Questions about Baxter & Sagart's Old Chinese

Axel SCHUESSLER

Professor Emeritus of History, Wartburg College

This article raises some questions about the methodology of Old Chinese reconstruction in general and Baxter and Sagart 2014 in particular. Issues are discussed using examples from Baxter and Sagart.

Keywords: Old Chinese reconstruction, Comparative method, Hypothetic-deductive method, Baxter and Sagart, Etymology, *Xiéshēng* series

List of Abbreviations:

B&S	Baxter and Sagart *Old Chinese: A New Reconstruction,* 2014[1]
IE	Indo-European
MC	Middle Chinese
MY	Miao-Yao 苗瑤 languages
NOC	Old Chinese, new reconstruction by Baxter and Sagart's 2014 book[2]
OB	Shang Dynasty oracle bone inscriptions
OC	Old Chinese
OCM	Minimal Old Chinese (Schuessler 2009)[3]
QYS	*Qieyun* system (traditionally 'Middle Chinese')
ST	Sino-Tibetan
TB	Tibeto-Burman
WT	Written Tibetan

1. Introduction

No one knows what Old Chinese (OC) was like, none of us was there. We can only interpret sparse data, and these interpretations are more subjective than we might wish. Every one who tries to reconstruct OC on the basis of, more or less, the same material (Middle Chinese=MC, *xiéshēng* 諧聲 series, *Shījīng* 詩經 rhymes) arrives at a different OC language. Perhaps OC was as proposed by Baxter & Sagart (B&S), or by Baxter[4], or by Wáng Lì 王力 , or Li Fang-kuei 李方桂 , Pān Wùyún 潘悟雲 , etc., or something entirely different. Or OC simply cannot be reconstructed with any degree of plausibility.

Note the OC phonological interpretations of the copula *wéi* (MC jiwi) 隹 維惟 'to be, it is/was' > 'only' 唯 (隹 has also the reading MC tświ< *tui):

---

1   William H. Baxter and Laurent Sagart, *Old Chinese: A New Reconstruction* (Oxford: Oxford University Press, 2014).

2   Ibid.

3   Axel Schuessler, *Minimal Old Chinese and Later Han Chinese: A Companion to Grammata Serica Recensa* (Honolulu: Hawai'i University Press, 2009).

4   William H. Baxter, *A Handbook of Old Chinese Phonology* (Berlin/New York: Mouton de Gruyter, 1992).

*dʑ̯wər   Karlgren[5]
*rəd   Li Fang-kuei[6]
*lul   Schuessler[7]
*iuəi(?)   Wáng Lì[8]
*ljuəj   Schuessler[9]
*wjij   Baxter[10]
*t(ə)-wij Sagart[11]
*k-lul   Pān Wùyún[12] etc. (cf. Schuessler[13])
*ɢʷi   Zhèngzhāng[14]
*wi   Schuessler[15] (cf. Baxter[16])
*ɢʷij   B&S (cf. Zhèngzhāng[17])

(Comments on the reconstructions: (1) The initials. The graph 隹 has two Middle Chinese readings *tświ* and *jiwi* (initial MC *ji-* is called *yù sì* 喻四), the phonetic series includes also 帷 MC *jwi-* (initial MC jw- = *yù sān* 喻三), among others. In 隹 'to be' Karlgren assumed initial *d from MC *ji-* in order to account for the presumed phonetic similarity with tświ< OC *tui. Later, Li Fang-kuei concluded that the *yù sì* initial goes back to OC *r- which he deemed phonetically close enough to a dental stop to warrant the graphic loan, therefore

5   Bernhard Karlgren, *Grammata Serica Recensa* (Stockholm: Museum of Far Eastern Antiquities, 1957).

6   Li Fang-kuei 李方桂 , "Shànggǔ yīn yánjiū"上古音研究 , in *Tsing Hua Journal of Chinese Studies* 清華學報 , n.s. 9 (Taipei: The Tsing Hua Journal of Chinese Studies, 1971), 1–61.

7   Axel Schuessler, "R and L in Archaic Chinese," *Journal of Chinese Linguistics* 2, 2(1974): 186–199.

8   Wáng Lì 王力 , "*Tóngyuán zìdian*" 同源字典 (Běijīng: Shāngwù yìnshūguǎn, 1982).

9   Axel Schuessler, *A Dictionary of Early Zhou Chinese* (Honolulu: Hawai'i University Press, 1987).

10 Baxter, *A Handbook of Old Chinese Phonology*.

11 Laurent Sagart. *The Roots of Old Chinese* (Amsterdam/Philadelphia, John Benjamins Publishing Company, 1999), 94.

12 Pān Wùyún 潘悟雲 , *Hànyǔ lìshǐ yīnyùnxué* 漢語歷史音韻學 (Shànghǎi: Shànghǎi jiàoyù chūbǎnshè, 2000).

13 Schuessler, "R and L in Archaic Chinese," 186–199.

14 Zhèngzhāng Shàngfāng 鄭張尚芳 , *Shànggǔ yīnxì* 上古音系 (Shànghǎi jiàoyù chūbǎnshè, 2003).

15 Schuessler, *Minimal Old Chinese*.

16 Baxter, *A Handbook of Old Chinese Phonology*.

17 Zhèngzhāng Shàngfāng, *Shànggǔ yīnxì*.

Li's *rəd. Others concluded that Li's *r- was really *l-, as suggested by the obvious OC *r- as source for MC l-, and by Tibeto-Burman (TB) cognates (*liù* 六 *ruk, cf. Tibetan *drug* 'six'). Meanwhile, it was realized that the OC source of *yù sān* could actually be reduced to *w-, as in *wéi* 帷 *wəi which made Li's *rəd and subsequent initial *l- look odd in this phonetic series. In the end, it was determined that *yù sì* in 隹 goes back to OC *wi; note that this graph was transcribing Indic *vi* in Buddhist texts, the initial *yù sì* (instead of the expected *yù sān* *w-) was cause by the high vowel /i/. As a result, Baxter and others reconstructed 'to be' as *wij or *wi. In order to make a *wij compatible with MC tświ, Sagart proposed at one time an OC prefix *t- that was later lost, hence *t(ə)-wij. Pān Wùyún suggested that the initial might have been *k-l-. Since Zhèngzhāng and others like B&S have introduced a voiceless uvular *q into OC, the search for a voiced counterpart resulted in repurposing the *yù* initials as OC *Gw-. I find the *q- and G-arguments too hypothetical and unnecessary; the more abstract the pile of theories, the farther the reconstruction seems to be removed from anything that ancient Chinese might actually have spoken. (2) The rhyme of 隹: MC -i occasionally alternates with MC -n in phonetic series and early poetic rhymes; note also the pronunciation *yī* 衣 for the Yīn 殷 dynasty. Therefore Karlgren thought that this MC -i must go back to an OC sound that was somehow similar to -n, therefore his -r (*diwər). Others suggested *-l, hence my one time suggestion *lul. Eventually, it was concluded that the rhyme was *-ij. The final -j is postulated in order to make the word conform to a theorized root structure CVC; here -j behaves like a final consonant which is only hypothetical. I would simply suggest *-i, therefore OC *wi).

Three elements go into a reconstruction of OC: (1) knowable data, linguistic facts; (2) evidence from *xiéshēng* series, i.e. graphic evidence; (3) theories, hypotheses, ideas, assumptions, including internal reconstruction, i.e. conclusions based on distributional patterns. The OC results (reconstructions) depend on which of these takes *precedence* over the others.

When *xiéshēng* take precedence over the others, the question asked is: what was the word that was written with this graphic element? Linguistic data may be of secondary significance. Then 隹維惟 may come out as OC *di̯wər (Karlgren[18]) or *t(ə)-wij (Sagart[19]) — because of *zhuī* 隹 MC tświ.

When *theory* takes precedence, the question is: how does the OC word

18 Karlgren, *Grammata Serica Recensa.*

19 Sagart, *Roots of Old Chinese,* 94.

fit into the theory? You get OC *rəd (Li Fang-kuei[20]), *k-lul (Pān Wùyún[21]) or
*ɢʷij (B&S; also Zhèngzhāng[22]) — the latter because of the uvular theory.

When *linguistic* data take precedence, the question is: why was the OC
word written with this graph? Then you get OC *wjij (Baxter[23]) — because of
MC, Sanskrit transcriptions, TB; the graph is not a linguistic problem, but a
philological one and may linguistically be nearly irrelevant. I prefer this last
approach, let linguistic data take precedence over writing and theories:

wéi	Modern Standard Chinese
vi [wi]	Sin Sukchu (standard reading; Ming Dynasty; Coblin[24])
ywi [yi]	'Phags-pa Chinese (Mongol Dynasty; Coblin[25])

(jiwi) *Qiēyùn* 601 AD, not a phonetically attested form, but a
reconstruction based on attested categories and dialects (Li[26])

iui	Old Northwest Chinese (400 AD; Coblin[27])
wi[2]	Common Dialectal Chinese (Norman[28])
wi	Hàn Buddhist Transcriptional Dialect (Coblin[29]); transcribing

Skt. *vi*

___ ?	OC
*wəy or *wi	Tibeto-Burman 'to be'

Thus I believe that the OC word for *wéi* 隹維惟 'to be' can hardly be any-
thing other than *wi (or *wij, *wjij — however you want to write it); the role
of *zhuī* 隹 MC tświ is a philological or graphic problem.

20 Li Fang-kuei, "Shànggǔ yīn yánjiū," 1–61.

21 Pān Wùyún, *Hànyǔ lìshǐ yīnyùnxué.*

22 Zhèngzhāng Shàngfāng, *Shànggǔ yīnxì.*

23 Baxter, *A Handbook of Old Chinese Phonology.*

24 W. South Coblin, *A Handbook of 'Phags-pa Chinese* (Honolulu: Hawai'i University Press, 2007).

25 Ibid.

26 Li Fang-kuei, "Shànggǔ yīn yánjiū," 1–61.

27 W. South Coblin, *Studies in Old Northwest Chinese,* Journal of Chinese Linguistics Monograph 4 (Berkeley, California: Project on Linguistic Analysis, 1991).

28 Jerry Norman, "Common Dialectal Chinese" in *The Chinese Rime Tables: Linguistic Philosophy and Historical-Comparative Phonology*, ed. David P. Branner (Amsterdam/ Philadelphia: John Benjamins Publishing Company, 2006), 233–254.

29 W. South Coblin, "Notes on the dialect of the Han Buddhist transcriptions," in *Proceedings of the International Conference on Sinology*, (Taipei: Academia Sinica, 1981), 121–183.

Baxter and Sagart give precedence to theory.[30] Their work can be read in two ways:

(1) As suggesting imaginative new ideas that explore possibilities for OC, legitimate and fascinating mental exercises (Gedankenspiele) if you wish. In this sense, the authors made an important contribution to the field. We appreciate, and are grateful for, their monumental efforts.

(2) Or one can understand B&S as a definitive reconstruction of OC. Even though B&S explicitly say that their work is not to be seen as such, it seems to imply this, and those who cite B&S reconstructions treat their new OC (NOC) as if it were the actual language of the Zhou 周 period. I have difficulty accepting B&S in this latter sense, because I for one am interested in what OC might *plausibly* have been like, not theoretical possibilities. Thus I look at OC as a language that people might actually have spoken. My perspectives are therefore contrary to much of what I find in B&S. This concerns over-all methodology and approach, as well as details. I will try to explain why I find B&S's new OC unconvincing, except, generally, those features which have been carried over from Baxter 1992.[31] Much of the rest is of a "may be/ may be not" nature.

2. Methodology

2.1 There are two fundamentally different methodological approaches to historical reconstruction: one is the traditional method with neogrammarian principles (no exceptions to laws) that starts with evidence and data, and draws its conclusions from these. For example, the Indo-European (IE) handbook by Szemerényi presents the evidence, data, the facts, summarizes others' proposals, and then typically continues "these data lead to the conclusion that..."[32] A reader can follow his arguments and insights, and agree or disagree with him on the merits of the evidence. For instance: what was the structure of an IE word root? Traditional neogrammarians study as evidence words like *pet- 'to fly', *kwi 'who', *aǵ- 'drive' and *i- 'go' in many IE languages; this empirical evidence leads to the conclusion that IE roots could have the structure CVC, CV, VC, V, or for short (C)V(C). For instance, in all IE languages the words for 'drive', like Latin *ag-ere*, Sanskrit *aj-āmi* etc.

---

30 Baxter and Sagart, *Old Chinese*.

31 Baxter, *A Handbook of Old Chinese Phonology*.

32 Oswald J. L. Szemerényi, *Introduction to Indo-European Linguistics*, 4th ed. (Oxford: Oxford University Press, 1990/1996).

consistently point to an IE *aǵ- as a common source.

In other methods hypotheses and theories have priority. Today it is the hypothetic-deductive method as formulated by Karl Popper. Here, hypotheses are set forth first, then the investigator tries to fit data into the hypothesis which is then tested by attempts to falsify it. Critics may point out that this whole approach is based on the 'confirmation fallacy': first one has an idea, and then one picks and choses data that confirm it, and ignores or rationalizes away (with additional hypotheses) what does not fit. This approach seems to amount to "anything goes" unless/until falsified. Thus the majority of IE linguists follow the laryngealist school that goes back to Saussure, so called because it postulates purely theoretical laryngeal phonemes like /h/ that have supposedly disappeared in IE languages everywhere, except for some examples in Hittite. They adhere strictly the purely hypothetical dogma that the IE root had the structure CeC, they fit every piece of evidence into this scheme with the help of laryngeals. The root *pet- seems to prove their point, but then *aǵ- becomes h_2eǵ-, *i- becomes *h1ei- (i and u are structurally consonants in laryngealism). Hittite sometimes confirms such laryngeals, but not always, as is the case with Greek *apó* 'behind' = Hittite *appa* without the hypothesized *h_2; in order to satisfy the CeC theory, laryngeal *h_4 was introduced which disappears everywhere, including in Hittite. This strikes me as a *reductio ad absurdum*. A representative of extreme laryngealism is Winfred Lehmann who claims that laryngeals explain all kinds of puzzling phenomena, including the consonantal strengthening in Germanic (as in German *Brücke* vs. 'bridge'),[33] and whom Szemerényi may have had in mind when he said that some laryngealists "find in them the final solution to all mysteries."[34] Lehmann introduces his arguments and propositions with "I assume that..." and piles assumption upon hypothetical assumption, and the reader has to take his word on faith. Szemerényi's conclusions rely on data as evidence, he follows the traditional method of historical reconstruction.

Baxter 1992 followed the traditional method. When the evidence seemed inconclusive, he used one of several criteria to identify the most plausible reconstruction among all other possibilities. The criteria or tools include parsimony (the least complex explanation = Ockham's = Occam's razor) and naturalness, beside internal structural patterns and universal phenomena. This way Baxter 1992 often followed earlier proposals.

---

33 Winfred P. Lehmann, *Theoretical Bases of Indo-European Linguistics* (London/New York: Routledge, 1993).

34 Szemerényi, *Introduction to Indo-European Linguistics*, 130.

B&S threw these tools by and large out: no mention of Ockham or naturalness. Instead, they explicitly use the hypothetic-deductive method, sounding like Lehmann, except instead of repeating "we assume", they assert "we reconstruct..." — basically asking the reader to take their proposals on faith (because they hardly ever show how theirs are preferable to alternative explanations).

Consider parsimony:

NOC uvulars (喉音) like *q: How do B&S get from MC kjɐŋˊ 景 to NOC *C.qraŋʔ 'bright'?

• Assumption 1: MC ʔ- and k- were more similar in OC in order to explain *xiéshēng* connections MC kjɐŋˊ 景 ~MC ʔjɐŋˊ 影 .

• Assumption 2: Therefore some MC ʔ derive from NOC *q, as in MC ʔjɐŋˊ 影 NOC *qraŋʔ.

• Assumption 3: Unlike Pān Wùyún 1997,[35] B&S believe that NOC *q- and *k- were too dissimilar for *xiéshēng* agreement, therefore some MC k- derive from *q- as well.

• Assumption 4: Because *q- is already used for a source of MC ʔ-, the initial must have been different, therefore the assumption of an unspecifiable prefix *C-: MC kjɐŋˊ 景 from NOC *C.qraŋʔ.

Every one of these 4 assumptions may or may not be correct; that means B&S's reconstruction has only a 1 in 4 chance of being true, i.e. only when all assumption are equally correct.

This stack of assumptions ("hypotheses") makes this NOC reconstruction rather implausible. Ockham's razor should have provided a red flag.

Alternative: the traditional reconstruction: Push known facts, like MC, back in time as far as possible, make adjustments only based on compelling evidence (*xiéshēng*, rhymes). There is no reason to assume that *k- and *q- in *xiéshēng* is any more compelling than *k- and *ʔ-. The simplest and most plausible interpretation is that

MC kjɐŋˊ 景 was OC *kraŋʔ, and

MC ʔjɐŋˊ 影 was OC *ʔraŋʔ

影 OC *ʔraŋʔ had 景 OC *kraŋʔ appear as "phonetic" because of the meaning 'bright ~ shadow'.

This interpretation requires only one or two assumptions.

Naturalness.

Take *q in *yǐng* 影 NOC *qraŋʔ again, or their Semitic (閃族語) pha-

---

35 Pān Wùyún 潘悟雲 , "Hóu yīn kǎo" 喉音考 , *Mínzú yǔwén* 民族語文 5, (1997):10–24.

ryngealization /ˤ/ (咽音) as the source in MC Division I/IV words like *gāng* 亢 NOC *k-ŋˤaŋ: these sounds are not what one typically finds in East Asia. There must be a better explanation for the MC Divisions — and the NOC *q are unwarranted anyway. These exotic sounds should be red flags. When one needs to appeal, in an attempt to be persuasive, to unique features in far away languages like Semitic (B&S), Athabascan (美洲亞大巴斯卡語 , Li Fang-kuei[36]) or Greek, that diminished the credibility of a proposal.

These claims and suppositions alone throw doubts on the plausibility of NOC.

2.2 It is my impression that the hypothetic-deductive method is based on the "confirmation fallacy", i.e. an argument starts out with an idea (hypothesis), and then proponents hunt for validating evidence where ever it may be found, be it alleged loan words in outlier languages (like Ruc or Lakkia), and phonetic over-interpretation of graphs — all that without showing how their ideas are superior solutions to traditional or alternative explanations. It is like putting the cart before the horse. B&S invoke Einstein's theories, but in the hard sciences and solid linguistic work (e.g. Baxter 1992),[37] the hypotheses are arrived at through induction, i.e. evaluation of data, and only then they are proposed as a "theory" and ready to be "falsified".

2.3 B&S have a new way of reconstructing OC, that is already foreshadowed in Baxter 1992 and Sagart 1999a.[38] They try to capture in their OC forms all possible alternative features that might agree with their hypotheses. Their OC reconstructions are therefore full of brackets, parentheses, hyphens and dots. Thus *jǔ* 矩 'carpenter's square' is NOC *[k] (r)aʔ which means that there are four possibilities:

*kʷaʔ
*qʷaʔ
*kʷraʔ
*qʷraʔ

Or take *quǎn* 犬 'dog' : NOC *[k]ʷʰˤ[e][n]ʔ, where there might have been q instead of k, i instead of e, and r instead of n, i.e. nine possibilities, if not more. The reader is now free to pick any of these possibilities, and has the option of connecting the Chinese word with Tibeto-Burman *kywal or *kwi,

36   Li Fang-kuei, "Shànggǔ yīn yánjiū," 1–61.

37   Baxter, *A Handbook of Old Chinese Phonology*.

38   Baxter, *A Handbook of Old Chinese Phonology*; Sagart, *The Roots of Old Chinese*.

ancient Greek *kýōn*, or who knows with what in other languages — there are lots of exciting possibilities. But seriously: what is a reader to do with a set of possibilities (not "reconstructions"), any one of which can only be 1/9 correct? (Or are they thought to be all correct simultaneously?) If you decide on *qʷʰˤerʔ, you will be 8/9 wrong, in the case of *jǔ* 矩 any possibility can only be 1/4 correct; and so on. What does this explain about OC? How can this ambivalent algebraic formula possibly be used for one's own research? How could I trust arguments based on NOC any better than NOC itself?

By the time-tested traditional method of historical reconstruction, we should be guided by the data we actually have, word forms that are attested (in actual Sinitic languages, Middle Chinese), not a pile of theoretical possibilities. 矩 was most likely *kʷaʔ (the uvulars are unwarranted, again a hypothetical invention; and medial "(r)" is only a possibility), and 犬 was, I think, most likely *khwînʔ (for uvulars see above, and because *some* MC final n may derive from OC r does not mean that there was a reasonable chance that most or all final n must be suspected of deriving from *r). These plausibilities are not proven OC forms, but at least weighted in favor of likelihood. It is the best we can do in any case.

2.4 The role of Non-Sinitic languages in OC reconstruction.

To reconstruct OC on its own merits without consideration of cognate TB languages is in principle a sound *first* step. But in the end OC has to fit into the overall frame of Sino-Tibetan languages. As Sagart points out: "There is also no question that advances in TB can help students of Chinese historical phonology constrain their hypotheses on the early history of Chinese."[39] Contrary to this admonition, B&S avoid TB, yet make liberal use of what they assume to be OC loan words in southern languages. Loan words are treated as if they were OC words, without clarifying two problems:

(1) The direction of borrowing was certainly not only one way from Chinese to others, especially not in the early period when China was expanding and absorbing other peoples with their languages. E.g. 'dog': the original Sino-Tibetan (ST) word in OC was *quǎn* 犬 *khwînʔ, *gǒu* 狗 is not a ST word. As B&S correctly point out, 狗 also shows up in Miao-Yao 苗瑤 (MY) languages.[40] B&S, as always, claim that MY has borrowed the word from OC (NOC *Cə.kˤroʔ), but I am sure this non-ST word has been absorbed with MY people and their

---

39 Laurent Sagart, "Review of Matisoff Handbook of Tibeto-Burman," *Diachronica* 23, 1(2006):221.

40 Baxter and Sagart, *Old Chinese*, 215.

language when Zhou period China expanded south. The OC form was *kloʔ (cf. Yao *klo^B 'dog'; see also Ostapirat),[41] a foreign medial *l always results in a MC Div. I/IV type syllable (Schuessler; NOC has an oddly unbalanced distribution of phonemes, medial /r/ abound, medial /l/ does not exist; and yes, here I project a foreign medial /l/ into OC because it left a trace in MC Div. I/IV).[42]

Baxter & Sagart: NOC *Cə.kˤroʔ
        >> Miao-Yao *qluwX, Proto-Yao *klo^B
Schuessler (also Ostapirat):   MY *k̠lu^B (?), Proto-Yao *klo^B 'dog'
        >> late OC *kloʔ

(2) Even if little-known southern languages like Ruc had borrowed OC words (how exactly was that supposed to have happened?), one first has to exclude nativization, i.e. addition of native prefixes etc., before projecting features (prefixes) of outlier languages straight back into OC.

3. Some Details

Although B&S's erudition is impressive and disarming, upon closer consideration so many of their assertions and hypotheses (though thought-provoking as suggestions) are questionable, so that I cannot help but be skeptical and distrustful of anything in their book. Yet generally, I trust the material that has been carried over from Baxter 1992 (with updates and modifications). Therefore, one might as well go straight back to Baxter 1992.

3.1 Interpretation of graphs.

When MC homophones are distinguished by two different graphs, one can suspect that these *might* have been phonetically different in OC. Here I agree with B&S. A good example is *yáng* 羊 vs. 昜陽 which were OC *jaŋ vs. *laŋ (the former is now NOC *ɢaŋ).

Sometimes, the difference, if any, as in wǔ 五 MC ŋuo´ vs. wǔ 午

---

41 Weera Ostapirat, "Issues in the reconstruction and affiliation of Proto-Miao-Yao," *The 14th International Symposium on Chinese Languages and Linguistics* (IsCLL-14), (Taipei: Academia Sinica, 2015), 357.

42 Axel Schuessler, *ABC Etymological Dictionary of Old Chinese* (Honolulu: Hawai'i University Press, 2007).

MC ŋuoˊ is difficult to identify. Again with B&S I would first explore the possibility of a phonetic distinction here. The *xiéshēng* series are no help because all words written with these graphic elements uniformly point to an OC root *ŋa, except for *chǔ* 杵 MC t hjwoˊ 'pestle', on whose interpretation everything hinges.[43] The graph 杵 MC t hjwoˊ seems phonetically irreconcilable with the element *wǔ* 午 MC ŋuoˊ (except for the rhyme), nor is there any possibility of a semantic connection ('pestle' vs. a cyclical sign). Graphs like 杵 and for that matter 午 are like Rorschach tests (羅夏墨迹測驗) where the interpretation may reveal more about the viewer than about OC.

Here B&S avail themselves of their uvular hypothesis to explain the difference between 五 *ŋˤaʔ vs. 午 *[m].qʰˤaʔ on the basis of their reconstruction of 杵 as *t.qʰaʔ (from where they take the q for *[m].qʰˤaʔ —?). But MC homophonous graphs must *not* necessarily go back to different OC syllables, as in this case where members of a word family 'to face, go against, oppose' seem to mix freely with both phonetic elements (contra B&S, as far as I can see):

wǔ *ŋâʔ 午 'resist' [*Lǐjì*]; wǔ, wù<*ŋâʔ, *ŋâ-s 仵 'equal' [*Zhuāngzǐ*]
wù *ŋâ-s 晤 'to meet, face to face' [*Shījīng*], 五 'to meet' [OB—Xú Zhōngshū, *Jiǎgǔwén zìdiǎn* 1528], 捂 'turn against' [*Yílǐ*], 悟 'go against' [*Lǚshì chūnqiū*], 'encounter' [*Chǔcí*], 迕 'go against' [*Lièzǐ*], 忤 'oppose' [*Hánfēizǐ*]
yǔ *ŋaʔ 禦 'defend, oppose, prevent' [*Shījīng*, BI]

Also cognates to TB words with initial *ŋ- are found written with both 五 and 午 (this is an illustration for how TB can "help... constrain... hypotheses on the early history of Chinese"):

wú *ŋâ 吾 'I, me' = TB *ŋa, Written Tibetan (WT) ŋa
wǔ *ŋâʔ 五 'five' = TB *(-)ŋa, WT lŋa
yà *ŋrâ-s 御 (迓) 'to meet, welcome, provide against' [*Shījīng*]
   = TB *ŋra, Written Burmese ŋraᴮ 'meet with, find'
yù *ŋa-s 御 'manage, serve, superintend, drive a chariot' = WT *mŋa'-ba* 'having, owning', *mŋa'* 'might, dominion' (i.e. 'having power over, control'), *mŋag-pa* 'to charge, send, serve'.
hǔ *hŋâʔ 滸 'river bank' [*Shījīng*], cf. WT dŋo 'shore, bank' (occasionally, Tibetan has the vowel /o/ for other languages' /a/).

---

43 Baxter and Sagart, *Old Chinese*, 128ff.

When words are consistently written with either one or the other as phonetic, that does not exclude homophony, but may reflect writing traditions. Anyone who has learned English will automatically write the /f/-sound in 'enough' with "gh", the /f/ in bluff with "f" and not confuse the two.

As to chǔ 杵 MC tśhjwoˊ, the correspondence of initial consonants with 午 does not follow a recognizable pattern, it is unique, needs therefore to be set aside as irregular. Doing otherwise amounts to phonetic over-interpretation. In any case, the phonetic complication in the phonetic series 午 is to be found in the single odd graph 杵 , and not in the 8 or so other graphs written with 午 where everything points to a uncomplicated syllable type *ŋa. According to the rule of parsimony, it makes more sense to concern ourselves with the one word *chǔ*, rather than trying to bring the other 8 words written with 午 in line with this one exception 杵 . Of course, the rule of parsimony does not always apply, but unless we have better evidence (data), this is the best we can do.

We see here, as elsewhere in Sagart's work and in B&S, how the issue is framed backwards. They pose the question: because 杵 has MC initial tśhj-, what was the initial of 午 ? Instead most investigators would begin by asking: because all words written with 午 have MC initial *ŋ, why/how is the initial in 杵 irregular? In other words: the authors assume that the exception 杵 is original or regular, the bulk of the data 午 must therefore be reinterpreted. Another example of this below.

3.2 Etymology vs. mental association.

Anyone studying etymology can easily be misled by mental association — one's own or those of others, even authors thousands of years ago. Like my own work, B&S is no exception; this concerns both morphology, and interpretation of graphs and their phonology.

Take for example *yǐng* 影 NOC *qraŋʔ 'shadow' which has *jǐng* 景 *C.qraŋʔ 'bright' as "phonetic". The authors claim that the words are derived from the same root, thus confirming their initial *q in these words.[44] But how could a word for 'bright' be derived from 'shadow'? Would *C- turn a word into its semantic opposite? (In my 2015 review I accidentally interchanged the NOC forms with the graphs.)[45] Hardly. For me, the OC forms are clearly 景 *kraŋʔ and 影 *ʔraŋʔ, 景 serves as a *partial* phonetic through *mental* association of the notion 'shadow' with 'bright' in the minds of the creators of the Chinese script and modern physicists alike.

---

44 Baxter and Sagart, *Old Chinese*, 28.

45 Axel Schuessler, "New Old Chinese," *Diachronica* 32, 4 (2015): 571–598.

Other example (not from B&S). It is tempting to consider a number of words genetically related that refer to the 'young of animals', so for example Wáng Lì: [46]

gǒu 狗   *kôʔ = *kloʔ 'puppy dog, dog', in eastern dialects gāo 羔 *kâu
= gǒu 豿   *kôʔ 'young of bears or tigers', is an extension of 'puppy'
jū 駒   *ko 'young horse'
hǒu 犼   *hôʔ 'calf'
gāo 羔   *kâu 'lamb'; in eastern dialects 'young of bears or tigers'

All these words look similar. But if they were etymologically related, then *ô MC Division I (一等 -ou vs. *-ju 三等) must mean 'relating to dogs' etc., *-âu must mean 'relating to sheep', initial *x- must mean 'relating to cows', and the unmodified form *ko must mean 'relating to horses'. Thus the words are unrelated. Inventors of writing may make here mental connections between unrelated words because these happen to share a logical category as well as sounds, and which are therefore written with the same phonetic 句 .

There is a simple rule of thumb, although it is not fool proof: When the items that happen to sound similar would be found on the same page in a biology-, anatomy- or physics (etc.) handbook, an etymological link is probably false, because strong mental categorization is an easy trap for etymologizing, note also the example of 'blood' and 'vein' below. In reality, concepts associated with certain objects or phenomena would be *distinguished* by being derived from *different* roots. Various parts of the body would not be derived from the same root (unless by transparent morphemes).

3.3 Illustration: 'Blood'

B&S offer the curious OC form *m̥ˤ(r)ik for xiě 血 , MC xiwet 'blood', which has been taken over from Sagart.[47] A reader may be disarmed by the erudition and sophisticated reasoning of a well-known author. It is an example for Sagart's and B&S's mode of operation that is encountered throughout the book.

Here are the words written with 血 as phonetic (or "phonetic") element

---

46 Wáng Lì, "*Tóngyuán zìdiǎn*", 182–183.

47 Baxter and Sagart, *Old Chinese*, 240; Laurent Sagart, "The Chinese and Tibeto-Burman Words for 'Blood,'" in *In Honor of Mei Tsu-Lin: Studies in Historical Syntax and Morphology*, eds. Alain Peyraube and Sun Chaofen (Pairs: Le Centre de recherches linguistique sur l'Asie orientale, EHESS, 1999), 165–181.

(by "Han" I suggest a likely Han-period phonemic form; OC = OCM from Schuessler 2009,[48] unless marked "NOC"):

血 Han huet, *hwît   'blood', cf. TB *s-hywəy, Proto-Lolo *swi, Magarihyu<hwi

洫 Han huit, *hwit   'moat' (MC xjwək)

恤卹 Han suit, *swit   'care about'

卹 Han suət, *sût   'rub, brush' [Lǐjì 禮記], cf. Lolo-Burmese *sut 'wipe, sweep'

洫 Han huɨk, huɨ`, *hwək, *hwək-s (?)   'still, quiet' [Shījīng], not a rhyme word

殈 Han huek, hyek, *hwêk, *hwek [Lǐjì] = 㘸 [Zhuāngzǐ]

The *xiéshēng* graphs indicate that the root initial of most words was *w (*sw-, *hw-), the rhymes suggest that *xiě* 血 (MC xiwet) was something like OC *hwit (so Baxter, or OCM *hwît),[49] confirmed by all Sinitic languages, dictionaries and texts. This word is obviously related to TB *s-hywəy or *hywəy (Matisoff 50; perhaps the TB word was simply *s-wi> *hwi). Chinese words occasionally have a final -t where TB counterparts end in an open vowel, e.g. also *rì* 日 *nit 'day, sun' = TB *ni. The case should be relatively clear-cut and settled.

In B&S, no explanations are provided for NOC *m̥ˤ(r)ik, as usual. I would have to accept their assertions on faith (contrary to Sir Francis Bacon's warning never to blindly accept authority), or dig out the background information myself, which, when I take the trouble, makes the B&S claims all too often appear of the "maybe/maybe-not" type, if not entirely unconvincing.

Now let us take a look at how Sagart arrives at this reconstruction.

3.3.1 Sagart's initial OC *m-. Two pieces of information persuade him that the initial of 血 was *m- (or rather *m̥-). Actually they involve the word 恤 MC sjwet, and 血 only indirectly by association.

(1) In a *Shūjīng* 書經 passage, one version's word *xù* 恤 MC sjwet (Karlgren 'carefully', Legge 'compassion') has been substituted in another by *mì* 謐 MC mjet 'gentle, mild'. This is taken as evidence of *m- in 恤.

---

48   Schuessler, *Minimal Old Chinese*.

49   Baxter, *A Handbook of Old Chinese Phonology*.

50   James A. Matisoff, *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction* (Berkeley: University of California Press, 2003).

Whenever possible, B&S interpret a graph phonetically; to me this looks like a type of phonetic over-interpretation of a graph. It happens in most or all languages that in texts one word is substituted for another whose meaning would fit the context. Chinese literature is full of examples (e.g. *Shījīng* quotes in *Zuǒzhuàn*). Therefore it is highly likely that in 恤 and 謐, we have here one of those associative substitutions. Before phonetic similarity is assumed with subsequent NOC reconstructions, this kind of substitution must be firmly ruled out. This has not even been attempted.

(2) The second piece of evidence for *m- in *xù* 恤 is provided by the *Shìmíng* 釋名 which glosses *xū* 戌 MC sjwet 'cyclical sign' with 恤 MC sjwet. Such a Han 漢 period gloss is of little or no help for OC, for two reasons: (a) Han period homophony in Liu Xi's (*Shìmíng*) time does not prove at all that 恤 and 戌 were homophones earlier. (b) Even if these words were homophones in OC, works like *Shìmíng* were collections of puns, the phonological parameters of which, if any, are unknown (cf. Coblin, n.d.).[51] Thus the *Shìmíng* proves nothing for OC. — As a consequence, all the considerations of additional evidence for *m- in 戌 (威 MC mjät in which 戌 seems to be phonetic, and has the Tai language Ahom equivalent *mit*. All this points to an OC *sm-cluster in 戌, according to Bodman[52]) are beside the point, because a possible *sm-initial in 戌 has no bearing on 恤 and 血 anyway.

And all these arguments only touch 血 indirectly and may even be irrelevant for 血.

Then Sagart relates *Shuōwén jiězì* 說文解字 *miè* 㵏 'defile with blood', *Guǎngyǎ* 廣雅 'blood' to 血 as further evidence for *m-. The former is not an OC word, as far as we can tell, it is apparently only known from later dictionaries. In addition Sagart cites a Tujia word *mie 35* 'blood'. The assumed etymological connection of these words with 血 should have no bearing on the OC word for *xiě*. These marginal words are adduced only as validation of a theory (confirmation fallacy).

An obstacle to this theory on 血 'blood' should be TB. But in his 1999 paper Sagart claims that much or most of the TB vocabulary consists of Chinese loans.[53] Therefore TB *hywəy must have been borrowed from later (post OC) Chinese (this requires yet another additional theory that OC final

---

51 W. South Coblin, "Beyond BTD: An excursion in Han Phonology," Unpublished Manuscript, n.d. PDF file.

52 Nicholas C. Bodman, *A Linguistic Study of the Shih Ming* (Cambridge, Mass.: Harvard University Press, 1954), 63,102.

53 Sagart, "The Chinese and Tibeto-Burman Words for 'Blood,'" 165–181.

*-t < *-k was lost in TB). I am among the many who believe in the plausibility of the Sino-Tibetan hypothesis, I find it difficult to imagine why and how in early dynastic times an OC lexical wave would have swept all the way into Southeast Asia and the Himalayas.

3.3.2 Sagart assumes the rhyme *-ik for 血 . As far as I can see, the linguistic data we have point consistently to the rhyme *-it (OC *hwit) in *Old Chinese*, whatever the rhyme may have been in Proto-Chinese. Again, I prefer to rely on linguistic evidence first, regardless what the *xiéshēng* indicate, for whatever reasons.

B&S suggest that final *-ik (B&S 血 *m̥(r)ik) is confirmed by cognancy with *mài* 脈 *mrêk (B&S *C.m <r>[i]k) 'vein' (why "[i]" and not "i" which they must have had in mind in order to establish the etymological link?).[54] This is the same confusion of etymology with mental association again that I have pointed out above (*gǒu* 狗), because NOC *xiě* and *mài*, being somewhat similar in the NOC reconstruction, would appear on the same page in a physiology textbook.

3.3.3 To conclude: B&S presents 血 OC *m̥ˤ(r)ik as if it were a fact.[55] I see no reason why 血 should be anything other than OC *hwît (or *hwit or how ever one wants to transcribe it).

4. Conclusion

To summarize. From my perspective, I find the NOC reconstructions and their rationales problematic, for several reasons, among them:

- Framing of the issue is backward, i.e. starting from the unusual and marginal (杵 and 午 ; the uvular hypothesis; marginal data are to validate a theory, e.g. 蠛 , Tujia *mie 35*).
- Etymology tends to be a somewhat subjective field (澠 ~ 減 , 血 ~ 脈 , 血 ~ 蠛 ; 影 ~ 景 , 圉圍).
- Phonetic over-interpretation of graphs (恤 , 血 ; 影 ~ 景 , 杵 ~ 午).
- Stuffing "reconstructions" with alternative features (犬 NOC *[k]ʷʰˤ[e][n]ʔ)

---

54 Baxter and Sagart, *Old Chinese*, 240.

55 Ibid.

Every student of OC has his own ideas and system, and disagrees with others. That is the nature of this highly interpretive field. Often other scholars' ideas are so incompatible with one's own that one can conveniently ignore them, it seems B&S's work has the great merit of offering countless stimulating suggestions, but from my perspective it does not hold up to close scrutiny as a plausible reconstruction of the OC language.

I have not combed through B&S hunting for the occasional item I can find fault with. Unfortunately, I run into problems and questions of the sort discussed above throughout the book, almost on page after page, even item after item — claims that at best may or may not be true. I do not find NOC a useful system. There are alternative and more plausible solutions (but by no means definitive) ones, such as Baxter 1992 (or a variant of it: Schuessler 2009).

Works Cited

Baxter, William H. *A Handbook of Old Chinese Phonology*. Berlin/New York: Mouton de Gruyter, 1992.

——and Laurent Sagart. *Old Chinese: A New Reconstruction*. Oxford: Oxford University Press, 2014.

Bodman, Nicholas C. *A Linguistic Study of the Shih Ming*. Cambridge, Mass.: Harvard University Press, 1954.

Coblin, W. South. "Notes on the Dialect of the Han Buddhist Transcriptions." In *Proceedings of the International Conference on Sinology*, 121–183. Taipei: Academia Sinica, 1981.

——. *Studies in Old Northwest Chinese*. Journal of Chinese Linguistics Monograph 4. Berkeley, California: Project on Linguistic Analysis, 1991.

——. *A Handbook of 'Phags-pa Chinese*. Honolulu: Hawai' i University Press, 2007.

——. "Beyond BTD: An excursion in Han Phonology." Unpublished Manuscript, n.d. PDF file.

Karlgren, Bernhard. *Grammata Serica Recensa*. Stockholm: Museum of Far Eastern Antiquities, 1957.

Lehmann, Winfred P. *Theoretical Bases of Indo-European Linguistics*. London/New York: Routledge, 1993.

Li Fang-kuei 李方桂. "Shànggǔ yīn yánjiū" 上古音研究. *Tsing Hua Journal of Chinese Studies n.s.* 9 (1971): 1–61.

Matisoff, James A. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press, 2003.

Norman, Jerry. "Common Dialectal Chinese." In *The Chinese Rime Tables. Linguistic Philosophy and Historical-Comparative Phonology*, edited by Davied P. Branner, 233–254. Amsterdam/Philadelphia: John Benjamins Publishing Company, 2006.

Ostapirat, Weera. "Issues in the Reconstruction and Affiliation of Proto-Miao-Yao." *The 14th International Symposium on Chinese Languages and Linguistics (IsCLL-14),* 347–358. Taipei: Academia Sinica, 2015.

Pān Wùyún 潘悟雲. "Hóu yīn kǎo" 喉音考. *Mínzú yǔwén* 民族語文 5 (1997):10–24.

——. *Hànyǔ lìshǐ yīnyùnxué* 漢語歷史音韻學. Shànghǎi: Shànghǎi jiàoyù chūbǎnshè, 2000.

Sagart, Laurent. *The Roots of Old Chinese*. Amsterdam/Philadelphia: John Benjamins Publishing Company, 1999.

————. "The Chinese and Tibeto-Burman Words for 'Blood.' " In *In Honor of Mei Tsu-Lin: Studies in Historical Syntax and Morphology*, edited by Alain Peyraube and Sun Cahofen, 165–181. Paris: Le Centre de recherches linguistique sur l'Asie orientale, EHESS, 1999.

————. "Review of Matisoff Handbook of Tibeto-Burman." *Diachronica* 23, 1 (2006): 206–223.

Schuessler, Axel. "R and L in Archaic Chinese." *Journal of Chinese Linguistics* 2, 2(1974): 186–199.

————. *A Dictionary of Early Zhou Chinese*. Honolulu: Hawai'i University Press, 1987.

————. *ABC Etymological Dictionary of Old Chinese*. Honolulu: Hawai'i University Press, 2007.

————. *Minimal Old Chinese and Later Han Chinese: A Companion to Grammata Serica Recensa*. Honolulu: Hawai'i University Press, 2009.

————. "New Old Chinese." *Diachronica* 32, 4 (2015): 571–598.

Szemerényi, Oswald J. L. *Introduction to Indo-European Linguistics*. 4th ed. Oxford: Oxford University Press, 1996.

Wáng Lì 王力. *Tóngyuán zìdǐian* 同源字典. Běijīng: Shāngwù yìnshūguǎn, 1982.

Zhèngzhāng Shàngfāng 鄭張尚芳. *Shànggǔ yīnxì* 上古音系. Shànghǎi: Shànghǎi jiàoyù chūbǎnshè, 2003.

「白—沙」的上古漢語語音構擬若干問題

許思萊

美國沃特堡學院

　　本文就上古漢語語音構擬，尤其是白一平、沙加爾於 2014 年提出的構擬方法提出若干問題，並就「白—沙」體系中的具體例證進行討論。

關鍵詞： 上古音構擬　比較方法　假設—推演方法　白一平 & 沙加爾　語源學　諧聲系統