

SOME PROBLEMS OF THE CONTEMPORARY CHINESE DICTIONARY (6<sup>th</sup> ed.):  
WITH A CALL FOR A CORPUS-BASED CHINESE LEXICOGRAPHY

Zequan Liu

*Yanshan University (Qinhuandao, China)*

Dandan Zhang

*Qiqihar University (Qiqihar, China)*

ABSTRACT

This paper sets out to examine the problems which exist with the newly published *Contemporary Chinese Dictionary* (6th edition), with particular reference to the establishment of the definitions, exemplifications and word class labels of its entries, the selection of its neologisms, as well as its reader-friendly considerations. The examination is conducted based on the query results from available Chinese corpora, with an aim to highlight the urgent necessity of the use of large-scale corpora for Chinese lexicography. The study concludes that it is now imperative to build open-ended, large-scale and multimodal corpora of contemporary Chinese in order to lend assistance to Chinese lexicography.

SUBJECT KEYWORDS

Corpus-based lexicography *The Contemporary Chinese Dictionary* (6th edition) Neologisms

School of Foreign Languages

Henan University

85 Minglun Street

Kaifeng, Henan Province 475001

CHINA

[zeqian@163.com]

中国河南省开封市明伦街 85 号

河南大学外国语学院 邮编 475001

School of Foreign Languages

Minzu University of China

27 Nanda St., Zhongguancun, Haidian District

Beijing, 100081, CHINA

[zhangdan1223@163.com]

北京市海淀区中关村南大街 27 号

中央民族大学外国语学院 邮编 100081

ON THE CLUSTER \*sr- IN SINO-TIBETAN<sup>#</sup>  
Guillaume Jacques  
CNRS/CRLAO/INALCO, Paris

ABSTRACT

This paper presents a critical overview of previously proposed etymologies involving the initial cluster \*sr- between Chinese and other Sino-Tibetan languages. It puts forth one new etymology, which confirms the simplification of the cluster \*sr- to s- in Kiranti and the preservation of this cluster in Rgyalrong languages.

SUBJECT KEYWORDS

Kiranti Rgyalrong Japhug Tibetan Clusters \*sr-

1. INTRODUCTION

The Middle Chinese 生 shēng initial consonant *s*, which originates from Old Chinese \*sr- in all modern systems of reconstruction,<sup>1</sup> is attested in a words of Sino-Tibetan origin, and corresponds to onsets either preserving a fricative +/r/ cluster or originating from one. Thus, \*sr appears to be one of the few consonant clusters uncontroversially reconstructible to proto-Sino-Tibetan.

In this paper, we first discuss previously proposed etymologies, and present the known correspondences of \*sr in languages other than Chinese. Second, we present a new etymology and discuss its significance for the conditioning of the sound laws in individual languages.

## 2. PREVIOUS COMPARISONS

Only three Chinese words with initial \*sr- correspond to forms that are widespread in the rest of the family and can be solidly reconstructed with initial \*sr-clusters. They were first proposed by Benedict (1972).

The first such etymon is the word 虱 \*srik → sri 'louse', which can be compared to Tibetan *sig*, Burmese *hrac*, Japhug *zruy* and Limbu *sif*. All words in this cognate set share the same meaning, and there is little doubt that they are related.

The second comparison is Chinese 色 \*srak → sik 'colour, sex, shame', which is compared to Tibetan *Me<sup>h</sup>ags*, *beags* 'confess', Burmese *hrak* 'shame', Japhug *tr-zrak* 'shame'. This comparison is however less convincing from the point of view of semantics, and in the case of Tibetan, philology suggests that the meaning 'confess' is secondary, and evolved from 'declare', the meaning attested in its oldest attestation, the bilingual Sino-Tibetan treaty inscription (example 1, translation after Li and Coblin 1987: 40, 80), where *beags* corresponds to Chinese 稽告 'make known, explain, declare'.

- (1) *ndi-lar bod rgi'a gnis ki'i rdze blon-g'is ka<sup>h</sup>taig*  
 this-like Tibet China two GEN sovereign minister-ERG together  
*beags mnañ bor-te*  
 PST:declare oath PST:throw-CONV  
 Thus the sovereigns and the ministers of both Tibet and China together declared and swore an oath. (Sino-Tibetan Treaty, West face, l. 71-72)

The third one is 生 \*N-srey → sey 'live, alive', corresponding to Burmese *hray<sup>2</sup>* 'alive' and other comparanda (see STEDT #71). This root has no cognate in Tibetan or Rgyalrong languages.

Other comparisons of Old Chinese \*sr- have been proposed by Coblin (1986) in particular, but they are restricted to Tibetan comparanda, and involve words with the onset sr- in Tibetan. Nearly all such comparisons can be shown to be invalid for various reasons.<sup>2</sup> The only promising such correspondence is 甥 \*sr<sup>h</sup>ey → sey 'sister's son' with Tibetan *sriy<sup>mo</sup>* 'sister' (the vowel correspondence is a consequence of Dempsey's law, see Hill 2014b).

The double correspondence \*sr- to sr- or ε- in Tibetan suggests that two proto-onsets must be reconstructed here: \*sa-r- with a reduced vowel yielding sr-, while the actual cluster \*sr- changes to ε-, perhaps through a stage \*[s].<sup>3</sup>

## 3. A NEW EXAMPLE OF PROTO-SINO-TIBETAN \*sr

All Rgyalrong languages share a common word for '(plant) root' attested by Japhug *tr-zrrm* 'root', Situ *-sram* and Zbu *-rzám*. It is an inalienably possessed noun with indefinite possessor prefix *tr-* (on which see Jacques 2014a: 4-5), and can be reconstructed as proto-Rgyalrong \**srrm*.

While some possessed nouns in Rgyalrong languages can derive from verb without any nominalization affix (see Jacques 2014a: 3-7), it is not the case for this noun, as no corresponding verb is found in any Rgyalrongic language.

Japhug has a variant *-srrm* which refers to the meaning 'root' in a more abstract sense of 'family lineage', as illustrated by the following example:

- (2) *nmuu gu-k<sup>h</sup>-ru mu~m<sup>h</sup>-pu<sup>h</sup>-tu<sup>h</sup>-cha gu ny,*  
 DEM TRANSLOC-INF-bling COND-NEG-IPFV-2-can FACT:be LINK  
*li n<sup>h</sup>-srrm n<sup>h</sup>-sro<sup>h</sup> ma me*  
 again 2SG.POSS-root 2SG.POSS-life apart:from FACT:not:exist  
 If you cannot bring it here, again, there is only your family and your life (for you to lose). (Slobdpon2, 2007)

This restricted meaning in a context involving a king and his subjects suggests that *-srrm* in Japhug is not inherited: it is borrowed from Situ Rgyalrong, which was the language of the local chieftain.<sup>4</sup> Note that borrowings from Tibetan, such as *-sro<sup>h</sup>* 'life' from *srog* have sr- in Japhug corresponding to Tibetan sr-, not zr- as in the inherited vocabulary.<sup>5</sup>

Apart from this example, voicing of s- in Japhug in this cluster and metathesis in Zbu is completely regular.

In Kiranti, we find a noun \**sam* attested by Khaling *sēm* 'root' (personal fieldwork), Yakha *sam* 'root', Kulung *sam* 'root' (Kongren 2007, Tolsma 2006). The correspondence of Kiranti initial \*s- to Japhug

sr- 'louse' is the same of that in the noun 'louse' (Japhug *zrru vs Kulung si*).

A search in STEDT reveals no similar form in any other Sino-Tibetan language. However, this word is phonologically comparable with Chinese 参 *shēn* (Middle Chinese *sim*). The character 参 has several readings, but Middle Chinese *sim* is associated with two meanings: one of the 28 constellations, and rhizomous medicinal plants such as Ginseng (still called in modern Chinese 人参 *rénshēn*). The earliest attestation of the use of 参 *sim* for a medicinal plant goes back to the Western Han dynasty, and some scholars have argued for an earlier date (for instance Xu 2011; Sun 1992).

Baxter and Sagart (2014: 75) reconstruct *\*srrm* for this character reading, but no evidence either from loanwords or phonetic series rule out the reconstruction *\*srrm*, which is the one adopted by other scholars (Schuessler 2009). Old Chinese *\*srrm* is a perfect match for proto-Rgyalrong *\*srrm* and proto-Kiranti *\*srrm* (as shown by Gong 1995 and Hill 2012, Old Chinese *\*a* regularly corresponds to *a* in Tibetan and other languages).

Chinese has innovated the noun 根 *[kʰa/n]* 'root', relegating the inherited word 参 *\*srrm* to rhizomous medicinal plants.

#### 4. LOSS OF \*-r-?

In addition to the correspondences seen in section 2, comparisons where Chinese *\*sr* corresponds to *s* in other languages have been proposed (in particular by Coblin 1986). Most of these examples either represent more complex correspondences (𪛗 *\*sr'am* → *sam* 'hair' corresponds to *s* in some languages, and to an affricate in others, as in Burmese *cham<sup>3</sup>* 'hair') or are spurious.<sup>6</sup>

Possible examples of the correspondence *\*sr* : *s* include the following:

- 殺 *\*sraɪ* → *set* 'kill' with Tibetan *gsod*, *bsad* 'kill', Japhug *sar* 'kill' etc. (on the vocalism of this word in Chinese, see Baxter and Sagart 2014: 214)

- 沙 *\*sr'aj* → *sae* 'sand' with Tibetan *sa* 'place' (see Hill 2014a concerning the rhyme correspondence).

- 漱 *\*sr'ok* → *sæwk* 'suck, drink' with Burmese *sok* 'drink'. If this comparison is valid, the original meaning probably was 'sip, suck', 'drink' being a parallel innovation in both languages.

The only attempt to explain the double correspondence of Chinese *\*sr-* to other languages is Handel (2002: 25). According to Handel, original PST *\*sr* changed to *s* in non-Chinese languages ('Tibeto-Burman') before nonfront vowels. This phonological solution has the merit of simplicity, and, if true, provides a common phonological innovation to all languages besides Chinese (the only one that has been explicitly proposed in print apart from the merger of *\*a* and *\*ə*, on which see Gong 1995, Handel 2008).

However, examples such as 色 *\*srək* 'colour, shame' or 参 *\*srrm* 'rhizome' refute Handel's theory, as they show that the conditioning factor that he proposed is not valid. There are three possibilities to account for the examples above.

First, it is possible that Handel is basically right, but that the conditioning is more restricted than he proposed: *\*sr-* is simplified to *\*s-* in languages other than Chinese only before *\*a* (and perhaps *\*o*), not before other non-front vowels such as *\*ə*. If confirmed, this would be another piece of evidence that the merger of *\*a* and *\*ə* is not a common innovation of non-Chinese languages (contra Gong 1995 and Handel 2008; see also Hill 2014a and Jacques 2014b: 75-6 for additional evidence of the preservation of the contrast in Lolo-Burmese and Tangut respectively). However, it would also constitute a potential common innovation for Sino-Tibetan languages other than Chinese.

Second, the *\*r-* could be secondary in Chinese. As proposed by Sagart (1999) (see also Baxter and Sagart 2014: 57-8), an infix *\*-r-* is reconstructible in Old Chinese, an alternative explanation is to consider Chinese here to be innovative in these three examples. In this alternative view, the three examples above represent infixed forms, while the original base forms without infix have been lost. Thus, there would no need to look for a phonological conditioning of this correspondence.

Third, an alternative possibility is that the present models of Old Chinese reconstruction (including Starostin 1989, Schuessler 2009, and Baxter and Sagart 2014) overestimate the quantity of syllables with medial or prefixed \**r*- in Old Chinese by overgeneralization. In all modern systems of reconstruction, \**r*- is reconstructed for all syllables with either second division rhyme, *chongniu* 3 and/or retroflex initials in Middle Chinese. While it has been convincingly demonstrated that clusters in \**r*- is indeed one possible origin for these syllables (Yakhontov 1961), there is no definite proof that \**r*- should be reconstructed in all cases.

As a measure of comparison, over 20% of syllables in Old Chinese as reconstructed by Baxter and Sagart (2014) contain a preinitial or a medial \**r*, while in Japhug and Tibetan, where consonant clusters including *r* are attested, we only find respectively 12% and 16% of syllables with non-initial *r*.

Given the limited number of reliable comparisons illustrating the correspondences at hand, it is too early to argue which of these three possibilities is the most probable, but each deserves to be investigated in detail.

## 5. CONCLUSION

The contribution of this paper is twofold. First, it provides a critical overview of previously proposed etymologies involving the onset \**sr*- in Old Chinese, and shows which etymologies are possible and which should be discarded, on the basis of philological and comparative data.

Second, it shows a new example of proto-Sino-Tibetan \**sr*-, and in particular the second comparison including Kiranti languages. It confirms that proto-Sino-Tibetan \**sr*- is simplified to \**s*- in proto-Kiranti. This work also contributes to the research on Sino-Tibetan subgrouping by exploring to what extent the correspondences at hand provide evidence for common innovations of non-Chinese Sino-Tibetan languages ('Tibeto-Burman').

## NOTES

<sup>#</sup> Old Chinese follows Baxter and Sagart 2014's system, Middle Chinese is in an API transcription based on Baxter (1992), and Tibetan is transcribed according to Jacques (2012). I wish to thank Gong Xun, Nathan W. Hill, Ma Kun, Lai Yunfan and Laurent Sagart for comments on previous versions of this paper. I am responsible for any remaining error.

1. Some reconstruction models, such as that of Baxter and Sagart (2014), allow more complex clusters such as \**snr*- where \**N* is a nasal with *g* as outcome in Middle Chinese. However, even in the case of these cluster an intermediate stage \**sNr*- → \**sr*- → *g*- has to be postulated.

2. The comparison of Chinese 率 \**s-rut* → *swit* 'rule' to Tibetanstrid 'government' is problematic for several reasons. The vowel correspondence is not a match (Gong 1995), and the Chinese verb is obviously related to 律 \**rut* → *witi* 'law, rule': the *s*- is here denominial. On the etymology of 律 \**rut*, see Sagart (2014).

The only other comparison, 產 \**s-ɲʰarʔ* → *sen* 'produce' to Tibetan *srel* 'bring up', which appears possible on the basis of Middle Chinese, is to be ruled out once Old Chinese reconstruction is taken into account.

3. Note that the causative *s*- forms of *r*- initial verbs in Tibetan is always *sr*-, never *ε*-.

4. Tusi 土司, in Japhug *ɽɽɽɽɽ* from Tibetan *ɽɽɽɽɽ* 'king'.

5. The only potential Tibetan borrowing with *zr*- is *zrritekw* 'bean' from Tibetan *srən(mə)* 'bean' (with the native diminutive suffix *-tew*), though it cannot be excluded that this word is a cognate between Japhug and Tibetan.

6. The comparison of Chinese 雙 \**sr'oj* → *gəwɲ* 'pair' to Tibetan *zuy* 'pair' proposed by Coblin is impossible as Tibetan *z* originates from pre-Tibetan \**dz*, see Hill (2014c).

## REFERENCES

- BAXTER, William H. III. 1992. *A Handbook of Old Chinese Phonology*. Trends in Linguistics, Studies and Monographs 64. Berlin: Mouton de Gruyter.

- BAXTER, William H., and Laurent Sagart. 2014. *Old Chinese: a New Reconstruction*. Oxford: Oxford University Press.
- BENEDICT, Paul K. 1972. *Sino-Tibetan: a Conspectus*. Cambridge: Cambridge University Press.
- COBLIN, Weldon South. 1986. *A Sinologist's Handlist of Sino-Tibetan Lexical Comparisons*. Netetal: Steyler Verlag.
- GONG, Hwang-chemg. 1995. The system of finals in Proto-Sino-Tibetan. In *The Ancestry of Chinese*, ed. William S.-Y. Wang, 41-92. Journal of Chinese Linguistics Monograph Series 8. Berkeley, CA: Project on Linguistic Analysis.
- HANDEL, Zev. 2002. Rethinking the medials of Old Chinese: Where are the r's? *Cahiers de Linguistique - Asie Orientale* 31(1):3-32.
- \_\_\_\_\_. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2:3:422-441.
- HILL, Nathan W. 2012. The six vowel hypothesis of Old Chinese in comparative context. *Bulletin of Chinese Linguistics* 6:1-69.
- \_\_\_\_\_. 2014a. Cognates of Old Chinese \*-n, \*-r, and \*-j in Tibetan and Burmese. *Cahiers de linguistique - Asie orientale* 43:91-109.
- \_\_\_\_\_. 2014b. Some Tibetan verb forms that violate Dempsey's law. *Revue d'Etudes Tibétaines* 29:91-101.
- \_\_\_\_\_. 2014c. Tibeto-Burman \*dz- > Tibetan z- and related proposals. In *Studies in Chinese and Sino-Tibetan Linguistics*, ed. Richard VanNess Simmons and Newell Ann Van Auker: 167-178. Taipei, Institute of Linguistics, Academia Sinica.
- JACQUES, Guillaume. 2012. A new transcription system for Old and Classical Tibetan. *Linguistics of the Tibeto-Burman Area* 35.2:89-96.
- \_\_\_\_\_. 2014a. Denominal affixes as sources of antipassive markers in Jarhug Rgyalrong. *Lingua* 138:1-22.
- \_\_\_\_\_. 2014b. *Esquisse de phonologie et de morphologie historique du tangout*. Leiden: Brill.
- KONGREN, Ranjee. 2007. *Yakha-Nepali-English Dictionary*. Kathmandu: Indigenous People Yakha Organisation.
- LI, Fang-kuei, and W South Coblin. 1987. *A Study of the Tibetan Inscriptions*. Nankang: Institute of History and Philology, Academia Sinica.
- SAGART, Laurent. 1999. *The roots of Old Chinese*. Amsterdam: Benjamins.

\_\_\_\_\_. 2014. A note on Tibeto-Burman bone words and Chinese pitch-  
pipes. In *Studies in Chinese and Sino-Tibetan Linguistics*, ed.  
Richard VanNess Simmons and Newell Ann Van Auker, 178-183.  
Taipei: Institute of Linguistics, Academia Sinica.

- SCHUESSLER, Axel. 2009. *Minimal Old Chinese and Later Han Chinese: A Comparison to Grammatica Serica Recensa*. Honolulu: University of Hawai'i Press.
- STAROSTIN, Sergej Anatol'evich. 1989. *Реконструкция древнетибетской фонологической системы*. Москва: Наука.
- SUN, Wencai 孙文采. 1992. Shi shep 释参 (Explaining the character "shep"). *Renshen yanjiu* 人参研究 2:38-41.
- TOLSMA, Gerard Jacobus. 2006. *A Grammar of Kiling*. Leiden: Brill.
- XU, Kunyu 徐坤宇. 2011. Shi "shep" 释“参” (Explaining the character "shep"). *Yuwen xuekan* 语文学刊 3:93-4.
- YAKHONTOV, Sergei Ye. 1961. Сочетания согласных в древнекитайском языке. In *Труды XXV Международного конгресса востоковедов*: 89-95.

### 关于汉藏语言中的 \*sr- 群

向柏霖

法国科学研究中心-法国国立东方语言文化学院

#### 提要

本文批判性的总结了以前关于汉语与汉藏语言之间关于各种 \*sr- 辅音  
声母群语源的提议。文章提出 \*sr- 新的语源；证实 \*sr- 到 s- 在基兰  
提语言里的简化及 \*sr- 在嘉绒语言中的保存。

#### 主题词

基兰提语 茶堡语 嘉绒语 藏语 \*sr- 群

INALCO/CRLAO

2 rue de lille  
75007 Paris

France

[ryvalongskad@gmail.com]