànggǔ yīnyùn biǎo gǎo 上古音韻表稿. *Zhōngyāng yánjiūyuàn lìshǐ yǔyán yánjiūsuǒ jíkān* 中央研究院歷史語言研究所集刊 [Bulletin of the Institute of history and philology, Academia Sinica] 18. 1–249

Doyle, Arthur Conan. 1894. *Memoirs of Sherlock Holmes*. New York: Harper & Brothers.

Graham, A. C. 1989. *Disputers of the Tao: philosophical argument in ancient China*. La Salle,

IL: Open Court.

Haudricourt, André-Georges 1954a. Comment reconstruire le chinois archaïque. *Word* 10. 351–364.

Haudricourt, André-Georges. 1954b. De l'origine des tons en Vietnamien. *Journal Asiatique* 242. 68–82.

Hú Píngshēng 胡平生 & Hán Zìqiáng 韓自强. 1988. *Fùyáng Hàn jiǎn Shījīng yánjiū* 阜陽漢簡詩經研究. Shànghǎi 上海: Shànghǎi gǔjí chūbǎnshè 上海古籍出版社.

Huáng Bùfán 黄布凡. 1992. *Zàng-Miǎn yǔzú yǔyán cíhuì* 藏缅语族语言词汇 = *A Tibeto-Burman lexicon*. Běijīng 北京: Zhōngyāng mínzú xuéyuàn chūbǎnshè 中央民族学院出版社.

Jakobson, Roman. [1957] 1971. Mufaxxama: The "emphatic" phonemes in Arabic. *Selected writings*, 1:510–522. The Hague: Mouton.

Jaxontov, Sergej Evgen'evič. 1960a. Consonant combinations in Archaic Chinese. *Papers presented by the USSR delegation at the 25th International Congress of Orientalists,* Moscow, 1–17. Moscow: Oriental Literature Publishing House.

Jaxontov, Sergej Evgen'evič. 1960b. Fonetika kitajskogo jazyka 1 tsysjačeletija do n. e. (labializovannye glasnye). *Problemy Vostokovedenija* 1960(6). 102–115.

Jaxontov, Sergej Evgen'evič. 1963. Sočetanija soglasnyx v drevnekitajskom jazyke [Consonant combinations in Archaic Chinese]. *Trudy dvadcat' pjatogo meždunarodnogo kongressa vostokovedov, Moskva, 9–16 avgusta 1960 g.* Vol. 5: *Zasedanija sekcij XVI–XX* [Works of the 25th International Congress of Orientalists, Moscow, 9–16 August 1960. Vol. 5: Sessions of sections XVI–XX, 89–95]. Moskva: Izdatel'stvo vostočnyj literatury.

Jì Xùshēng 季旭昇, Chén Línqìng 陳霖慶, Zhèng Yùshān 鄭玉姍 & Zōu Jùnzhì 鄒濬智. 2004. *Shànghǎi bówùguǎn cáng zhànguó Chǔ zhúshū (yī) dúběn* 上海博物館藏戰國楚竹書 (一) 讀本. Táiběi 臺北: Wànjuàn lóu 萬卷樓.

Jīngmén shì bówùguǎn 荊門市博物館. 1998. *Guōdiàn Chǔ mù zhújiǎn* 郭店楚墓竹簡. Běijīng 北京: Wénwù chūbǎnshè 文物出版社.

Karlgren, Bernhard. 1954. *Compendium of phonetics in ancient Chinese and archaic Chinese.* Stockholm: Museum of Far Eastern Antiquities.

Karlgren, Bernhard. 1957. Grammata serica recensa. *Bulletin of the Museum of Far Eastern*

Antiquities 29. 1–332.

Kroll, Paul W., William G. Boltz, David R. Knechtges, Y. Edmund Lien & Antje Richter. 2014. *A student's dictionary of Classical and Medieval Chinese*. Leiden: Brill.

Li, Fang-kuei [Lǐ Fāngguì 李方桂]. 1971. Shànggǔ yīn yánjiū 上古音研究. *Tsing Hua Journal of Chinese Studies* n.s. 9. 1–61.

Lǐ Jiāhào 李家浩. 1979. Shì "biàn" 釋"弁". *Gǔ wénzì yánjiū* 古文字研究 1. 391–395.

Mǎ Chéngyuán 馬承源. 2001–. *Shànghǎi bówùguǎn cáng Zhànguó Chǔ zhúshū* 上海博物館藏戰國楚竹書. Shànghǎi 上海: Shànghǎi gǔjí chūbǎnshè 上海古籍出版社.

Norman, Jerry. 1973. Tonal development in Min. *Journal of Chinese Linguistics* 1(2). 222–238.

Norman, Jerry. 1974. The initials of Proto-Min. *Journal of Chinese Linguistics* 2(1). 27–36.

Norman, Jerry. 1981. The Proto-Min finals. *Zhōngyāng yánjiūyuàn guójì Hànxué huìyì lùnwén jí: Yǔyán wénzì zǔ* 中央研究院國際漢學會議論文集:語言文字組 [Proceedings of the International Conference on Sinology : section on linguistics and paleography], 35–73. Táiběi: Academia Sinica.

Norman, Jerry. 1994. Pharyngealization in Early Chinese. *Journal of the American Oriental Society* 114(3). 397–408.

Pān Wùyún 潘悟云. 1997. Hóuyīn kǎo 喉音考. *Mínzú yǔwén* 民族语文 1997(5). 10–24.

Pulleyblank, Edwin G. 1962–1963. The consonantal system of Old Chinese. *Asia Major* (n. s.) 9. 58–144, 206–265.

Pulleyblank, Edwin G. 1973. Some new hypotheses concerning word families in Chinese. *Journal of Chinese Linguistics* 1(1). 111–125.

Sagart, Laurent. 1993. REVIEW ARTICLE - A Handbook of Old Chinese Phonology. *Diachronica* 10(2). 237–260.

Sagart, Laurent. 1999. *The roots of Old Chinese*. Amsterdam: John Benjamins.

Starostin, Sergei A. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy*. Moskva: "Nauka," Glav. red. vostochnoĭ lit-ry.

Wáng Lì 王力. 1999. *Gǔdài Hànyǔ (jiàodìng chóngpái běn)* 古代漢語 (校訂重排本). 4 vols. Běijīng 北京: Zhōnghuá shūjú 中華書局.

Xiàng Xī 向熹. 1986. *Shījīng cídiǎn* 詩經詞典. Chéngdū 成都: Sìchuān rénmín chūbǎnshè 四川人民出版社.

Yú Wànlǐ 虞萬里. 2002. Shàngbó jiǎn, Guōdiàn jiǎn "Zi yi" yǔ chuánběn héjiào bǔzhèng (shàng) 上博簡、郭店簡《緇衣》與傳本合校補證（上）. *Shǐ lín* 史林 2002(2).

Zhèngzhāng Shàngfāng 鄭張尚芳. 1987. Shànggǔ yùnmǔ xìtǒng hé sìděng, jièyīn, shēngdiào de fāyuán wèntí 上古韻母系統和四等、介音、聲調的發源問題. *Wēnzhōu shīyuàn xuébào (Shèhuì kēxué bǎn)* 溫州師院學報 (社会科学版) 1987(4).

[Zhèngzhāng Shàngfāng 鄭張尚芳. 2003. *Shànggǔ yīnxì* 上古音系. Shànghǎi 上海: Shànghǎi jiàoyù chūbǎnshè 上海教育出版社.]

Zōu Jùnzhì 鄒濬智. 2004. "Zī yī" yìshì 《緇衣》譯釋. In Jì Xùshēng 季旭昇 (ed.), *Shànghǎi bówùguǎn cáng zhànguó Chǔ zhúshū (yī) dúběn* 上海博物館藏戰國楚竹書 (一) 讀本, 84–155. Táiběi 臺北: Wànjuàn lóu 萬卷樓.