

# How Old is the Chinese in Bái?

Reexamining Sino-Bái under the Baxter-Sagart reconstruction

白-沙汉语构拟视角下的汉源白语音系

Xun GONG

CRLAO, EHESS/INALCO, Paris

*Workshop 'Recent advances in Old Chinese Historical Phonology'*

*SOAS, 5 – 6 November 2015*

## Topics

- The position of Sino-Bái: can the Chinese-like vocabulary of Bái legitimately inform the reconstruction of Old Chinese?
- Introducing a new reconstruction of Proto-Bái
  - Compared to the wealth of initial clusters reconstructed by (Wāng, 2006; Yuán, 2006), this reconstruction has a CGVC, modern-Chinese type of phonology.
  - From Pre-Proto-Bái to Proto-Bái
    - \*  $*\text{æ} < \text{Pre-Bái } *rV, \text{Pre-Bái } *e(\eta, k)$ .
    - \*  $*iaC < *æt$
- Comparing the new Proto-Bái with the new Old Chinese
  - Bái systematically reflects the contrast between Baxter-Sagart  $*r-$  vs.  $*C.r-$ , and probably also  $*r̥-$ .
  - A few Bái words show uvular stops in Division-III syllables. Possible correlation with Baxter-Sagart uvulars.
  - Individual forms of interest: Bái  $*piN^1$  'wind' < 風 OC<sup>1</sup>  $*prəm$ , Bái  $*tʃ^hæN^1$  <  $*tʃ^heŋ^1$  'light' < 輕 OC  $*[k^h]eŋ$ .
  - For cases where 3th-century Mainstream Chinese differs dialectally from Shāndōng Chinese, current Mainstream Chinese always shows the Shāndōng forms while Bái consistently shows 3th-century Mainstream forms.

---

<sup>1</sup>I write 'OC' when I feel the Baxter-Sagart reconstruction is consensus, and 'Baxter-Sagart' when it differs significantly from its precedents.

## 1 Introduction

- Bái vocabulary, later Chinese borrowings excluded, can be divided into a Sino-Bái vocabulary, which shows a close resemblance to Chinese, and a Pre-Bái or Eteo-Bái vocabulary, which is not etymologizable into Chinese. There are hence three possible positions on the genetic position of Bái:
  - Non-Sinitic: Bái is a non-Sinitic language (probably Qiāngic, cf. the ethnonym Bái = Proto-Bái *\*bæC* ‘white’, like the local Qiāngic speaking peoples in contrast to the Loloish-speaking peoples who call themselves blacks) with a huge amount of Chinese loanwords.
  - Sinitic: Bái is a Chinese dialect.
  - Sino-Bái group: Bái is a close cousin to Chinese and contains important archaisms not reconstructible for Old Chinese.
- In support of the third position, Wāng (2006) invoked what he termed the *Inexplicability Principle* (Wāng, 2006, p. 2):

(iii) By **Inexplicability Principle** we mean the inability to describe a recipient language in terms of the phonological system of the donor language. The inexplicable elements by borrowing are considered to be inherited from the ancestor language.

Elements from his Proto-Bái, inexplicable from the Chinese comparanda, are interpreted as evidence of Pre-OC Bái archaism.

- Lexical archaism of Bái. These everyday Bái words, identical to ‘school’ Classical Chinese (not obscure terms attested only in the lexica), are unattested in any of the 42 dialects in Lǐ et al. (2002). Taiwanese and Wēnzhōu Wú are given as examples of conservative Southern dialects.

Gloss	Bái	Mandarin	Taiwanese	Wēnzhōu
firewood	*siN 薪 till 13c.	chái 柴	chhâ 柴 frequent from 7c.	za∨ 柴
steal	*daH 盜 3c. BC – 4c.	tōu 偷	thau 偷 as ‘to steal’ from 1c.	t <sup>h</sup> au-∨ 偷
put on	*ʔiH 衣 till 1c.	chuān 穿 from 6c.	chhēng ?? etym. unk.	tɕa∨ 著 from 2c. BC
sleep	*ts <sup>h</sup> iNX 寢 till 2c. BC	shuì 睡 from 10c.	khùn 困 from 11c.	k <sup>h</sup> y∨ 困

## 2 Memoranda

- Sources: Wāng (2006) (w̄), Yuán (2006) (ȳ), the Bái-Chinese dictionary (bh). Both the dialect transcriptions and the reconstructions of Wāng and Yuán show a high degree correspondence to each other, which testifies to the reliability of the fieldwork.
- Notation: \*kr-: \*kwr- means \*kr- (Wāng, 2006) and \*kwr- (Yuán, 2006).

- Dialectology (Xú and Zhào, 1984; Wāng, 2006, 2013):
  - **Bái proper:** (= Eastern Bái, Wāng, 2006) spoken by easterly Bái people, called Míngjiā in Chinese, about 955 600 in 1984.
    - \* **Central Bái:** (= Jiànchuān dialect & Dàlǐ dialect, Xú and Zhào, 1984) One single language with little diversity. The written language is standardized around Jīnhuá dialect.
    - \* **Mǎzhěnlóng Bái:** deviant dialect closely related to Bái proper, spoken by people called Míngjiā in Mǎzhěnlóng (文山州丘北县双龙营镇马者龙村), ~500km southwest from Bái-speaking territory.
    - \* (other diaspora Bái dialects)
  - **Western Bái:** (Wāng, 2006, = Bijiāng dialect, Xú and Zhào, 1984) spoken by westerly Bái people, called Lè mò, Láibǔ, Nàmǎ, Láimò, etc. Large internal diversity, likely several mutually unintelligible independent languages.
  - **Càijiā** (Lǚ Shānshān, p.c.): large Sino-Bái-like vocabulary, but preserves phonological archaisms disappeared in Bái dialects surveyed.
- Tone
  - The tone correspondence between dialects is clear;
  - Different treatment in the comparison to Chinese: Yuán (2006), following Zhèngzhāng Shàngfāng ≠ Wāng (2006).
  - For now I follow Yuán (2006).
  - The tone letters are from the 1984 orthography, used notably in Zhào and Xú (1996). High-register qù tone merged with low-register checked tone in Central Bái, but they are distinguished in some Western Bái dialects, for example EQ<sub>w</sub>=Yíngpán<sub>y</sub>.

Chinese	Class	Notation	Central Bái	Yíngpán
平 (*-Ø)	*C̣-	*-Ø	⊔ modal ( <i>l</i> )	⊔
	*C̣-		⊓ harsh ( <i>Ø</i> )	⊓
上 (*-ʔ)	any	*-X	⊖ modal ( <i>x</i> )	⊖
去 (*-s)	*C̣-	*-H	⊓ tense ( <i>rt</i> )	⊓
	*C̣-		⊓ breathy ( <i>t</i> )	
入 (*-T)	*C̣-	*-C	⊖ tense ( <i>rx</i> )	⊖
	*C̣-		⊓ tense ( <i>rt</i> )	⊓

### 3 A revised Proto-Bái

#### 3.1 The cluster-rich 不谋而合

Wāng (2006) and Yuán (2006) are independently executed. Both, though, reconstructed a large number of consonant clusters.

Wāng (2006):						Yuán (2006)											
						<i>p</i>	<i>p<sup>h</sup></i>	<i>b</i>	<i>mb</i>	<i>m</i>	<i>hm</i>						
						<i>pr</i>	<i>p<sup>h</sup>r</i>	<i>br</i>			<i>hmr</i>						
						<i>pj</i>		<i>bj</i>	<i>?bj</i>								
<i>p</i>	<i>p<sup>h</sup></i>	<i>b</i>	<i>m</i>	<i>m-p</i>			<i>brj</i>	<i>sbr</i>									
<i>pr</i>	<i>p<sup>h</sup>r</i>	<i>br</i>				<i>kp</i>	<i>kp<sup>h</sup></i>	<i>gb</i>	<i>km</i>								
<i>pj</i>	<i>p<sup>h</sup>j</i>	<i>bj</i>	<i>mj</i>			<i>t</i>	<i>t<sup>h</sup></i>	<i>d</i>	<i>n</i>	<i>hn</i>							
<i>t</i>	<i>t<sup>h</sup></i>	<i>d</i>	<i>n</i>	<i>l</i>		<i>tr</i>	<i>t<sup>h</sup>r</i>	<i>dr</i>									
<i>tr</i>	<i>t<sup>h</sup>r</i>	<i>dr</i>				<i>str</i>	<i>st<sup>h</sup>r</i>	<i>sdr</i>									
<i>tj</i>	<i>t<sup>h</sup>j</i>	<i>dj</i>	<i>nj</i>			<i>ts</i>	<i>ts<sup>h</sup></i>	<i>dz</i>	<i>s</i>	<i>z</i>	<i>hs</i>						
<i>ts</i>	<i>ts<sup>h</sup></i>	<i>s</i>	<i>s<sup>h</sup></i>	<i>z</i>	<i>dz</i>	<i>tsr</i>	<i>ts<sup>h</sup>r</i>	<i>dzr</i>	<i>sr</i>	<i>zr</i>							
<i>tsr</i>	<i>ts<sup>h</sup>r</i>	<i>sr</i>	<i>s<sup>h</sup>r</i>	<i>zr</i>	<i>dzr</i>	<i>tɕ</i>	<i>tɕ<sup>h</sup></i>	<i>dɕ</i>	<i>ɕ</i>	<i>ɲ</i>	<i>hɲ</i>	<i>j</i>	<i>?j</i>				
<i>tsj</i>	<i>ts<sup>h</sup>j</i>	<i>sj</i>				<i>k</i>	<i>k<sup>h</sup></i>	<i>g</i>	<i>ŋ</i>	<i>hŋ</i>	<i>ɣ</i>						
<i>k</i>	<i>k<sup>h</sup></i>	<i>g</i>	<i>ŋ</i>	<i>x</i>	<i>ɣ</i>	<i>kw</i>	<i>k<sup>h</sup>w</i>	<i>gw</i>	<i>ŋw</i>	<i>hŋw</i>							
<i>kr</i>	<i>k<sup>h</sup>r</i>	<i>gr</i>				<i>kfi</i>		<i>mg(w)</i>									
<i>q</i>	<i>q<sup>h</sup></i>	<i>ɕ</i>	<i>ɕ<sup>h</sup></i>	<i>N</i>	<i>ɸ</i>	<i>kwr</i>	<i>k<sup>h</sup>wr</i>	<i>gwr</i>									
<i>tɕ</i>	<i>tɕ<sup>h</sup></i>	<i>ɕ</i>	<i>ɕ<sup>h</sup></i>		<i>dɕ</i>	<i>q</i>	<i>q<sup>h</sup></i>	<i>ɕ</i>	<i>ɕ</i>								
<i>h</i>	<i>?</i>					<i>qw</i>	<i>q<sup>h</sup>w</i>										
								<i>mfiɕ</i>									
						<i>?</i>	<i>h</i>	<i>fi</i>									
						<i>?w</i>	<i>hw</i>										

The most important clusters in the two reconstructions correspond with each other.

<i>Wāng</i>	<i>Yuán</i>	<i>Mine</i>
<i>*Pr-</i>	<i>*Pr-</i>	<i>*P-(i)</i>
<i>*Pj-</i>	<i>*Pj-</i>	<i>*P-(iu,y)</i>
<i>*Tr-</i>	<i>*sTr-</i>	<i>*Tj-</i>
<i>*TSr-</i>	<i>*TSr-</i>	<i>*Tj-</i>
<i>*TSr-</i>	<i>*Tɕ-</i>	<i>*Tj-</i>
<i>*Kr-</i>	<i>*Kwr-</i>	<i>*K-(y)</i>

### 3.2 *\*Pr-: \*Pr-*

Gloss	Bái Proper		Western Bái		Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w</sub>	LBZ <sub>y</sub>	GX <sub>w</sub>	Wāng	Yuán	Mine	
wind	<i>pi</i>	<i>pien</i>	<i>tɕui</i>	<i>tɕyen</i>	<i>*pren</i>	<i>*priN</i>	<i>*pin</i>	風 <i>piuŋ</i> < <i>*prəm</i>
lung	<i>p<sup>h</sup>ia</i>	<i>phio</i>	<i>tɕ<sup>h</sup>ua</i>	<i>tɕ<sup>h</sup>ua</i>	<i>*p<sup>h</sup>raC</i>	<i>p<sup>h</sup>raS</i>	<i>*p<sup>h</sup>iaC</i>	肺 <i>phiɹjH</i> < <i>*pho[t]s</i>
low	<i>pi</i>	<i>pi</i>	<i>dzui</i>	<i>dzy</i>	<i>*briX</i>	<i>briX</i>	<i>*biX</i>	庫 <i>bjieX</i> < <i>*N-pe?</i>

*\*Pr-: Pi-* in Central Bái, but *Tɕy-* or *Tɕu-* in Western Bái dialects. Some Chinese comparanda are reconstructed with *\*-r-*:

- *\*pren: \*priN* ‘wind’ < 風 OC *\*prəm*;

- *\*praC: \*praS* ‘eight’ < 八 OC *\*p<sup>h</sup>ret*.

Most, however, are not (cf. also examples above):

- *\*bræC²: \*braS* ‘to bark’ < 吠 OC *\*Cə.bo[t]s*;
- *\*priC: \*priS* ‘female sexual organ’ < 屖 OC *\*pij<sup>3</sup>*.

Wāng (2006, p. 170) himself recognized that many of his Proto-Bái *\*Pr-* cannot be explained by the Chinese, a fact used as one of his arguments for the position that Bái cannot derive from Old Chinese.

This reconstruction is a greater problem for Yuán, who also reconstructed a *\*-re-* (*\*-raS* in checked syllables) for the vowel  $\text{æ} \sim \text{ə}$  in Central Bái, where the *\*-r-* is etymologically justified (see 3.6). The difference between 八 and 𠂔, Central Bái *pia* vs *pæ*, becomes that between *\*Pr-aS* and *\*P-raS*, a not improbable one, but to be strongly dispreferred.

The giveaway: a majority of words with *\*Pr-* (9 among 15 words shared by Wāng and Yuán) have *i*-type rhymes: *\*-i: \*i*, *\*-en: \*iN*, *\*-i: \*iS*. On the other hand, no words are reconstructed with these rhymes with *\*P-: \*P-*.

Hypothesis: what is reconstructed as *\*Pr-: \*Pr-* is nothing but *\*P-: \*P-* before *i*. In Western Bái, there was a very common sound change *\*Pi > Tçui > Tşui*.

For words like *\*praC: \*praS* ‘eight’, I reconstruct *\*piaC*, i.e. with an initial consonant *\*p-* and a rhyme *\*iaC*. The *\*-i-* should be phonologized as part of the rhyme, because its distribution is severely constrained by the rhyme.

The initial correspondence sets that reflect a Proto-Bái *\*i* are: *\*Pr-: \*Pr-*, *\*mr-: \*hmr-* (palatalized *\*hm-*) and *\*tç: \*tç* (palatalized *\*k-*, *\*ts-* and probably *\*t-*). They mainly occur in the following contexts (indexed by my own proto-Bái rhymes):

- *\*i*, *\*iC*, *\*iN*: of course.
- *\*io*: nothing interesting to say yet.
- *\*iæC*, *\*iæN* (always with a coda, no open-vowel *\*iæ*): comes from Pre-Bái *\*\*ek*, *\*\*eŋ*, cf. Section 3.6.
- *\*iaC*: reflects genuine etymological *ia* in many cases: *\*p<sup>h</sup>iaC* ‘lung’ < 肺 *phi<sup>h</sup>ŋH* < *\*p<sup>h</sup>jats* *\*p<sup>h</sup>o[t]s*. The remaining ones, notably *piaC* ‘eight’ < 八 MC *pæt*, OC *p<sup>h</sup>ret*, probably come from older *\*\*pæt*, cf. Section 3.7.

### 3.3 Labiodentalizing *\*Pj-: \*Pj-*

Both Wāng and Yuán correctly reconstructed *\*Pj-: \*Pj-* to account for labiodentals in most dialects and labials in Mǎzhěnlóng<sub>w</sub>. Now that *\*Pr-: \*Pr-* is revised to *\*P-* before front vowels, there is a danger of collision with their *\*Pj-: \*Pj-*.

<sup>2</sup>Error for *\*braC*.

<sup>3</sup>The absence of *\*-r-* in this word is indirectly attested in the Mandarin pronunciation of words like 卑 *pjiē* as *bēi* as opposed to *bī*.

Gloss	Bái Proper		Western Bái		Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w̄</sub>	LBZ <sub>ẏ</sub>	GX <sub>w̄</sub>	Wāng	Yuán	Mine	
fly	<i>fv</i>	<i>pi</i>	<i>fe</i>	<i>pfu</i>	<i>*pje</i>	<i>*pj-</i>	<i>*py</i>	飛 <i>piij</i> < <i>*Cə.pəj</i>
divide	<i>fv</i>	<i>pien</i>	<i>fɛ</i>	<i>pfū</i>	<i>*pjen</i>	<i>*pjiN</i>	<i>*pyN</i>	分 <i>piun</i> < <i>*pən</i>
stomach	<i>fv</i>	<i>puw</i>	<i>fo</i>	–	<i>*pjuC</i>	<i>*pjruS</i>	<i>*piuS</i>	腹 <i>piuk</i> < <i>*p(r)uk</i>
bee	<i>fv</i>	<i>p<sup>h</sup>en</i>	<i>xō</i>	<i>pfuŋ</i>	<i>*p<sup>h</sup>juŋ</i>	<i>*p<sup>h</sup>jruN</i>	<i>*p<sup>h</sup>iuN</i>	蜂 <i>p<sup>h</sup>ioŋ</i> < <i>*p<sup>h</sup>(r)oŋ</i>

Fortunately, labiodentalizing words always have *v* in Central Bái, so from the Central Bái point of view, they are in complementary distribution with the palatalized *\*P*-series. The Proto-Bái rhymes, on the other hand, need to be revised, cf. *\*briX*:*\*briX* ‘low’, *pren*:*priN* ‘wind’ above.

In non-labial words, Central Bái *v* are mostly reconstructed as *\*o*:*\*rwu*. The Chinese comparanda, however, sound *y*-ish – always Division-III with late Middle Chinese rounded elements: *\*njoX*:*\*nrwuX* ‘woman’ < 女, *\*ŋoA*:*\*hŋrwu* ‘fish’ < 魚, *\*sroA*:*\*srrwu* < 鼠, Written Jiànchuān *gv<sup>rt</sup>ze<sup>l</sup>* (*kv.tsu*)<sub>bh</sub> for Guìzhōu 貴州 – the same range of Middle Chinese rhymes as the labial words here.

Solution (which looks rather well cf. the etymologies):

- Central Bái *v* < *\*y* when they have *i*-ish reflexes in dialects (Wāng *\*e* or *\*ɛ*, Yuán *\*i*);
  - *\*py* 飛
  - *\*pyn* 分
  - *\*pyC* 筆
  - *\*byX* 負
- Central Bái *v* < *\*iu* when they have *u*-ish reflexes in dialects (Wāng *\*o* or *\*u*, Yuán *\*u*).
  - *\*piuC* 腹
  - *\*phiuN* 蜂

Two words are reconstructed with *\*pr*:-*\*pr*- by Wāng and Yuán: ‘six’ and ‘saw’.

Gloss	Bái Proper		Western Bái		Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w̄</sub>	LBZ <sub>ẏ</sub>	GX <sub>w̄</sub>	Wāng	Yuán	Mine	
six	<i>fv</i>	<i>fv</i>	<i>fo</i>	<i>pfu</i>	<i>*pjoC</i>	<i>*pjruS</i>	<i>*piuS??</i>	六 <i>liuk</i> < <i>*k.ruk</i>
saw	<i>fv</i>	<i>fv</i>	<i>fv</i>	<i>pfu</i>	<i>*pjo<sup>5</sup></i>	<i>*pjruH</i>	<i>*piuH??</i>	鋸 <i>kiΛH</i> < <i>*k(r)as</i>

They don’t look right:

- The initial is ultimately from *\*k*- (for ‘six’, see Section 4).
- No affrication in Mǎzhěnlóng<sub>w̄</sub>. Likely genuine, as Càijiā has the same contrast: *fv* ‘six’ ≠ *pv* ‘fly’.

Would be nice to reconstruct *\*kiuC* and *kiuH*. However, there is ‘live, sit’ *\*kiuH* < 踞 *kiΛH* < *\*k(r)as* with usual out labiodentalization and with identical rhyme reflexes.

Gloss	Bái Proper		Western Bái		Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w̄</sub>	LBZ <sub>ẏ</sub>	GX <sub>w̄</sub>	Wāng	Yuán	Mine	
sit	<i>kv</i>	<i>kv</i>	<i>kv</i>	<i>kao</i>	<i>*kjo<sup>5</sup></i>	<i>*krwuH</i>	<i>*kiuH</i>	踞 <i>kiΛH</i> < <i>*k(r)as</i>

### 3.4 \*Kr-: \*Kwr- and \*r-w: \*brj-

The hypothesis that Central Bái *v* always come from either \**y* or \**iu* also permits the resolution of another problem.

Gloss	Bái Proper		Western Bái		Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w̄</sub>	LBZ <sub>ý</sub>	GX <sub>w̄</sub>	Wāng	Yuán	Mine	
solider	<i>kv</i>	–	<i>tʂɛ</i>	–	* <i>kroŋ</i>	* <i>krwiN</i>	* <i>kyn</i>	軍 <i>kiun</i> < * <i>k<sup>w</sup>ər</i>
ghost	<i>kv</i>	<i>tɕu</i>	<i>tʂɛ</i>	<i>tʂi</i>	* <i>kroX</i>	* <i>krwiX</i>	* <i>kyX</i>	鬼 <i>kiuijX</i> < * <i>k.ʔujʔ</i>
snake	<i>k<sup>h</sup>v</i>	<i>tɕ<sup>h</sup>iu</i>	<i>tʂ<sup>h</sup>ɛ</i>	<i>tʂ<sup>h</sup>i</i>	* <i>k<sup>h</sup>roX</i>	* <i>k<sup>h</sup>rwiX</i>	* <i>k<sup>h</sup>yX</i>	虺 <i>xiuijX</i> < * <i>[r<sub>o</sub>]u[j]</i> ʔ
crowd	<i>kv</i>	–	<i>tʂɛ̃</i>	–	–	* <i>grwiN</i>	* <i>gyN</i>	羣 <i>giun</i> < * <i>[g]ur</i>
rain	<i>vu</i>	<i>ju</i>	<i>dʒɛ</i>	<i>ʒi</i>	* <i>r-woX</i>	* <i>brj-</i>	* <i>yyX</i>	雨 <i>yiux</i> < * <i>C.G<sup>w</sup>(r)aʔ</i>
stomach	<i>vu</i>	<i>ju</i>	<i>ʒɛ</i>	<i>ʒi</i>	* <i>r-wuC</i>	* <i>brj-</i>	* <i>yyC</i>	胃 <i>yiuijH</i> < * <i>[G]<sup>w</sup>ə[t]s</i>

Note that for \**yyX* ‘rain’ < 雨 *yiux* and \**byX* ‘carry on the back, be in debt’ < 負 *biuwX*, the rather unexpected vowel reflex in Bái correlates with the Shàng tone.

### 3.5 \*Tr-: \*sTr-, \*TSr-: \*TSr- and \*TSr-: \*Tɕ-

TS- in Central Bái, apart from its usual origin \**TS-*: \**TS-*, shows in three different correspondence sets.

Gloss	Bái Proper		Western Bái			Proto-Bái			Chinese comparanda
	Cent.	MZL <sub>w̄</sub>	LBZ <sub>ý</sub>	GX <sub>w̄</sub>	YP <sub>ý</sub>	Wāng	Yuán	Mine	
weave	<i>tsu</i>	<i>tsəu</i>	<i>tʃu</i>	<i>tʂi</i>	<i>tiu</i>	* <i>truwC</i>	* <i>struS</i>	* <i>tʃuwC</i>	織 <i>tɕik</i> < * <i>tək</i>
needle	<i>tsi</i>	<i>tseŋ</i>	<i>tʂɛ̃</i>	<i>tʂɛ̃</i>	<i>tʂɿ</i>	* <i>tsreŋ</i>	* <i>tsriN</i>	* <i>tʂiN</i>	鍼 <i>tɕim</i> < * <i>t.[k]əm</i>
boil	<i>tsv</i>	<i>tsv</i>	<i>tɕv</i>	<i>tsao</i>	<i>tɕu</i>	* <i>tsrɔX</i>	* <i>tɕrwu</i>	* <i>tʂiu</i>	煮 <i>tɕΛX</i> < * <i>[t]a</i>

Chinese comparanda indiscriminately come from 章, 庄 and 知 series, mostly in Division III – no connection with etymological *-r-*.

Again, these proto-Bái initials show complementary distribution wrt Central Bái vowels.

	<i>e</i>	<i>ɛ</i>	<i>i</i>	<i>o</i>	<i>u</i>	<i>u</i>	<i>v</i>
* <i>Tr-</i> : * <i>Str-</i>	+	+		+	+	+	
* <i>TSr-</i> : * <i>TSr-</i>			+				
* <i>TSr-</i> : * <i>Tɕ-</i>							+

I reconstruct all of them as \**Tʃ*.<sup>4</sup>

Càijiā, more archaic than Bái, has *T<sup>h</sup>-* for etymological \**Tr-* and *TS-* for etymological \**Tɕ-*.

<sup>4</sup>In fact more likely \**Tɕ*, but there are some residue cases of \**Tɕ*: \**Tɕ*, so they are retained for the moment as \**Tɕ*, awaiting further etymological research.

### 3.6 The double origin of Proto-Bái \*æ

Yuán reconstructed \*-re, \*-reN, \*-raC for the vowel which is  $\varepsilon$  or  $\vartheta$  (depends on the dialect) in Central Bái, a correct insight whose incompatibility with the \*pr- hypothesis is now eliminated. I nevertheless reconstruct \*æ for the vowel, because a large set of \*æ come from Chinese division III/IV words with no OC \*-r-.

They compare to Chinese syllables with -eŋ and -ek:

- \*hmiæN < 名 mjieŋ, Baxter-Sagart C.meŋ;
- \*tɕiæN < 釘 teŋ, OC t<sup>ɕ</sup>eŋ;
- \*tʃ<sup>h</sup>æC < 赤 tɕ<sup>h</sup>ek, Baxter-Sagart [t-q<sup>h</sup>](r)Ak;
- \*piæC < 壁 pek, Baxter-Sagart C.p<sup>ɕ</sup>ek.

Hence, I reconstruct the sound change \*iæN < \*\*eŋ and \*iæC < \*\*ek.

- \*kæC ‘to catch’, is etymologized in (Wāng, 2006; Yuán, 2006) as 格 kæk < OC k<sup>ɕ</sup>rak, a problematic etymology:
  - k- for k<sup>ɕ</sup>- irregular;
  - 格 doesn’t mean ‘to catch’ (格物致知 notwithstanding).
- 𠵹 kiek < krak ‘to grasp’ a better match, which implies that \*kiæ became \*kæ, not \*tɕiæ < \*\*tiæ, \*\*tsiæ.
- Unfortunately a dictionary-word without real attestation.

### 3.7 Proto-Bái \*iaS < Pre-Proto-Bái \*æt

- The correspondence piaC ‘eight’ < 八 MC pɛt, OC p<sup>ɕ</sup>ret is strange: the Division-II vowel should be Proto-Bái \*æ.
- One (speculative) solution: all words with \*æC with safe etymologies come from Chinese \*rV<sup>ɕ</sup>k, never \*rVt, so Proto-Bái \*iaS < Pre-Proto-Bái \*æt.
- Examples:
  - \*piaS ‘eight’ < \*\*pæt < 八 pɛt, OC p<sup>ɕ</sup>ret
  - \*ɕaS ‘to kill’ < \*\*sæt < 殺;
  - \*tɕ<sup>h</sup>aS ‘otter’ < \*\*t<sup>h</sup>æt (or, probably, as I will argue in Section 4, \*\*k<sup>h</sup>æt) < 獺 t<sup>h</sup>æt < \*r<sup>ɕ</sup>at.
- \*\*æp should ideally pattern with \*\*æt, however, \*qæC 甲, \*gæC 夾. Probably an effect of the uvular initial.
- Note also \*ʔiaC 壓 and \*ʔaC 鴨.

## 4 OC sonorants in Bái

High-register vs. low-register sonorants in modern Bái, in all proto-tones except Shàng (-X), Yuán reconstructed \*hN- and \*N- series.

Do they correspond with the similar distinction in Southern Chinese? Seems not: both Proto-Mǐn \*N- and \*hN- series correspond to \*hN- in Bái.



- ‘grind’ 磨: Bái \*mo : Proto-Mǐn \*m-;
- ‘honey’ 蜜: Bái \*hmiC : Proto-Mǐn \*m-;
- ‘wheat’ 麥: Bái \*hmwC : Proto-Mǐn \*m-;
- ‘life’ 命: Bái \*hmiæNH : Proto-Mǐn \*m-;
- ‘face’ 面: Bái \*hmiNH : Proto-Mǐn \*hm-;
- ‘mosquito’ 蚊: Bái \*hmw : Proto-Mǐn \*hm-;
- ‘name’ 名: Bái \*hmiæN : Proto-Mǐn \*hm-;
- ‘fish’ 魚: Bái \*hɲiu : Proto-Mǐn \*ɲ-;
- ‘outside’ 外: Bái \*hɲuaC : Proto-Mǐn \*ɲ-;
- ‘silver’ 銀: Bái \*ɲin : Proto-Mǐn \*ɲ-;
- ‘moon’ 月: Bái \*hɲuaC : Proto-Mǐn \*ɲ-;
- ‘forehead’ 額: Bái \*hɲæC : Proto-Mǐn \*hɲ-;

The reflexes of Chinese 来母 (MC l-) is more interesting.

Etymon	Bái	Jiànyáng	Baxter-Sagart
流 ‘flow’	*gu	l-	*ru
柳 ‘willow’	*ɣwC	l-	*(mə.)ru?
漏 ‘leak’	*ɣwC	l-	*[Nə-r]ʰoks
落 ‘fall’	*ɣoC	l-	*kə.rʰak
镰 ‘sickle’	*jiN	l-	*[r]em
泪 ‘tears’	*jiC	l-	*[r][ə]ps
栗 ‘chestnut’	*jiC	l-	*[r]i[t]
力 ‘force’	*ɣwC	l-	*k.rək
利 ‘sharp’	*jiH	l-	*C.rits
聾 ‘deaf’	*kuN	s-	*C.rʰoŋ
露 ‘dew’	*koH	s-	*p.rʰaks
老 ‘old’	*kuX	s-	*C.rʰu?
兩 ‘two’	*koNX	s-	*p.raŋ?
六 ‘six’	*piuC??	s-	*k.ruk
犁 ‘plow’	*tɕi	l-	*[r]ʰ[i]j
虺 ‘snake’	*kʰyX	–	*[r]u[j]?
獺 ‘otter’	*tɕʰiaC	–	*rʰat

< \*\*[t/k]ʰæt

Proto-Bái	Old Chinese
*ɣ-, *g-	*r-
*k-	*C.r-
*kʰ-	*rʰ-

What it means for the phonetic reality of Late Old Chinese sonorants.

## 5 Baxter-Sagart uvulars in Bái

Typically, Bái uvulars correspond to Baxter-Sagart  $K^{\text{f}}$ . Exceptions:

- $*q\text{æ}NX$  ‘shade, light’ < 景 *yǐng* Baxter-Sagart  $qraŋ?$ , also *jǐng* C.  $qraŋ?$ .
- $*q\text{æ}N$  ‘today, this year’ < 今 Baxter-Sagart  $*[k]r[\text{ə}]m$ .
- $*q\text{æ}NX$  ‘copper’ < 金 Baxter-Sagart  $*k(r)[\text{ə}]m$ , cf. Càijiā *kǎ* ‘gold’.

For 今 and 金, two weeks ago in Shanghai, Yè Yùyīng reported the consensus among paleographers that the phonetic series contains both MC  $k$ - and  $ʔ$ - words.

## 6 Miscellanea

### 6.1 $*tʃ^h\text{æ}N^1 < *tʃ^h\text{e}\eta^1$ ‘light’ 輕

- $*tʃ^h\text{æ}N^1 < *tʃ^h\text{e}\eta^1$  ‘light’ < 輕  $k^hje\eta$ , OC  $*[k^h]e\eta$ .
- The Bái form corresponds to MC  $T\text{ç}$ - series, not rare  $Kj$ -.

### 6.2 $*p\text{i}N^1$ ‘wind’ 風: the dialectal position of Sino-Bái within Sinitic

< OC  $*pr\text{ə}m$

In Mainstream Chinese, cf. Baxter and Sagart (2014, p. 310 – 311):

The graphic connections of 風  $*pr\text{ə}m > pjuwng > f\text{ē}ng$  ‘wind (n.)’ suggest that it had  $*-r-$  before the vowel; but at the time  $*-r-$  disappeared, the vowel was already rounded and thus  $*-r-$  did not leave a trace on the final (as with original  $*-ruŋ$ ; see section 5.4.6.3).

In Bái, on the other hand, things happened in their right order, which should have produced a hypothetical Mandarin  $b\text{ī}n < MC p\text{i}m$ .

This is indeed an attested pronunciation in 3th-century Chinese. In *Shì míng* (~ 200AD), the pronunciation of 風 is described thus, the pronunciation anachronically noted in Middle Chinese:

風、兗豫司橫口合唇言之…青徐言風、踈口開唇推氣言之……

‘Wind 風 ( $p\text{i}m$ ): said in Yǎnzhōu, Yùzhōu and Sīzhōu with broadened mouth shape ( $i$ ) and closed lips ( $-m$ )...In Qīngzhōu and Xúzhōu (Modern Eastern Shāndōng), it is pronounced  $piuŋ$ , with contracted mouth ( $u$ ), open lips ( $-ŋ$ ) and forceful aspiration<sup>5</sup>.

<sup>5</sup>The meaning of this phrase is not clear. It may indicate an early labiodentalization and related with the fact that even in Mǐn dialects, the word is almost never noted with a  $p$ - initial.

In central China, including the Sīzhōu inspectory which contains the two capitals, the 3th century form resembles Bái and no known modern Chinese dialect.

The text *Shímíng* also documented another case of 3th-century dialect difference, discussed at length in Baxter and Sagart (2014, p. 310 – 311), one of the two places where Bái evidence is invoked in this work. For 天 ‘sky’, Baxter-Sagart  $l̥ʰin$ , central Chinese has *xen* while Shāndōng Chinese has *tʰen*. Again, present-day Chinese dialects agree with 3th-century Shāndōng dialect while Bái *\*xeN* agrees with the central dialect.

A final word to be discussed is 西 Baxter-Sagart  $*s.nʰər$ , which rhymed in *-n* during the early 3th century<sup>6</sup>. According to Baxter and Sagart (2014, p. 264 – 268), the reflex of final *\*-r* as *-n* is a Shāndōng dialect feature. If the nasalization in Jiànchuān *sein*<sup>l</sup> is genuine, it can be reconstructed to Proto-Bái *\*seN*. Bái nasalization is too unreliable, though, for these kind of assertions to be safely made.

Compared with mainstream modern Chinese, Bái consistently show features belonging to the Archaic Central dialect and di Qīng-Xú dialect features. It is the only witness of the form of Chinese once current in the greater part of China, including its two capitals.

## Bibliography

W.H. Baxter and L. Sagart. *Old Chinese: A New Reconstruction*. OUP USA, 2014.

Róng Lǐ et al. *Xiàndài Hànyǔ Fāngyán Dà Cídiǎn* 现代汉语方言大词典. Jiāngsū Jiàoyù Chūbǎnshè, 2002.

Fēng Wáng. *Comparison of Languages in contact: the distillation method and the case of Bai*. Language and linguistics monograph: Frontiers in linguistics. Institute of Linguistics, Academia Sinica, 2006.

Fēng Wáng. Bái yǔ yǔ báizú de liúbiàn: Duōjiǎodù jiéhé de shìyě 白語與白族的流變：多角度結合的視野. In Fēng Shí and Gāng Péng, editors, *Dàjiāng Dōngqù: Wáng Shìyuán Jiàoshòu Bāshísuì Hèshòu Wénjí* 大江東去：王士元教授八十歲賀壽文集. City University of Hong Kong Press, 2013.

Lín Xú and Yǎnsūn Zhào. *Báiyǔ Jiǎnzhì* 白語簡志. Mínzú Chūbǎnshè, 1984.

Míngjūn Yuán. *Hànbáiyǔ diàochá yánjiū* 汉白語調查研究. Zhōngguó Wénshǐ Chūbǎnshè, 2006.

Yǎnsūn Zhào and Lín Xú. *Bái-Hàn Cídiǎn* 白漢詞典. Sìchuān Mínzú Chūbǎnshè, 1996.

## Appendix: Comparaison of Proto-Bái rhymes

Zhōuchéng	Present	Wāng	Yuán	Notes
-----------	---------	------	------	-------

<sup>6</sup>For example,

耿耿伏枕不能眠	披衣出戶步東西	自謂終天路	忽然下沉淵
展詩清歌聊自寬	樂往哀來摧肺肝	驚飆接我出	故歸彼中田
	曹丕(187-226)·燕歌行	當南而更北	謂東而反西
		宕宕當何依	忽亡而復存
		曹植(192-232)·吁嗟篇	

a	(au)	a	rwa	o-ish reflexes in dialects
a	(auN)	aŋ	rwaN	o-ish reflexes in dialects
a	a	a	a	
a	a	–	ra	ie-ish reflexes in dialects
a	a	–	raN	ie-ish reflexes in dialects
a	aC	a	aS	
a	aN	aŋ	aN	
ia	iaC	Pr-æ, Pr-a	Pr-a	
ua	uaC	ua	at, w-aN	
ua	uaN	aŋ	ran, w-raN	ye-ish reflexes in dialects
ua	uaN	aŋ	waN	Q- only
e	(ei)	e	e	i-ish reflexes in dialects
e	e	ɛ	e	ɛ-ish reflexes in dialects
e	eC	e, ɛ	eS	
e	eN	en	eN	
ɛ	æ	æ	re	
ɛ	æC	æ	raS	
ɛ	æN	æŋ, æŋ	reN	
iɛ	iæC	Pr-æ	Pr-reS	
iɛ	iæN	Pr-æŋ	Pr-eN	
ue	uæ	uæ	w-re	
i	i	i	i	
i	iC	i	iS	
i	iN	en	iN	
i	iN	eŋ	iN	after hushing series
ui	ui	ui, uɛ	uii	
ui	uiC	uɛ	uit	
ui	uiN	uen	uun	
ou, o	o	ɔ	o	
ou, o	oC	o, ɔ	oS	
ou, o	oN	oŋ	oN	
u	u	u	u, ru	
uu	uu	uu	uu	
uu	uuC	uu	uuS	
uu	uuN	ũ	uuN	
v	iu	o	rwu	
v	iuC	o	ruS	
v	iuN	uŋ	ruN	labiodentals
v	uN	uŋ, oŋ	ruN	
v	y	o, ɛ	i	labiodentals and labiovelars
v	yC	–	iC	
v	yN	oŋ	i	labiodentals and labiovelars