



**HAL**  
open science

## Review of Wang Feng (2006). Comparison of languages in contact: The distillation method and the case of Bai.

Ekaterina Chirkova

### ► To cite this version:

Ekaterina Chirkova. Review of Wang Feng (2006). Comparison of languages in contact: The distillation method and the case of Bai.. 2007. hal-00187281

**HAL Id: hal-00187281**

**<https://hal.science/hal-00187281>**

Submitted on 14 Nov 2007

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Wang Feng (2006). *Comparison of languages in contact: The distillation method and the case of Bai*. Language and linguistics monograph series B, Frontiers in linguistics III. Taipei: Institute of Linguistics, Academia Sinica. 227 pp. ISBN 986-00-5228-X.

In the past decades, Sino-Tibetan historical linguistics has increasingly become a flourishing field of studies, both in China and elsewhere. Recent trends include growing awareness of the importance of shared common innovations for subgrouping within Sino-Tibetan (Sagart 1995; van Driem 2001, 2006), interest in computational methods derived from evolutionary biology to test hypotheses about the sequence and timing of proposed subgroups (Sagart, Blench and Sanchez-Mazas 2005), and steady advances in Old Chinese reconstructions (Baxter 1992; Baxter and Sagart 1997; Sagart 1999; Zhengzhang 2003). Wang Feng's *Comparison of languages in contact: The distillation method and the case of Bai* (the revised version of his 2004 Ph.D. dissertation) marries all these developments, while focusing on a language of considerable importance for a better understanding of Sino-Tibetan: Bai. The genetic affiliation of the Bai language has been disputed due to its prolonged contact history with Chinese. However, despite an overwhelming number of Chinese loanwords, most scholars see Bai as an independent Sino-Tibetan language (cf. Lacouperie 1887; Li 1937; Zhao 1982; Dell 1981; Lee and Sagart 1998; Matisoff 2003; van Driem 2001). In *Comparison of languages in contact: The distillation method and the case of Bai*, Wang postulates a set of Sino-Bai related words and on the basis thereof argues that the relationship between Bai and Chinese is due to inheritance from a common ancestor. As is evident from the title, Wang considers the major contribution of his book to be methodological. That is, he advances a general method, the Distillation Method, for establishing the genetic affiliation of a language whose origins are unknown due to long and intensive contact with one or more other languages. This review accordingly focuses on the methodological aspects of Wang's study.

The book is organized into five chapters. Chapter 1 outlines the Distillation Method, a new procedure in historical linguistics, designed to “exclude elements from horizontal transmission [i.e. borrowing], consequently obtaining reliable genetic evidence

from vertical transmission [i.e. common inheritance] on the basis of language comparisons” (p. 1). The Distillation Method comprises three steps: (i) Intra-comparison; (ii) Inter-comparison; and (iii) Recognition.

Chapter 2 presents the first step. Intra-comparison is defined by Wang as “detection of corresponding lexical items across dialects of a language with which the proto-form of that language may be constructed” (p. 1). In connection to Bai, he uses data from nine Bai dialects to reconstruct their linguistic ancestor. In addition, he performs a subgrouping of these dialects using two phylogenetic algorithms, PAUP and PENNY.

Chapter 3 discusses the second step. Inter-comparison aims at detection of a set of corresponding lexical items across proto-languages with which the core of that language may be identified (cf. pp. 1-2). This procedure, according to Wang, allows him to stratify Sino-Bai related words in order to exclude layers of Chinese borrowings in Proto-Bai.

Chapter 4 deals with the third step, Recognition. In this chapter, Wang attempts to detect cognates and borrowings in the oldest layer of Sino-Bai and to assess their ratio based on the Inexplicability Principle and Chen Baoya’s (1996) Rank Theory. The Inexplicability Principle refers to “the inability to describe a recipient language in terms of the phonological system of the donor language” (p. 2). The “inexplicable elements” are then considered to be inherited from the ancestor language. Chen Baoya’s Rank Theory consists in distinguishing between a 100-word list of more basic and a 100-word list of less basic words. Chen then argues that if two languages are genetically related, the proportion of cognates in the 100-word list of more basic words will be greater than the proportion thereof in the 100-word list of less basic words. This is apparently an independent duplication of Sergei E. Yakhontov’s analysis of the Swadesh 100-word list as consisting of more basic (35) and less basic (65) words and of his similar idea of the relationship between the ratio of related and unrelated words within the more and less basic words respectively and the genetic relatedness of the compared languages (Starostin 1991: 59-60).

Chapter 5 summarizes the major findings of the study. In addition, Appendix 1 offers a list of 2.124 lexical items in nine Bai dialects, accompanied by corresponding reconstructions of proto-forms; Appendix 2 lists Sino-Bai related words (2.109 items);

and Appendix 3 presents layers of Sino-Bai related words within the Swadesh 200 word list.

This is an innovative work, based on first-hand data of no less than nine Bai dialects. The author creatively combines Chinese and Western scholarship, dares to question established methods and proposes unorthodox solutions. There are, however, fundamental problems with the approach, of which these five are most significant: (1) unclarity of background information, key terms and underlying assumptions; (2) proposed principle of loanword stratification; (3) chronology of Proto-Bai and its stages; (4) criteria for dialect subgrouping; (5) use of Old Chinese reconstructions.

For someone who is neither versed in Bai, nor familiar with phylogenetic algorithms, the book is not reader-friendly. The author does not provide information on the history of Bai or on the distribution and history of neighboring languages and their relation to Bai. To give one example, Wang bases his work on nine Bai dialects, but does not specify the total number of Bai dialects or explain, if more than nine, why these particular dialects have been chosen. This information is highly pertinent to some of the arguments pursued in the book, as will be shown further on.

Equally unclear are also some terms and assumptions underlying the study. Let us examine one crucial issue. Given that this book aims to distinguish Chinese borrowings in Bai from Sino-Bai cognates, the notion of “borrowing” is one of its fundamental concepts. Wang uses this term in different contexts and introduces fine but not always clearly defined gradations. The gradation of “later borrowing” is the most frequent and the most essential to the author’s line of reasoning. At first, it seems that “later borrowing” refers to loanwords that entered Bai dialects *after* their split from the proto-language (p. 9). Therefore, later borrowing should be any kind of borrowing that postdates the split of the proto-language. Further on, however, the term “later borrowing” apparently refers to borrowings within the reconstructed Proto-Bai lexicon, i.e. *before* the split of the proto-language (Chapter 3). This unclarity of terms and assumptions is one of the major flaws of this book, as it impedes full appreciation of the posed aims, suggested methods and underlying logics of this study.

To deal with numerous Chinese loanwords in Bai, *Comparison of languages in contact: The distillation method and the case of Bai* advances a new principle of

loanword stratification. As a starting point, Wang examines Laurent Sagart and Xu Shixuan's (2001) method of loanword classification and concludes that their Principle of Coherence (i.e. the initial, final and tone correspondences on a borrowed syllabic morpheme obey the same set of correspondences) and their Extended Principle of Coherence (i.e. the initial, final and tone correspondences on all syllables of one borrowed polysyllabic morpheme obey the same set of correspondences, provided the morpheme is semantically noncompositional) are unsuitable for his purposes. Notably, Sagart and Xu's method has been successfully tested on several languages (Bai-Chinese [Lee and Sagart 1998], Hani-Chinese [Sagart and Xu 2001], rGyal-rong-Tibetan [Jacques 2004]). Rejecting their method therefore requires solid reasons and substantial alternatives. Wang offers neither.

As far as we can judge from the book, Wang's only motivation for questioning the Principle of Coherence and the Extended Principle of Coherence is one unique example: the two variants for 'lute',  $\text{pi}^{21}\text{pho}^{55}$  and  $\text{phi}^{21}\text{pe}^{21}$ , in Jianchuan Bai (p. 20).<sup>1</sup> He takes the Middle Chinese form  $\text{bjij}1^{\text{b}}\text{b}\text{æ}1^{\text{b}}$  to be the donor form for these two words. To him, they exemplify the effect of the so-called bidirectional diffusion, whereby the initial, final and tone of the syllable do not change at the same rate.<sup>2</sup> This example supposedly violates Sagart and Xu's Extended Principle of Coherence, as the initials and tones of the two syllables of the borrowed polysyllabic words for 'lute' do not obey the same set of sound correspondences. Thus, the two words for 'lute' in Jianchuan Bai are argued to exhibit interaction between two different strata of borrowings from Chinese, i.e. between the hypothetical layers  $*\text{phi}^{55}\text{pho}^{55}$  and  $*\text{pi}^{21}\text{pe}^{21}$ .

In a footnote (p. 20 n. 6), Wang mentions that he discussed this particular example with Sagart, who suggested that the two variants in Jianchuan Bai are not borrowings from Chinese, but loans from another surrounding language. Nonetheless,

---

<sup>1</sup> Overall, it is unclear why a Jianchuan Bai example is used, since Jianchuan Bai is not on the list of nine Bai dialects surveyed by Wang. Equally unclear is what the word for 'lute' is in the surveyed dialects (this word is not included into the lexical list in Appendix 1).

<sup>2</sup> The term bidirectional diffusion is not defined in the text. It refers to a case of competing tone changes in three Chinese dialects discussed by William S.-Y. Wang and Chinfa Lien (1993). It is argued that in this case, two competing changes (internal and contact-induced) spread out from different foci within a single speech community by lexical diffusion.

Wang prefers the explanation of bidirectional diffusion, for he did not find a possible donor language.

It is impossible for the reader to evaluate the validity of Wang's preferred explanation, since no systematic historical information on donor languages for Bai is offered in the book. However, if we take a brief look at Bai history, we find that the Jianchuan area was under Tibetan political control during the early eighth century (Wiersma 2003: 652). This would suggest Tibetan as the donor language for the word 'lute'. In Tibetan, the word for 'lute', *pi-wang*, is also a loan, probably from a Khotanese source (Beyer 1992: 140, n. 29). According to Beyer, *pi-wang* is an early transfer, already attested in ninth-century Tibetan texts. This word suggests itself as a regular donor (in terms of sound correspondences) for the variants of the word 'lute',  $\text{pi}^{21}\text{ph}\alpha^{55}$  and  $\text{phi}^{21}\text{p}\epsilon^{21}$ , in Jianchuan Bai. (The two variants probably reflect two different strata of borrowings or two distinct dialectal renderings of the Tibetan original). Note that the difference between the two bilabials in the donor word (*p/w*) is reflected as that between the voiceless unaspirated and voiceless aspirated bilabials in Bai (*p/ph* in  $\text{pi}^{21}\text{ph}\alpha^{55}$  or *ph/p* in  $\text{phi}^{21}\text{p}\epsilon^{21}$ ) and the Extended Principle of Coherence is therefore neatly observed.

Since Wang claims that in loanwords, due to bidirectional diffusion, sub-components of a syllable (initial, final and tone) change at different rates, he proposes to use initial, final and tone separately as basic units for loanword stratification. This is his alternative to Sagart and Xu's Principle of Coherence. Wang also suggests that different layers of borrowings in the recipient language should be explainable in terms of the phonological development of the donor language (pp. 25-26). The only example illustrating this point is the Old and Middle Chinese initial *\*d-*, which corresponds to two initials, *\*d-* and *\*th-*, in Wang's reconstructed Proto-Bai lexicon. Wang hypothesizes that the latter represent two different layers in Proto-Bai, one older (*\*d-*) and one later (*\*th-*), the order of which can be explained by the process of devoicing in Chinese. The same principle is also directly applied in detection of the oldest layer of Bai. (Overall, in a situation of stratified borrowing, the oldest layer is understood to be maximum free of borrowings and is thus diagnostic for the genetic affiliation of the language in question.) Notably, Wang's principle of loanword stratification is obviously unable to isolate the

oldest layer in those instances, where a particular subcomponent of a syllable remains unchanged throughout different stages of Chinese, as is the case in the example above. And indeed, Wang admits being unable to determine whether the Proto-Bai word *\*dɔl* ‘peach’ with the initial *\*d-* (corresponding to the Old *and* Middle Chinese *d-*), is borrowed from Chinese or came from a single shared ancestor of Proto-Bai and Chinese (p. 26).

A related serious problem is that Wang never specifies the time frame of his Proto-Bai. Neither does he attempt to chronologically order “later borrowings”, despite his explicit promise to do so on p. 26. What one finds instead in Chapter 3, which aims at isolating the oldest layer, is a layer established on the basis of Old Chinese reconstructions, and which Wang then considers to be *the* oldest layer, versus all the rest, labeled “later borrowing”. This absence of a clear chronological parameter in Wang’s study, in my opinion, does not only impede his stratification, but also affects his phylogenetic subgrouping argument.

Wang explicitly aims at a phylogenetic subgrouping of Bai dialect data on the basis of unique shared innovations (p. 12). Overall, phylogenetic subgrouping essentially relies on relative chronology to determine what changes are diagnostic for subgrouping and what changes are the results of convergence or independent development (cf. Bowerman and Koch 2004: 2-7). Presumably, as a consequence of his chronological underspecification of Proto-Bai and its stages, Wang never explicitly addresses the issue of relative chronology of changes between Proto-Bai and modern Bai dialects. Paradoxically, Wang sees relative chronology of these changes not as an underlying criterion, but as a by-product of a subgrouping, i.e. something which is inferred on the basis of an already performed subgrouping (p. 121). It is conceivable that the algorithms he uses impose limitations on his choice of features underlying phylogenetic calculations, but the reader is not informed. Overall, the criteria for selecting 19 features out of the total of 35 phonological changes and several morphosyntactic and lexical changes between the reconstructed proto-form of the language and modern dialects are unspecified. (The statement to the effect that “the selection of subgrouping criteria is better to be treated probabilistically” (p. 12) has little explanatory value.)

Note that throughout the discussion of dialect data in Chapter 2, Enqi Bai appears to stand out from the rest. For example, it has the most complete uvular series of all Bai dialects and it is on the basis of this dialect \**n* is reconstructed for Proto-Bai (p. 71). Only this dialect merged Proto-Bai \*1b and \*4b tones (p. 57) and it appears to have preserved many conservative lexical forms already replaced in other dialects, such as the form for ‘mountain’ (p. 107). In a similar fashion, Mazhelong Bai has some unique characteristics too and at some point, while discussing the word for ‘breast’ (p. 110-111), the author even speculates that Mazhelong Bai was the earliest dialect to have split from Proto-Bai. However, since none of these seemingly unique characteristics of Enqi or Mazhelong Bai listed above have been taken into consideration in the process of feature selection for the algorithms, the potentially special position of these dialects does not show in the resulting family trees. In sum, the unspecified and at times problematic nature of feature selection casts doubt on the results of the subgrouping.

A large part of this book is devoted to comparisons between reconstructed Proto-Bai forms and corresponding forms in Old and Middle Chinese. In the concluding chapter, Wang states that the comparison between Bai and Chinese carried out in this study sheds new light on Old Chinese phonology as well as on Old Chinese morphosyntax (p. 176). Generally speaking, this type of Sino-Bai comparisons is bound to enrich our understanding of Old Chinese phonology and morphosyntax, but, in my opinion, more caution should have been exercised when using Old Chinese reconstructions. For his comparisons with Proto-Bai, Wang uses different and at points mutually incompatible systems of Old Chinese (Li Fang-kuei 1971; Gong Hwang-cherng 1989; Baxter 1992; Sagart 1999). In addition, he uses discussions of separate Middle and Old Chinese phenomena in Shao (1991, 1995) and Zeng (2003). He evaluates these competing proposals against his Proto-Bai data. While presumably trying to be exhaustive, the author takes the freedom to choose from different systems whatever matches best his own reconstructions. Regrettably, Wang never questions how his suggestions affect the overall integrity of a particular system of Old Chinese or how well it fits internal or external evidence which brought a certain reconstruction about.

To conclude, this book provides new data on heretofore little-studied varieties of Bai and thus advances our knowledge of Bai. It creatively applies recent methods in



historical linguistics. It combines Western and Chinese scholarly traditions and promotes theories by Chinese linguists (such as Chen Baoya), which are still of little renown beyond China's borders. But the author pursues too many goals at the same time. If he had been working with just one dialect of Bai and had consistently applied Sagart and Xu's method of loanword classification, he would have probably arrived at a more consistent and chronologically sound oldest genetic layer. And if he had compared the resulting layer with corresponding Old Chinese forms according to just one system of Old Chinese reconstructions, he may have had stronger arguments for postulating sound correspondences between the two. By and large, Wang tackles complex issues in language comparison. The biggest contribution of his study, to my mind, is that he brings these issues to discussion, and thereby stimulates quests in the genetic affiliation of lesser known Sino-Tibetan languages.

## References

- Baxter, William H. 1992. *A handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.
- Baxter, William and Laurent Sagart. 1997. "Word formation in Old Chinese". Packard, Jerome, ed. *New approaches to Chinese word formation*. Berlin: Mouton de Gruyter, pp. 35–76.
- Beyer, Stephan V. 1992. *The classical Tibetan language*. Albany: State University of New York Press.
- Bowern, Claire and Harold Koch. 2004. "Introduction: Subgrouping methodology in historical linguistics". Bowern, Claire and Harold Koch, eds. *Australian languages: Classification and the comparative method*. Amsterdam-Philadelphia: John Benjamins Publishing Company, pp. 1-15.
- Chen Baoya 陈保亚. 1996. *Lun yuyan jiechu yu yuyan lianmeng* 《论语言接触与语言联盟》 [Language contact and language union]. Beijing 北京: Yuwen Chubanshe 语文出版社.
- Dell, François. 1981. *La langue bai, phonologie et lexique*. Paris: Editions de l'Ecole des Hautes Etudes en Sciences Sociales.

- Gong Hwang-cherng 龚煌城. 1989 [2002]. “Cong Han-Zangyu de bijiao kan Hanyu shanggu ruogan shengmu de nice” “从汉藏语的比较看汉语上古若干声母的拟策” [The reconstruction of some initials in Old Chinese from the viewpoint of Sino-Tibetan comparison]. *Collected papers on Sino-Tibetan linguistics*. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Jacques, Guillaume. 2004. *Phonologie et morphologie du japhug (rGyalrong)*, Ph.D. dissertation: Université Paris VII - Denis Diderot.
- Lacouperie, de, Albert Terrien. 1887. *The languages of China before the Chinese*. London: David Nutt, 270, Strand. Reprinted by Taipei: Ch'engwen Publishing Company.
- Lee Yeon-Ju and Laurent Sagart. 1998. “The strata in Bai”, a paper presented to the 31st International Conference on Sino-Tibetan languages and Linguistics, Lund (handout).
- Li Fang-kuei 李方桂. 1937. “Chinese languages and dialects”. *The Chinese yearbook* 1937. Shanghai: Commercial Press, pp. 59-65. (Reprinted in *Journal of Chinese Linguistics* (1973) 1.1:1-13.) [East Asia Serials]
- . 1971 [1980]. *Shanggu yin yanjiu* 《上古音研究》 [Studies on Old Chinese phonology]. Beijing 北京: Shangwu Yinshuguan 商务印书馆.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman*. Berkeley, Los Angeles, London: University of California Press.
- Sagart, Laurent. 1995. “Comments from Sagart”. In: Wang, William S.-Y., ed. *The ancestry of the Chinese language*. *Journal of Chinese Linguistics* monograph series no. 8. Berkeley, CA: Project on Linguistic Analysis, University of California, pp. 337-372.
- . 1999. *The roots of Old Chinese*. Amsterdam, Philadelphia: John Benjamins.
- Sagart, Laurent, Blench, Roger and Alicia Sanchez-Mazas, eds. 2005. *The peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*. London: Routledge Curzon.

- Sagart, Laurent and Xu Shixuan. 2001. "History through loanwords: The loan correspondences between Hani and Chinese". *Cahiers de linguistique - Asie orientale* 30.1: 3-54.
- Shao Rongfen 邵荣芬. 1995. "Xiamuzi shanggu yifen wei'er zaizheng" "匣母字上古一分为二再证" [More evidence for the two Old Chinese origins of the Middle Chinese initial *h*-]. *Zhongguo yuyan xuebao* 《中国语言学报》 7: 121-134.
- Starostin, S. 1991. *Altajskaja problema i proisxoždenie japonskogo jazyka* [*The Altaic problem and the origin of the Japanese language*]. Moscow: Nauka.
- van Driem, George. 2001. *Languages of the Himalayas: An ethnolinguistic handbook*. Leiden: Brill.
- . 2006. "The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese". A paper presented at the joint meeting of the 14th Annual Conference of the International Association of Chinese Linguistics and the 10th International Symposium on Chinese Languages and Linguistics at the Academia Sinica in Taipei, May 2006.
- Wang, William S.-Y. and Chinfu Lien. 1993. "Bidirectional diffusion in sound change". Jones, Charles, ed. *Historical linguistics: Problems and prospectives*. London: Longman, pp. 345-400.
- Wiersma, Grace. 2003. "Yunnan Bai". Thurgood, Graham and Randy J. LaPolla. *The Sino-Tibetan languages*. London and New York: Routledge, pp. 651-673.
- Zeng Xiaoyu 曾晓渝. 2003. "Jianmu de shanggu yinzhì" "见母的上古音质" [The phonetic value of the initial *k*- in Old Chinese]. *Zhongguo Yuwen* 《中国语文》 2: 109-120.
- Zhao Yansun 赵衍荪. 1982. "Baiyu de xishu wenti" "白语的系属问题". *Minzu yuwen yanjiu wenji* 《民族语文研究文集》. Xining 西宁: Qinghai Minzu Chubanshe 青海民族出版社, pp. 150-188.
- Zhengzhang Shangfang 郑张尚芳. 2003. *Shanggu yinxi* 《上古音系》 [Old Chinese phonology]. Shanghai 上海: Shanghai Jiaoyu Chubanshe 上海教育出版社.

Research Fellow  
Centre de recherches linguistiques sur l'Asie orientale  
CNRS  
chirkova.k@gmail.com