Российский государственный гуманитарный университет

Russian State University for the Humanities

Journal of Language Relationship

International Scientific Periodical

№ 17/1

Вопросы языкового родства

Международный научный журнал

№ 17/1

ВЯР

Table of Contents / Содержание

Articles / Статьи

Contributors

Christopher J. FOSTER — Stanley Ho Junior research fellow, Pembroke College, University of Oxford, Oxford, christopher.foster@pmb.ox.ac.uk

Xun GONG — postdoctoral researcher, Department of East Asian Languages and Cultures, School of Oriental and African Studies, University of London, minus273cn@gmail.com

HUO Wenwen — graduate student, School of Humanities, Wenzhou University, Zhejiang, hww13736970100@163.com

Nathan W. HILL — Head of the Department of East Asian Languages and Cultures, School of Oriental and African Studies, University of London, nh36@soas.ac.uk

Guillaume JACQUES — research director, Centre National de la Recherche Scientifique, Paris, rgyalrongskad@gmail.com

JIN Lixin — professor, School of Humanities, Wenzhou University, Zhejiang, ytjlx@hotmail.com

LAI Yunfan — postdoctoral researcher, Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, lai@shh.mpg.de

Johann-Mattis LIST — senior researcher, Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, mattis.list@lingpy.org

MA Kun — postdoctoral researcher, Department of Chinese, Sun Yat-sen University, Guangzhou, 9filippo@gmail.com

SHEN Ruiqing — assistant professor, Department of Chinese Studies, National University of Singapore, shenruiqings@hotmail.com

George STAROSTIN — senior researcher, Institute of Oriental and Classical Studies, Higher School of Economics / Russian State University for the Humanities, gstarst1@gmail.com

ZHANG Shuya — doctoral student, Institut National des Langues et Civilisations Orientales, Paris, zhang.shuyafr@gmail.com

Сведения об авторах

ГУН, Сюнь — постдокторант отдела языков и культур Восточной Азии, Школа восточных и африканских исследований, Лондонский университет, minus273cn@gmail.com

ЖАК, Гийом — заведующий отделом исследований, Национальный центр научных исследований, Париж, rgyalrongskad@gmail.com

ЛАЙ, Юньфань — постдокторант отдела лингвистической и культурной эволюции, Институт Макса Планка, Йена, lai@shh.mpg.de

ЛИСТ, Йоханн-Маттис — старший научный сотрудник отдела лингвистической и культурной эволюции, Институт Макса Планка, Йена, mattis.list@lingpy.org

МА, Кунь — постдокторант отдела Китая, Университет Сунь Ятсена, Гуанчжоу, 9filippo@gmail.com

СТАРОСТИН, Г. С. — старший научный сотрудник Института классического Востока и античности ВШЭ, заведующий кафедрой истории и филологии Дальнего Востока ИВКА РГГУ, gstarst1@gmail.com

ФОСТЕР, Кристофер Дж. — научный сотрудник Пемброкского колледжа, Оксфордский университет, christopher.foster@pmb.ox.ac.uk

ХИЛЛ, Нэйтан У. — заведующий отделом языков и культур Восточной Азии, Школа восточных и африканских исследований, Лондонский университет, nh36@soas.ac.uk

ХО, Вэньвэнь — студент Школы гуманитарных наук, Вэньчжоуский университет, Чжэцзян, hww13736970100@163.com

ЦЗИНЬ, Лисинь — профессор Школы гуманитарных наук, Вэньчжоуский университет, Чжэцзян, ytjlx@hotmail.com

ЧЖАН, Шуя — докторант, Национальный институт восточных языков и цивилизаций, Париж, zhang.shuyafr@gmail.com

ШЭНЬ, Жуйцин — доцент отдела китаистики, Национальный университет Сингапура, shenruiqings@hotmail.com

Note for Contributors

Journal of Language Relationship welcomes submissions from everyone specializing in comparative-historical linguistics and related disciplines, in the form of original articles as well as reviews of recent publications. All such submissions should be sent to the managing editor:

G. Starostin
Institute for Oriental and Classical Studies
Russian State University for the Humanities
125267 Moscow, Russia
Miusskaya Square, 6
E-mail: gstarst@rinet.ru

Articles are published preferably in English or Russian, although publication of texts in other major European languages (French, German, etc.) is possible. Each article should be accompanied with an abstract (not exceeding 300 words) and keywords.

For more detailed guidelines on article submission and editorial policies, please see our website at: http://www.jolr.ru or address the editorial staff directly at gstarst@rinet.ru.

Будущим авторам

Журнал *Вопросы языкового родства* принимает заявки на публикацию оригинальных научных статей, а также рецензий от всех, кто специализируется в области сравнительно-исторического языкознания и смежных дисциплин. Рукописи можно высылать непосредственно заместителю главного редактора по адресу:

125267 Москва
Миусская площадь, д. 6
Российский государственный гуманитарный университет
Институт восточных культур и античности
Г. Старостину
E-mail: gstarst@rinet.ru

Предпочтительные языки публикации — английский или русский, хотя возможна также публикация статей на других европейских языках (французский, немецкий и т. п.). К каждой статье обязательно прикладывается резюме (не более 300 слов) и список ключевых слов.

Подробнее о требованиях к оформлению рукописи, редакционной политике журнала и т. п. вы можете узнать на нашем сайте по адресу: http://www.jolr.ru или же непосредственно, обратившись к редакции по электронной почте (gstarst@rinet.ru).

Preface

"Old Chinese and Friends": new approaches
to historical linguistics of the Sino-Tibetan area

> [La] restitution d'une « langue commune » dont le chinois, le tibétain, etc., par exemple, seraient des formes postérieures, se heurte à des obstacles quasi invincibles.
>
> (Antoine Meillet 1866–1936, [1925] 1954, 26f)

Despite almost 200 years of research, we still know very little about the Sino-Tibetan proto-language and its descendants, be it its age, its geographic origin or its history of spread. In this regard, our knowledge contrasts sharply with our insights in such well-established language families of similar size as Indo-European, where the major subgroups are well-established and scholarly discussions concentrate on the age of the language family which they link to archaeological records (Anthony and Ringe 2015), or Austronesian, where most scholars even seem to have reached a consensus with respect to the proto-language's age and *Urheimat* (Blust 2019). The poor state of Sino-Tibetan historical linguistics is perhaps best reflected in the dispute about the names of the language families. Some scholars deliberately exclude or marginalize the importance of the Sinitic subgroup, focusing on what they consider the *core* of the language family, *Tibeto-Burman* (Matisoff 2015). Other scholars emphasize the limits of our current knowledge with respect to the detailed affiliation of the numerous subgroups, demanding an unbiased bottom-up approach to what they call the *Trans-Himalayan* language family (Driem 2014).

Neither of the two approaches seems satisfying. Excluding Chinese as a subgroup from the beginning and concentrating on some kind of a core language family before any agreement has been reached of whether this Tibeto-Burman core ever existed is obviously imposing an unwanted bias on our research. Ignoring all insights and discussions that have been made into higher subgroups, however, seems also overly pessimistic. Furthermore, renaming a language family that is well-established enough as a whole (Jacques 2015) has the clear disadvantage of broadening the frontiers between the different camps in our field, depriving those scholars who want to remain independent of the possibility to do so by forcing them to choose one of the two camps. For the Trans-Himalayan camp, the use of the term "Sino-Tibetan" is seen as a subscription to the Tibeto-Burman paradigm, while the term "Trans-Himalayan" is seen as a clear opposition to the work of the Tibeto-Burman camp. For those scholars who refuse to camp, the best option seems to stick to the older term Sino-Tibetan, as it is done here.

Investigating the history of Sino-Tibetan languages is particularly hard for three reasons: (1) Language contact is widespread. (2) Sporadic processes of morphological and analogical change mask regular sound change processes. (3) The Sino-Tibetan languages exhibit a high degree of typological diversity.

In the Sino-Tibetan language family, language contact is the rule rather than the exception (Thurgood 2003), including contact inside subgroups, among subgroups, or with neighboring language families, like Tai-Kadai, Hmong-Mien, or Austro-Asiatic (see Figure 1). Due to intensive contact within Sinitic, for example, most Chinese dialectologists agree with Norman (2003, 76f) that Chinese is "not entirely amenable to a Stammbaum formulation". Due to more than a thousand years of intensive contact between Bai and Chinese languages, Sino-Tibetan
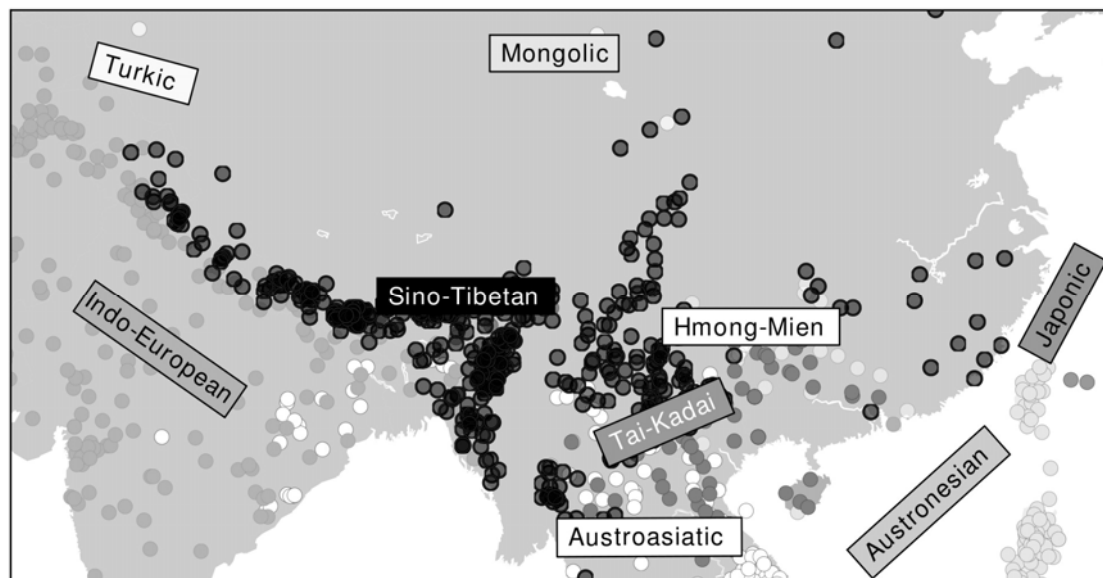
Sino-Tibetan and neighboring language families

linguists disagree whether the Bai languages are the closest relative of Chinese (Starostin [1995] 2007; Wang 2006) or a "normal" but heavily siniticized subgroup of Tibeto-Burman (Lee and Sagart 2008; Matisoff 2003).

If language contact can be excluded, sound change is a predominantly regular process that spreads across the whole lexicon of a language (Blevins 2004: 260–68; Kiparsky 1988; Labov 1981). Morphological processes, like suffixation, compounding, or analogy, however, are predominantly sporadic. Morphological processes can *mask* the regularity of sound change processes and obstruct the identification of regular sound correspondences. Compounding, for example, is a major process of word formation in the Sino-Tibetan family (Matisoff 2003: 153f). If compounds are reduced due to contraction (List 2017; Trask 2000: 92), they obscur regular sound correspondences, and this may explain the large-scale inconsistencies in sound correspondences among Sino-Tibetan languages (Handel 2008: 425f). When carrying out a lexical comparison based on word lists, compounding exacerbates the difficulties of identifying cognates, since words across different languages may share only one morpheme which may yield complex patterns of partial cognacy (List 2015: 56–58; Matisoff 2000: 341f; Satterthwaite-Phillips 2011: 99f).

Sino-Tibetan languages are typologically quite diverse. Tonogenesis (Abramson 2004; Haudricourt 1954), the process by which languages develop tone, occurred frequently and independently in the history of the Sino-Tibetan languages, and sometimes, as in the case of Tibetan, even subgroups have dialects with tone and dialects lacking tone. There are languages with rich inflectional morphology, like the Kiranti languages (Ebert 2003), and languages that are completely isolating, like Chinese (Sun 2006) or Bai (Wiersma 2003). Since words can be easily borrowed, many linguists, including Meillet, see morphology and morphosyntax as stronger evidence for subgrouping than shared vocabulary (Nichols 1996). However, since many Sino-Tibetan languages lack complex morphology, it is difficult to assemble evidence for deeper affiliations apart from the lexicon.

As a result, proposed subgroupings for the Sino-Tibetan language family differ widely (Handel 2008), as does the evidence scholars use to support their hypotheses (LaPolla 2012). Even the seemingly robust claim that the Sinitic was the first branch to split off (Matisoff 2003; Thurgood 2003) has been challenged on the basis of morphological and lexical evidence (Blench and Post 2013; Driem 1997) or lack of positive evidence (Driem 2011; Jacques 2006). On

the other hand, new approaches to the reconstruction of Old Chinese, proposed since the late 1980s (Baxter 1992; Starostin 1989; Zhengzhang 2000) and now broadly accepted (Pan 2000; Schuessler 2007), have revolutionized the field. The new reconstructions reveal closer similarities among Old Chinese, Tibetan and Burmese (Hill 2014), and come closer to recent reconstructions of Proto-Tibeto-Burman (Matisoff 2003), so that few scholars now doubt that Sino-Tibetan is a valid family (Jacques 2015).

Given the huge degree of discord among practitioners of Sino-Tibetan linguistics, it is unlikely that the disputes will be settled anytime soon. Even two recently published phylogenetic studies, which both conclude that Sino-Tibetan originated in North China (Zhang et al. 2019; Sagart et al. 2019), will not change this situation. Although similar on the surface, the studies differ crucially with respect to the inferred divergence times and the detailed subgroupings. Given that one study used data which was collected under the assumption that Chinese was the first group to branch off the Sino-Tibetan phylogeny (Zhang et al. 2019), while the other study could confirm the outgroup role of Chinese only in 30% of the cases (Sagart et al. 2019), it is unlikely that these findings "might settle this debate" (LaPolla 2019: 45). On the contrary, we should instead hope that they will revive the debate about the origin and dispersal of Sino-Tibetan, ideally by leading to less polemic and more objective debates among scholars.

Just as open to constant discussion and improvement is the issue of phonological reconstruction and its phonetic interpretation within Chinese (Sinitic) itself. As new waves of historical evidence continue to roll in, from diverse and often startling paleographic discoveries to the study of the oldest layers of Chinese loans in neighboring Hmong-Mien, Austro-Asiatic, and Kra-Dai languages, significant and stimulating amendments to older models of the Old Chinese reconstruction are offered by scholars on an almost daily basis — some will be found in the very papers that comprise the bulk of this issue — so that even the most recent and arguably most internationally well-known system, namely, Baxter and Sagart (2014), may at times already seem somewhat antiquated to those who diligently follow all the new proposals, too numerous to list. While occasional *caveats* are voiced, insisting that "harder" evidence accumulated in a systematic and quantitatively satisfying manner (e.g. rhyme systems and phonetic series of Chinese characters) should be strictly distinguished from "softer" evidence that often forces the researcher to suggest unique and unprovable scenarios for individual items (e.g. Starostin 2015), it goes without saying that no type of evidence should be ignored, and that, as more and more pieces are accumulated, some of the proposed hypotheses will be more strongly vindicated in the future, while others might be discarded in favor of more convincing interpretations.

This volume can be seen as an attempt to contribute to the debate on Sino-Tibetan historical linguistics — but not in form of offering big picture solutions that promise to end all debates once and forever, and rather by showing how research in different aspects of Sino-Tibetan linguistics and its different subfields can be improved. The volume emerged from a conference titled "Old Chinese and Friends", held in April 26/27 2018 in Jena at the Max Planck Institute for the Science of Human history as part of the research project "Computer-Assisted Language Comparison" (http://calc.digling.org) funded from 2017 to 2022 by the European Research Council in form of an ERC Starting Grant. The conference itself was conceptualized as the successor of a conference titled "Recent Advances in Old Chinese Phonology", held three years earlier at the School of African and Oriental Sciences in London (November 5-6, 2015), organized by Nathan W. Hill as part of the ASIA research project, funded from 2016 until 2021 by the European Research Council in the form of an ERC Synergy Grant.

The idea of both conferences was to reevaluate and reconsider the role of Old Chinese phonology in particular and Sinitic languages in general for the historical-comparative study of Sino-Tibetan languages. In this sense, both conferences stay in contrast to the paradigm of the Tibeto-Burman school, most prominently represented by the Sino-Tibetan Etymological Dictionary and Thesaurus (STEDT) project (Matisoff 2015), which has neglected or downplayed the importance of Sinitic to understand the history of Tibeto-Burman languages. Given the diversity of invited speakers on both events, reflecting a broad range of opinions, however, neither of the two events was planned as defending a specific agenda in Sino-Tibetan historical linguistics. Instead, the organizers were hoping to have fruitful discussions about new perspectives on the well-known problems of comparative research in the Sino-Tibetan area.

While quite a few of the papers that were presented at the first conference devoted to Chinese and Sino-Tibetan have now been published as part of a special issue of the Bulletin of Chinese Linguistics (Volume 9, Issue 2), this volume makes public some of the papers that were presented at the second conference. As the readers will see already from the table of contents of this volume, the collection of studies is quite diverse, reflecting the broad range of research questions which the field of Sino-Tibetan and Chinese historical linguistics offers, ranging from studies on Old Chinese etymologies (Ma), the development from Old to Middle Chinese (Jin and Huo), ancient language contact of Chinese (Gong), and the evolution of the Chinese writing system, via cognates between Chinese and other Sino-Tibetan languages (Zhang, Jacques, and Lai), up to frameworks for data annotation in Chinese historical phonology (List, Hill, and Foster).

Two studies concentrate on Old Chinese phonology, using different sources of evidence. Jin Lixin and Huo Wenwen discuss "The Old Chinese origin of Middle Chinese voiced sibilants z/ʑ", proposing that the alveolopalatal fricative ʑ was secondarily derived from the alveodental fricative z, for which four distinct sources in Old Chinese are proposed. The "Study of phonological issues in the text variants of *Xiaochu* and *Dachu* hexagrams, *Zhouyi*" by Ma Kun presents revised etymologies for Old Chinese, following the system by Baxter and Sagart (2014). Shen Ruiqing discusses "The monosyllabicization of Old Chinese and the birth of Chinese writing", hypothesizing that the Chinese language coevolved along with its writing system. Zhang Shuya, Guillaume Jacques, and Lai Yunfan present "A study of cognates between Gyalrong languages and Old Chinese", which proposes a range of new Sino-Tibetan etymologies. Xun Gong presents a study on "Chinese loans in Old Vietnamese with a sesquisyllabic phonology", proposing new evidence for sesquisyllabic words in Chinese based on a careful examination of Old Vietnamese. Johann-Mattis List, Nathan W. Hill, and Christopher J. Foster present their ideas in "Towards a standardized annotation of rhyme judgments in Chinese historical phonology (and beyond)", emphasizing the importance of a careful annotation of scholarly judgments on rhyming for the reconstruction of Old Chinese phonology and its historical stages. Another paper from the same conference, George Starostin's "Chinese basic lexicon from a diachronic perspective", shall be published separately in the next volume of the Journal for reasons of space as well as thematic distance (unlike the others, it concentrates on Chinese historical lexicology rather than phonological issues, but may just as well have implications for the entire field of Sino-Tibetan studies).

Despite the diversity of topics and methodological considerations discussed in this volume, it is clear that we only touch the top of the iceberg with respect to the diversity of Sino-Tibetan historical linguistics in general. We know well that this collection cannot give a full account on the field (which would be impossible, given the limited space). Instead, we

hope it can present readers interested in historical language comparison with some insights into the fascinating diversity of research questions and research methods which the fields of Sino-Tibetan linguistics and Chinese historical linguistics have to offer.

Johann-Mattis LIST, LAI *Yunfan, George* STAROSTIN

References

Abramson, A. S. 2004. The Plausibility of Phonetic Explanations of Tonogenesis. In: Gunnar Fant, Hiroya Fujisaki, Cao Jianfen, Xu Yi (eds.). *From Traditional Phonology to Modern Speech Processing: Festschrift for Professor Wu Zongji's 95th Birthday*: 17–29. Beijing: Foreign Language Teaching; Research Press.

Anthony, David W., Don Ringe. 2015. The Indo-European Homeland from Linguistic and Archaeological Perspectives. *Annual Review of Linguistics* 1: 199–219.

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: De Gruyter.

Baxter, William H., Laurent Sagart. 2014. *Old Chinese. A New Reconstruction*. Oxford University Press.

Blench, Roger, Mark W. Post. 2013. Rethinking Sino-Tibetan Phylogeny from the Perspective of North East Indian Languages. In: Thomas Owen-Smith, Nathan W. Hill (eds.). *Trans-Himalayan Linguistics*: 71–104. Berlin/New York: Mouton de Gruyter.

Blevins, Juliette. 2004. *Evolutionary Phonology. The Emergence of Sound Patterns*. Cambridge University Press.

Blust, Robert. 2019. The Austronesian Homeland and Dispersal. *Annual Review of Linguistics* 5: 417–34.

Driem, George van. 1997. Sino-Bodic. *Bulletin of the School of Oriental and African Studies* 60 (3): 455–88.

Driem, George van. 2011. Tibeto-Burman Subgroups and Historical Grammar. *Himalayan Linguistics* 10(1): 31–39.

Driem, George van. 2014. Tibeto-Burman. In: William S.-Y Wang, Sun Chaofen (eds.). *The Oxford Handbook of Chinese Linguistics*: 135–48. Oxford/New York: Oxford University Press.

Ebert, Karen H. 2003. Kiranti Languages: An Overview. In: Graham Thurgood, Randy J. LaPolla (eds.). *The Sino-Tibetan Languages*: 3–21. London/New York: Routledge.

Handel, Zev. 2008. What Is Sino-Tibetan? Snapshot of a Field and a Language Family in Flux. *Language and Linguistics Compass* 2(3): 422–41.

Haudricourt, André-Georges. 1954. De l'origine des tons en viêtnamien. *Journal Asiatique* 242: 69–82.

Hill, Nathan W. 2014. A Note on Voicing Alternation in the Tibetan Verbal System. *Transactions of the American Philosophical Society* 112(1): 1–4.

Jacques, Guillaume. 2006. *La morphologie du Sino-Tibétain*. Ms.

Jacques, Guillaume. 2015. The Genetic Position of Chinese. In: Rint Sybesma et al. (eds.). *Encyclopedia of Chinese Language and Linguistics*. Leiden/Boston: Brill.

Kiparsky, Paul. 1988. Phonological Change. In: Frederick J. Newmeyer (ed.). *Linguistics: The Cambridge Survey* 1: 363–415. Cambridge University Press.

Labov, William. 1981. Resolving the Neogrammarian Controversy. *Language* 57(2): 267–308.

LaPolla, Randy. 2019. The Origin and Spread of the Sino-Tibetan Language Family. *Nature* 569: 45–47.

LaPolla, Randy J. 2012. Comments on Methodology and Evidence in Sino-Tibetan Comparative Linguistics. *Language and Linguistics* 13(1): 117–32.

Lee, Yeon-Ju, Laurent Sagart. 2008. No Limits to Borrowing: The Case of Bai and Chinese. *Diachronica* 25(3): 357–85.

List, Johann-Mattis. 2015. Network Perspectives on Chinese Dialect History. *Bulletin of Chinese Linguistics* 8: 42–67.

List, Johann-Mattis. 2017. Contraction. In Rint Sybesma et al. (eds.). *Encyclopedia of Chinese Language and Linguistics*: 1: 672–75. Leiden/Boston: Brill.

Matisoff, J. A. 2000. On the Uselessness of Glottochronology for the Subgrouping of Tibeto-Burman. In: Colin Renfrew, April McMahon, Larry Trask (eds.). *Time Depth in Historical Linguistics*: 333–71. Cambridge: McDonald Institute for Archaeological Research.

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

Matisoff, James A. 2015. *The Sino-Tibetan Etymological Dictionary and Thesaurus Project*. Berkeley: University of California.

Meillet, Antoine. (1925) 1954. *La méthode comparative en linguistique historique*. Reprint. Paris: Honoré Champion.

Nichols, Johanna. 1996. The Comparative Method as Heuristic. In: Mark Durie (ed.). *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*: 39–71. New York: Oxford University Press.

Norman, Jerry. 2003. The Chinese Dialects: Phonology. In: Graham Thurgood, Randy J. LaPolla (eds.). *The Sino-Tibetan Languages*: 72–83. London/New York: Routledge.

Pān Wùyún (潘悟云). 2000. *Hànyǔ lìshǐ yīnyùnxué*. Shànghǎi: Shànghǎi jiàoyù (上海教育).

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin Ryder, Valentin Thouzeau, Simon J. Greenhill, Johann-Mattis List. 2019. Dated Language Phylogenies Shed Light on the Ancestry of Sino-Tibetan. *Proceedings of the National Academy of Science of the United States of America* 116(21): 1–6.

Satterthwaite-Phillips, Damian. 2011. Phylogenetic Inference of the Tibeto-Burman Languages or on the Usefulness of Lexicostatistics (and "Megalo"-Comparison) for the Subgrouping of Tibeto-Burman. PhD thesis, Stanford: Stanford University.

Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawaï'i Press.

Starostin, George. 2015. Review of William H. Baxter, Laurent Sagart: Old Chinese. A New Reconstruction. *Journal of Language Relationship* 13(4): 383–389.

Starostin, Sergej A. (1995) 2007. The Historical Position of Bai. In: *S. A. Starostin: Trudy Po Jazykoznaniju*: 580–90. Moscow: Languages of Slavic Cultures.

Starostin, Sergej A. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy (Reconstruction of the Phonological System of Old Chinese)*. Moscow: Nauka.

Sun, Chaofen. 2006. *Chinese: A Linguistic Introduction*. Cambridge University Press.

Thurgood, Graham. 2003. A Subgrouping of the Sino-Tibetan Languages: The Interaction Between Language Contact, Change, and Inheritance. In: Graham Thurgood, Randy J. LaPolla (eds.). *The Sino-Tibetan Languages*: 3–21. London/New York: Routledge.

Trask, Robert L. (ed.). 2000. *The Dictionary of Historical and Comparative Linguistics*. Edinburgh University Press.

Wang, Feng. 2006. *Comparison of Languages in Contact. The Distillation Method and the Case of Bai*. Taipei: Institute of Linguistics Academia Sinica.

Wiersma, Grace. 2003. Yunnan Bai. In: Graham Thurgood, Randy J. LaPolla (eds.). *The Sino-Tibetan Languages*: 649–73. London/New York: Routledge.

Zhang, Menghan, Shi Yan, Wuyun Pan, Li Jin. 2019. Phylogenetic Evidence for Sino-Tibetan Origin in Northern China in the Late Neolithic. *Nature* 569: 112–15.

Zhengzhang, Shangfang. 2000. *The Phonological System of Old Chinese*. Translated by Laurent Sagart. Paris: École des Hautes Études en Sciences Sociales.

*J*IN *Lixin*[†], *H*UO *Wenwen*[‡]

[†] School of Humanities, Wenzhou University, Zhejiang; ytjlx@hotmail.com

[‡] School of Humanities, Wenzhou University, Zhejiang; hww13736970100@163.com

The Old Chinese origin of Middle Chinese voice sibilants *z/ʐ*

This paper discusses the Old Chinese origin of voiced sibilants *z-* and *ʐ-* in Middle Chinese. First, based on careful examination of *Guǎngyùn* (广韵) and other texts, we argue that, although the distinction between voiced affricates and fricatives was largely kept in the *Guǎngyùn* system, in some words the voiced affricates *dz-/dʐ-* had already merged into fricatives *z-/ʐ-* due to copying from other texts and/or sound change. Second, we argue that *z-* and *ʐ-* in Middle Chinese are either in complementary distribution or show free variation in some occasions. Therefore, they had one single origin (/z/) before the onset of the sound change *z-* > *ʐ-*. Based on these arguments, we have carried out a detailed analysis of the *xiéshēng* series and found four different origins for /z/. They are **r-j-*, **s-ɢ-*, **z-*, and **s-d-* respectively.

Keywords: Old Chinese phonology; historical reconstruction; sound change; coronal affricates.

I

Unlike stop consonants, the fricatives in Chinese often have comparatively complicated origins, with the voiced fricative *z-/ʐ-* at the top of the list. First of all, a small number of syllables with affricate initials *dz-/dʐ-* merge with those with fricative initials *z-/ʐ-*. Already in the *Jīngdiǎn shìwén* (经典释文), compiled by Lù Démíng 陆德明, there are traces of their merging (Wáng 1984). The *Yùnjìng* (韵镜) is exceptional in this respect, as it misarranges the sequence for the affricate *dʐ-* and the fricative *ʐ-*. Second, in the rhyme dictionary *Guǎngyùn* (广韵), although the affricates and fricatives are distinct phonemes, there is occasional confusion of the two sets. This can be explained in part by assuming that the compilers were negligent in differentiating the data when copying from previous rhyme dictionaries and character dictionaries, and in part by the factor of sound change. On the other hand, there are cases of fricatives *z-/ʐ-* merging with affricates *dz-/dʐ-*. Such mutual developments between fricatives and affricates are common in Chinese and even Sino-Tibetan languages; we shall not expound on this here with concrete examples.

Lǐ Fāngguì 李方桂 (1980: 16) has proposed that the palatal fricative *ʐ-* is an allophone of the affricate *dʐ-*, with a shared origin in Old Chinese. This viewpoint is rarely supported by Chinese phonologists. Even though there are traces of confusion between *ʐ-* and *dʐ-* in *Guǎngyùn*, they are distinctly different from each other. More importantly, each of them corresponds to different *xiéshēng* series in Old Chinese. Their *xiéshēng* sets are conspicuously different from each other. The initials of the *xiéshēng* sets of the affricate *dʐ-* include *t-*, an alveolar stop, and *tɕ-*, a palatal affricate, whereas *t-* and *tɕ-* rarely occur in *xiéshēng* sets of *ʐ-*. Their demarcation is clear. Thus, Old Chinese fricative *ʐ-* and affricate *dʐ-* should originate from different consonants in Old Chinese. We cannot jump to the conclusion that they are reflexes of the same initial consonant in Old Chinese based on sporadic cases of their confusion as reflected in *Guǎngyùn*.

II

The fricative *z-* in Middle Chinese, different from sounds like *ts-/tsh-/dz-/s-*, which share the same place of articulation with it, can co-occur only with finals that have the main vowel *-i-* or the medial *-i-*, that is, *z-* is an initial consonant occurring in a specific phonetic environment. This is explainable only in two ways: either *z-* changes into other consonants in certain environments, parallel with the velar stop *g-* in Old Chinese, or it has an origin different from *ts-/tsh-/dz-/s-* in Old Chinese. The alveolars *ts-/tsh-/dz-/s-* belong in the same *xiéshēng* series with the retroflexes *tʂ-/ tʂh-/dʐ-/ʂ-*, forming a relatively exclusive *xiéshēng* set. Except for sporadic *xiéshēng* contacts with the alveolar fricative *s-*, initial *z-* forms no *xiéshēng* series with alveolar affricates *ts-/tsh-/dz-*, and never occurs together with retroflex *tʂ-/tʂʰ-/dʐ-/ʂ-*. The *xiéshēng* behavior of *z-* parallels that of the semi-vowel *j-* in that it forms *xiéshēng* series with alveolar stops *th-/d-* and the semi-vowel *j-*. Therefore, Lǐ Fānguì 李方桂 (1980: 14) holds that *z-* results from the same Old Chinese consonant group that also gave rise to *j-*, instead of *ts-/tsh-/dz-*.

The Middle Chinese consonant *ʑ-* is distinct from *tɕ-/tɕh-/dʑ-* (affricates of the same place of articulation), and they generally do not occur within the same *xiéshēng* sets. The alveolars *t-/th-/d-/n-* in Old Chinese, when co-occurring with the palatal medial *-i-* (*-j-*), evolve into Middle Chinese palatals *tɕ-/tɕʰ-/dʑ-/ɲ-*. In the same vein, the alveolar fricatives *s-/z-* in Old Chinese, when co-occurring with *-i-* (*-j-*), ought to evolve into palatal fricatives *ɕ-* and *ʑ-* in Middle Chinese, that is to say, the earlier forms of Middle Chinese *ɕ-* and *ʑ-* should be the alveolar fricatives *s-/z-*. The palatal fricative *ʑ-* and the alveolar fricative *z-* are usually in complementary distribution, and even when they are in contrastive distribution, they are still but phonetic variations; moreover, they share the same *xiéshēng* behavior. Hence, Jīn Líxīn 金理新 (2002: 237–238) argues that the palatal fricative *ʑ-* and the alveolar fricative *z-* are allophones of the same consonant initial. According to the chronological order of sound change, the palatal fricative *ʑ-* in Middle Chinese should be a later form evolved from the alveolar fricative *z-*, viz. *z-* > *ʑ-*.

The two fricatives *z-* and *ʑ-* are allophones of the exact same consonant initial. That is why they show identical *xiéshēng* behavior and are in complementary distribution. The two variants are in contrastive distribution only in the case where one phonetic variant of *shè* 射 ('shoot') and one of *shè* 麝 ('musk') are distinct from *xiè* 谢 ('thank') and *xiè* 榭 ('pavillion'), all with the phonetic/phonological radical *shè* 射 ('shoot'). In the *Guǎngyùn shè* 射 has four pronunciations, including *ʑiɛk*, *jiɛk*, *ʑiaᶜ* and *jiaᶜ*, and *shè* 麝 has two pronunciations, *ʑiɛk* and *ʑiaᶜ*. According to the *Jīngdiǎn shìwén* (经典释文), as a verb, *shè* 射 is pronounced as *ʑiɛk*, whereas as a noun, it is pronounced as *ʑiaᶜ*. It is evident that *ʑiaᶜ* is a later aberrant form, distinct from the most common, customary pronunciation because of the difference in meanings. Parallel to this, the two distinct pronunciations of *shè* 麝 (*ʑiɛk* and *ʑiaᶜ*) have been developed to differentiate 'musk' from 'musk-producing animal'. Except for the two aberrant pronunciation variants, the alveolar fricative *z-* from the same origin does not occur in contrastive distribution with the palatal fricative *ʑ-*. Under the same phonetic conditions, *z-* never occurs where there is *ʑ-*, and vice versa.

In the *xiéshēng* sets formed with above-mentioned phonetic radicals, we sporadically observe velar stops, but no traces of palatal affricates can be seen, neither can the alveolar stops be encountered (including the three initials *zhī* 知, *chè* 彻 and *chéng* 澄). According to Dǒng Tónghé's (1948: 33) Old Chinese reconstruction system, in those cases where the semi-vowel *j-* forms part of a *xiéshēng* series with velars, it should be reconstructed as **g-*. However, the fricative *z-* within the same *xiéshēng* sets is still reconstructed as **z-* in Dǒng's system. Obviously, this reconstructed form fails to take into account the *xiéshēng* connection between the

Table 1. Xiéshēng behavior of the fricatives z- and ʑ-

	z-	j-	k-	kh-	g-	t-	th-	d-	ṭ-	ṭh-	ḍ-	tɕ-	tɕh-	dʑ-	ɕ-	ʑ-
頤	洍	頤	姬													
食	食	食													飾	食
蠅		蠅														繩
孕		孕														
囟	囟	游														
叀		叀														
與	嶼	與		舉												
牙	邪	邪														
亦		亦														
夜		夜														
夕	夕															
射	榭	射														射
羊	祥	羊	姜		羌											
贏		贏														
衍		衍		愆								(餐)				
沿		沿														船
兗		兗														
允		允														吮
寅		寅														
壬		壬														
䛬		䛬														
裔		裔														

fricative *z-* and the semi-vowel *j-*. Lǐ Fānguì (1980: 60) reconstructs *yáng* 羊 and *xiáng* 祥 respectively as **raŋ* and **rjaŋ*. Similar to Dǒng's reconstruction, Li's reconstructed forms are also incapable of explaining the *xiéshēng* connection between *yáng* 羊 **raŋ*, *xiáng* 祥 **rjaŋ* and *jiāng* 姜 **kjaŋ*. Jīn Lǐxīn (2013: 283–294) discusses the pronunciations of the Old Chinese antecedents of the Middle Chinese fricative *z-/ʑ-* and the semi-vowel *j-* in the above-mentioned *xiéshēng* sets. He proposes that, as can be seen from those *xiéshēng* sets, the Old Chinese origin of Middle Chinese *j-* is still a semi-vowel, reconstructed as **j-*, while that of the Middle Chinese fricative *z-/ʑ-* is a semi-vowel **j-* with a prefix **r-* (the velar stop k- which forms the same *xiéshēng* series with it originates from Old Chinese **kj-*, where **k-* is a prefix). These reconstructed forms adequately settle the *xiéshēng* relationship between all involved characters. Thus, *yáng* 羊, *xiáng* 祥 and *jiāng* 姜 form *xiéshēng* series only because the three share an identical root **jaŋ*, which is explicit from their reconstructed forms **jaŋ*, **r-jaŋ* and **k-jaŋ*. We shall not expound on this issue any further here; however, besides **rj-*, there are other Old Chinese origins for Middle Chinese fricatives *z-/ʑ-*.

Within the *xiéshēng* series in Table 2, all of them representing a typical harmonic set of soft palate sounds, we only encounter those consonants which evolve into Middle Chinese velars *k-/kh-/g-/x-/ɣ-*, and never those which develop into Middle Chinese alveolar stops *t-/th-/d-*, palatal affricates *tɕ-/tɕʰ-/dʑ-*, or palatal fricative *ɕ-*. As to the finals, outside of those *yùnbù*'s (Old Chinese rhyme categories) which do not develop into Ancient Chinese *shè*'s (Middle Chi-

Table 2. Xiéshēng behavior of the fricatives z- and ʑ-

	j-	s-	z-	k-	kh-	g-	x-	ɣ-	t-	th-	d-	tɕ-	tɕh-	dʑ-	ɕ-	ʑ-
久	羑			久		柩										
公	頌	鬆	訟	公												
容	容		松													
雈	萑			榷	推		膻	雈								
谷	谷		俗	谷												
頃	潁			潁	頃											
矞	矞			橘												潏
唯	唯	雖			匯		睢	淮								
睿	睿	䀼	璿													
旋			旋													
敫	敫			激	竅			檄								
役	役						殳	椴								
肙	捐		圓	涓		蜎	圓	鞙								
熒	營					熒		熒								
彗		雪	彗				嘒	慧								
尹	尹							靬								
勻	勻	昀	昀	均				昀								
旬		詢	旬				眴	眴								
瞏		瞏	還	擐			嬛	還								
穴	鴪						沈	穴								祋

nese rhyme categories) that contain both *kāikǒu* 开口 (open) and *hékǒu* 合口 (closed) rhymes, all the *yùnbù*'s which can co-occur with the consonants are those which evolve into Middle Chinese *she*'s with *hékǒu* rhymes and can be encountered in characters outside of the 3rd division, such as 谷 *gǔ*. Jīn Lǐxīn (2013: 473-483) provides a detailed discussion of the pronunciations of the Old Chinese ancestors for this groups of initials. The Middle Chinese fricatives z-/ʑ- evolve from the Old Chinese voiced uvular stop *ɢ- with a prefix *s-, and the vowels co-occurring with it are tense vowels.

<div align="center">

III

</div>

Besides the two origins touched upon in the previous section, a third origin of the Middle Chinese fricatives z-/ʑ- is the Old Chinese fricative *z-. This ancestor of z-/ʑ- forms *xiéshēng* series with those Old Chinese consonants which change into Middle Chinese alveolar stops *th-/d-* (the unaspirated stop *t-* and palatal affricates *tɕ-/tɕh-/dʑ-* are not encountered in these *xiéshēng* series). The *xiéshēng* series for this variant of z-/ʑ- are distinct from those of the previous two. The Middle Chinese reflexes of this *xiéshēng* series include altogether seven consonants, namely, *th-/d-*, *ɕ-/ʑ-/j-* and *s-/z-*. For details, Jīn Lǐxīn (2013: 377–379) can be referred to. First of all, we shall examine the distribution of the three initials z-/ʑ-/j-.

Table 3. The distribution of the initials *z-/ʑ-/j-* across rhyme groups

	之			职		蒸		
	平	上	去	去	入	平	上	去
j	台	以	异		弋			膡¹
ʑ								膡
z		似						

	幽			觉		冬		
	平	上	去	去	入	平	上	去
j	攸	誘	柚			融		
ʑ								
z			袖²					

	宵			药				
	平	上	去	去	入	平	上	去
j	姚		筊	燿	躍			
ʑ								
z								

	侯			屋		东		
	平	上	去	去	入	平	上	去
j	俞	瘉	覦			庸	勇	用
ʑ								
z								诵

	鱼			铎		阳		
	平	上	去	去	入	平	上	去
j	余	予	豫		譯	易		煬
ʑ		抒						
z	徐	抒					象	

	支			锡		耕		
	平	上	去	去	入	平	上	去
j	㦤	㦤³		易	易	盈	郢	
ʑ		鍚						
z								

---

¹ The character *shèng* 膡 has two variant pronunciations in the *Guǎngyùn*: *jiŋ^C* and *ʑiŋ^C*, both sharing the same meaning.

² According to the pronunciations provided in *Jīngdiǎn shìwén*, the characters 袖, 岫 (*xiù*) have two pronunciation variants, *ziu^C* and *jiu^C*.

³ The character *yí* 㦤 is attested only in the *Shuōwén*, without any attestation in any other document from the pre-Qin era to the two Han dynasties.

	歌			月		元		
	平	上	去	去	入	平	上	去
j	蛇	池	施	扡	扡⁴	延		羨
ʑ	(蛇)				舌			
z						次	繇	羨⁵

	歌			月		元		
	平	上	去	去	入	平	上	去
j	蠵⁶	蠵		鋭	說	緣		掾
ʑ								
z	隋							

	微			物		文		
	平	上	去	去	入	平	上	去
j	遺	壝	遺					
ʑ					術			
z				豙				

	脂			質		真		
	平	上	去	去	入	平	上	去
j	夷		肆		逸		引	蚓
ʑ						神		
z		兕						

				葉		談		
	平	上	去	去	入	平	上	去
j					葉	(閻)		
ʑ								
z						燅		

				緝		侵		
	平	上	去	去	入	平	上	去
j				隶		撢⁷	潭	
ʑ								
z					隰	尋		

---

⁴ The character *yè* 抴 has two pronunciations in the *Guǎngyùn*: *jiɛi*^{C} and *jiɛt*^{D}. The former pronunciation makes it a character variant of *yè* 曳 (拽), whose pronunciation in the *Guǎngyùn* and in the *Jīngdiǎn shìwén* is also *jiɛi*^{C}.

⁵ The characters *xiàn* 羨 and *yàn* 遪 have two pronunciation variants, *ziɛn*^{C} and *jin*^{C}, in the *Guǎngyùn*.

⁶ The character *wéi* 蠵 is attested only in the *Shuōwén*, without any attestation in any other document from the Pre-Qin era to the two Han dynasties.

⁷ The character *dǎn* 撢 has two pronunciation variants, *jim*^{A} and *thom*^{A}, in the *Guǎngyùn*; likewise, *xín* (or *tán*, or *chán*) 镡 and *xún* 鱏 also have two pronunciation variants, *jim*^{A} and *zim*^{A}.

The fricatives *z*- and *ʑ*- are contrastive only in one case; otherwise, they are in complementary distribution. Three characters with the fricative initial *ʑ*- (*shū* 纾, *shū* 抒 and *zhù* 杼) are in contrastive distribution with *shū* 舒 that shows an initial *z*-; all four characters belong to the Old Chinese *yú* 鱼 rhyme category and to the Middle Chinese *yǔ* 语 rhyme group with *shǎng-shēng* (上声). The characters *shū* 纾, *shū* 抒 and *zhù* 杼 all have variant pronunciations. In the *Guǎngyùn*, the two pronunciations of *shū* 纾 are *ʑio^B* and *ɕio^A*. According to the *Jīngdiǎn shìwén*, *shū* 纾 should be pronounced *ɕio^A* [8]. In the *Guǎngyùn*, the two pronunciations of *shū* 抒 are *zio^B* and *ʑio^B*. According to the glossing of *shū* 抒 in the *Shuōwén*, the *Jīngdiǎn shìwén*, and the *Guǎngyùn*, *zio^B* and *ʑio^B* are merely phonetic variants without any difference in meaning [9]. The character *zhù* 杼 has two pronunciations in the *Guǎngyùn*: *dio^B* and *ʑio^B*. Except for these three characters with variant pronunciations, the two fricatives *z*- and *ʑ*- are generally complementary. It can be safely inferred that they represent variants of the same initial.

When the fricative *z*- is palatalized, it changes into the fricative *ʑ*-, which in turn is prone to more change, resulting in the semi-vowel *j*-. The reverse cycle of sound change between these three sounds is also quite common. The semi-vowel *j*- in Chinese is reflected in most dialects of the Hmong language as *ʑ*- (for example, 羊 is pronounced *ʑoŋ^2* in Yǎnhāo language), whereas in Sino-Vietnamese the typial reflex is *z*-, e.g. 养 is pronounced *zaŋ^4*. In some Kam-Tai languages they even shift to alveolar plosives, e.g. 养 is pronounced *t:aŋ^4* in the Mulam language.

As we have mentioned many times in previous sections, the two fricatives *z*- and *ʑ*- are in complementary distribution. Apart from the two origins mentioned in section 2, that is, *rj- and *sG-, and several characters with initials *z*- and *ʑ*-, which evolve from the alveolar stops, *z*-/*ʑ*- are in obvious complementary distribution with the semi-vowel *j*-. Syllabically, the fricatives *z*-/*ʑ*- are in contrastive distribution with the semi-vowel *j*- only in a few cases. Among these exceptional cases, except for two characters *yǐ* 以 and *sòng* 诵, which do not have variant pronunciations, all the others are characters with variant pronunciations, the majority of which are sheer variants without any difference in meaning. Aside from these characters with variant pronunciations, the fricatives *z*-/*ʑ*- are complementary with the semi-vowel *j*-. Moreover, *z*-/*ʑ*- and *j*- also form phonetic variants; such variant pronunciations are common in Ancient Chinese rhyme dictionaries such as the *Guǎngyùn* (see Table 4).

As we have indicated, these variant pronunciations, most of which are not accompanied by any difference in meaning, only reflect sound variations in different dialects, or sound change in different historical periods. During the Later Hàn 汉 Dynasty and the Three Kingdoms period, the semi-vowel is already used to transcribe *y*- in Sanskrit, and semi-vowels from two different origins have already merged (for detailed reference, see Yú Mǐn 俞敏 1999: 59). Therefore, the variant pronunciations mentioned above can be explained as sound change from different historical periods. The fricatives *z*-/*ʑ*- and the semi-vowel *j*- seem to be contrastive only in rhyme *yú* 鱼 from Old Chinese rhyme category *yú* 鱼 (*yú bù* 鱼部): the character *xú* 徐 has the fricative *z*- as the initial, while *yú* 余 has the semi-vowel *j*- as the initial. In light of the examination of the usage of the character *xú* 徐 in historical documents and of the pronunciations given for this character by teachers of Confucian Classics from various dynasties, it can be ascertained that the initial of *xú* 徐 in the *Guǎngyùn* is the fricative *z*-, which is a retention from Old Chinese pronunciation; we shall not dwell on this here.

---

[8] In the *Zuǒzhuàn* (左传)：令尹自毁其家以纾楚国之难 "The chief sacrificed his home so as to ease the pain of the state of Chu". *Jīngdiǎn shìwén*：纾, 音舒, 一音直汝反.

[9] In the *Zuǒzhuàn*：难必抒矣 "The trouble must be relieved". *Jīngdiǎn shìwén*：抒，直吕反，又时吕反.

Table 4. Examples of variant pronunciations for *z-/ʐ-* and j-

Character		Pronunciation with *j-*		Pronunciation with *z-/ʐ-*	
羠	*yí*	以脂切	广雅云：犍羊也 'wether'	徐姊切	犍羊 'wether'
妯	*yòu*	余救切	牛黑眦'black-canthused oxen'	似佑切	牛黑眦'black-canthused oxen'
改	*yǐ*	羊已切	大坚 'adornments worn by offi-cials in the Han dynasty to drive off evil spirits'	详里切	
巳	*yǐ*	羊已切	止也. 此也, 甚也, 讫也,又音似 'stop; here / this; very much; come to an end; another pro-nunciation is *sì*	详里切	
镡	*xín*	余针切	剑鼻'sharp tip of a sword'	徐林切	剑鼻'sharp tip of a sword'
鱏	*xún*	余针切	鱼名 'name of fish'	徐林切	鱼名 'name of fish'
勷	*xiàng*	余两切	勉也 'endeavor'	徐两切	勉也 'endeavor'
斜	*xié*	以遮切	斜谷, 说文抒也 'place name; *shū* 'push, scoop' in the Shuōwén'	似嗟切	上同（邪）'same as the above (*xié*)'
蔛	*yé*	以遮切	穗也 'ear of a grain, spike, tassel'	似嗟切	蔛蒿 'a kind of artemisia'
爓	*yàn*	以赡切	光也 'light (n.)'	徐盐切	同燅 'same as xún 'cook with fire, half-cooked meat''
羡	*xiàn*	予线切	延也, 进也 'invite to enter, let enter'	似面切	贪慕, 又馀也'envy, admire; also 'have extra''
递	*yàn*	予线切	移也 'move'	似面切	遮也 'shelter, cover'
蛇	*yí*	弋支切	蜲蛇 'move in a winding way'	食遮切	毒虫 'poisonous insect'
貤	*yí*	羊至切	重物次第 'sequence of repetition'	神至切	重物次第
野	*yě*	羊者切	田野, 说文云郊外也'field; *Shūowén* says 'at the outskirts''	承与切	田野 'field'
鱅	*yōng*	余封切	鱼名 'name of fish'	蜀庸切	鱼名 'name of fish'
鋋	*chán*	以然切	小矛 'small spear'	市连切	小矛 'small spear'
縢	*shēng*	以证切	增益, 一曰送也, 又物相赠 'increase, or send; also also pre-sent gifts'	实证切	
媵	*yìng*	以证切	送女从嫁 'send a daughter to the in-laws (to get married)'	实证切	

The three Middle Chinese initials *z-/ʐ-/j-* come from the same Old Chinese consonant ini-tial. For the sake of differentiating the fricatives *z-/ʐ-* and the semi-vowel *j-*, Jīn Lǐxīn (2013: 377–385) reconstructs the Old Chinese origin of *z-/ʐ-* as a voiced fricative *z- with a prefix *m-. It is obvious that if the Middle Chinese fricatives *z-/ʐ-* have a different Old Chinese origin from the semi-vowel *j-*, they will not be in complementary distribution across syllables. As such, the reconstructed prefix *m- should be deleted. The common origin of the alveolar fricative *z-*, the palatal fricative *ʐ-* and the semi-vowel *j-* in Middle Chinese is the Old Chinese alveolar *z- which co-occurs with tense vowels:

场 MC *daŋ* < OC*zaŋ;
汤 MC *thaŋ* < OC*saŋ;
易 MC *jiaŋ* < OC*zaŋ;
汤 MC *ɕiaŋ* < OC*saŋ.

IV

Besides the three origins discussed in sections 2 and 3, the fricatives *z-/ʑ-* had a fourth origin in Old Chinese.

Table 5. Xiéshēng behavior of the fricatives z- and ʑ-.

	j-	s-	z-	k-	kh-	g-	x-	ɣ-	t-	th-	d-	tɕ-	tɕh-	dʑ-	ɕ-	ʑ-
習	熠		習								慴	慴	謵	褶		
襲			襲										矗			
寺			寺						等		待	時		時	詩	
詹	檐								擔	𧤺	澹	詹	幨	贍		
占	阽								坫	沾		占			苫	
勺	约								钓			勺		勺		
川			巡				训			釧			川			舛
盾			循						幁	楯	盾					盾
甚	醓								娸	黮	湛	斟	觇	甚		
尤	尤								肬		沈	枕		訧	沈	
多	移								多		趍	眵	侈	哆		
召	軺								貂	超	召	招	怊	召		
炎	炎									毯	谈		裧	剡		
者			緒						都	楮	屠	诸		署	暑	
粥	鬻											粥				
育	育															
辰												振		辰	娠	唇
酉	酉												醜			

The character *yuē* 约 has two pronunciations in the *Guǎngyùn*: *iak* and *iau^C*. The relationship between the two pronunciations is such that the former is a verb while the latter is a noun. The *Shuōwén* describes it as 约, 从勺声*yuē, cóng sháo shēng* "*yuē* follows phonetic radical *sháo* 勺". In the *Guǎngyùn*, *yuē* is listed under the initial *yǐng* 影 (zero initial), reflecting an irregular sound change. When rhymes of division III (三等韵 *sānděngyùn*) occur in conjunction with zero initial, if the medial *-i-* is pronounced with a stronger than normal friction, it changes into the semi-vowel *j-*. As a consequence, characters with initial *yǐng* 影 (zero initial) change into characters with initial *yù* 喻 (*j-*). A case in point is *yì* 溢, pronounced as *jit* in the *Guǎngyùn*. *Yi* is a later derivate of the character *yì* 益, which is pronounced as *iak* in the *Guǎngyùn*. Conversely, when the semi-vowel *j-* is reduced in strength of friction, it changes into initial *yǐng*. As a result, the initial which should be the semi-vowel *j-* changes to a zero initial, i.e. *jiɑk > *iak*.

As can be seen from Table 5, the *xiéshēng* set formed with the phonetic radicals in the leftmost column is a typical set largely comprised of alveolar stops and palatal affricates. What is slightly different is that the *xiéshēng* set also includes fricatives *z-/ʑ-* and semi-vowel *j-*, which do not form *xiéshēng* sets series with palatal affricates *tɕ-/tɕ^h-/dʑ-* and the alveolar stop *t-*. It may thus be inferred that the fricatives *z-/ʑ-* and the semi-vowel in the above *xiéshēng* sets has an Old Chinese origin different from what we have talked about in the previous sections. Their root consonant in Old Chinese should be the alveolar stop *d-.

The fricative *z-* is in complementary distribution with *ʐ-*, but in contrastive distribution with the palatal affricate *dʑ-*, e.g. *sì* 寺 vs. *shì* 寺. Old Chinese possessed a very flexible prefix *s- which could occur before velar and uvular stops, so, naturally, there is no reason why it could not occur before alveolar stops. Thus, this Old Chinese origin of the fricatives *z-/ʐ-* is reconstructed as *sd- in Jīn Lǐxīn 2013: 352. In light of the *xiéshēng* series of alveolar stops, the fricatives *ʐ-* and *z-* in Middle Chinese are valid developments of *sd- in Old Chinese. In addition to fricatives *z-/ʐ-*, Jīn Lǐxīn (2013: 352) argues that *sd- is also the Old Chinese ancestral form of the Middle Chinese semi-vowel *j-* in the above *xiéshēng* sets.

The character *yì* 熠 has two pronunciations in the *Guǎngyùn*, namely *jip* (*yáng rù qiè* 羊入切) and *ɣip* (*wéi lì qiè* 为立切). The latter is a gloss reading of the character *yù* 煜 (MC *ɣip*). This character is attested in the *Shījīng* 诗经 ('The Book of Songs'): 町畽鹿场, 熠耀宵行 *tīng tuǎn lù chǎng, yì yào xiāo xíng* (from the *Dongshan* Ode). According to *Jīngdiǎn shìwén*, *yì* is pronounced as *jip* (*yǐzhí fǎn* 以执反). In addition to the two pronunciations *jip and* *ɣip*, *yì* has a third reading *zip* in the *Jíyùn* (集韵). *Yì* 熠 derives its reading from the phonetic radical *xí* 習; the semi-vowel initial *j-* is only a variant pronunciation of the fricative *z-*. This is not at all different from what we have previously observed about the variant pronunciations between the semi-vowel *j-* and the fricatives *z-/ʐ-*. The variants reflect the different stages of sound change: *z-* > *ʐ-* > *j-*.

In the *xiéshēng* sets mentioned above, the semi-vowel *j-* is contrastive relative to both the alveolar stop *d-* and the palatal affricate *dʑ-*. On the other hand, except for one character *yì* 熠 with variant pronunciations, it is thoroughly complementary with the fricatives *z-/ʐ-*. Consequently, the fricatives *z-/ʐ-* and the semi-vowel *j-* should be evolved from the exact same Old Chinese initial. This particular aspect has been discussed in detail by Jīn Lǐxīn (2013: 349–354), so it is not necessary to repeat the argumentation here.

References

Dǒng, Tónghé (董同龢). 1948. *Shàng gǔ yīn yùn biǎogǎo* (上古音韵表稿). Zhōngyāng yánjiūyuàn lìshǐ yǔyán yánjiūsuǒ jíkān (中央研究院历史语言研究所集刊).

Huáng, Xiàoshān (黄笑山). 1997. Qièyùn yúmǔ dúlì shìxī ("切韵"于母独立试析). *Gǔhànyǔ yánjiū* (古汉语研究) 3(1997): 7–14.

Jīn, Lǐxīn (金理新). 2002. *Shànggǔ hànyǔ yīnxì* (上古汉语音系). Zhèngjiāng: Huángshān shūshè (黄山书社).

Jīn, Lǐxīn (金理新). 2013. *Shànggǔ yīnlüè* (上古音略). Zhèngjiān: Huángshān shūshè (黄山书社).

Jīn, Lǐxīn (金理新). 2005. Shànggǔ hànyǔ shēngmǔ qīngzhuó jiāotì hé dòngcí de tǐ (上古汉语声母清浊交替和动词的体). *Yǔwén yánjiū* (语文研究) 1(2005): 27–31.

Jīn, Lǐxīn (金理新). 2009. Hànyǔ "sǐ" yǐjí yǐmǔ dúyīn wèntí (汉语"死"以及以母读音问题). *Yǔwén yánjiū* (语文研究) 4(2009): 38–52.

Jīn, Lǐxīn (金理新). 2016. Gǔyīn gòunǐ de yígè wèntí (古音构拟的一个问题). *Huázhōng guóxué* (华中国学) 6(2016): 28–33.

Lǐ, Fāngguì (李方桂). 1980. *Shànggǔyīn yánjiū* 上古音研究. Beijing: Shāngwù yìnshūguǎn (商务印书馆).

Luó, Chángpéi (罗常培). 1939. *Jīngdiǎn shìwén hé yuánběn Yùpiān fǎnqiè zhōng de xiá yú liǎng mǔ* (经典释文和原本玉篇反切中的匣于两母). Zhōngyāng yánjiūyuàn lìshǐ yǔyán yánjiūsuǒ jíkān(中央研究院历史语言研究所集刊).

Méi, Zǔlín (梅祖麟). 1988. Nèibù nǐgòu hànyǔ sānlì (内部拟构汉语三例). *Zhōngguó yǔwén* (中国语文) 204(3): 169–181.

Pān, Wùyún (潘悟云). 2006. Jìngzhēngxìng yīnbiàn yǔ lìshǐ céngcì (竞争性音变与历史层次). *Dōngfāng Yǔyánxué* (东方语言学) 1: 152–165.

Shěn, Jiānshì (沈兼士). 1985. *Guǎngyùn shēngxì* (广韵声系). Beijing: Shāngwù yìnshūguǎn (商务印书馆).

Shū, Zhìwǔ (舒志武). 1988. Shànggǔ hànyǔ s- qiánzhuì gōngnéng shìtàn (上古汉语 s- 前缀功能试探). *Zhōngnán mínzú xuéyuàn xuébào* (中南民族学院学报) 6(1988).

Wǎng, Fǔshì (王辅世). 1994. *Miáoyǔ gǔyīn gòunǐ* (苗语古音构拟). Dongjing: Guólì yàfēi yǔyán wénhuà yánjiūsuǒ (东京国立亚非语言文化研究所).

Wáng, Lì (王力). 1982. Jīngdiǎn shìwén fǎnqiè kǎo (经典释文反切考). In: *Lóngchóng bìngdiāo zhāi wénjí* (龙虫并雕斋文集) 3: 135–211. Beijing: Zhōnghuá shūjú (中华书局).

Wáng, Shìyuán (王士元). 1983. Yǔyán de yǎnbiàn (语言的演变). In: *Yǔyánxué lùn cóng*(语言学论丛) 11. Beijing: Shāngwù yìnshūguǎn (商务印书馆).

Wěi, Shùguān (韦树关). 2004. *Hànyuèyǔ guānxīcí shēngmǔ xìtǒng yánjiū* (汉越语关系词声母系统研究). Guangxi: Guǎngxī mínzú chūbǎnshè (广西民族出版社).

Yú, Mǐn (俞敏). 1999. *Yú Mǐn yǔyánxué lùnwénjí* (俞敏语言学论文集). Beijing: Shāngwù yìnshūguǎn (商务印书馆).

Yú, Shìcháng (喻世長). 1984. Xié-Yúxiāngtōng and Dòng-Míng Xiāngzhuǎn. In: *Yīnyùnxué Yánjiū*(音韵学研究), vol. I. Beijing: Zhōnghuá shūjú (中华书局).

Zhèngzhāng, Shàngfāng (郑张尚芳). 1997. Chóngniǔ de láiyuán jíqí fǎnyìng (重纽的来源及其反映). In: *Shèngyùn lùncóng*声韵论丛, vol. 6. Taiwan: Xuéshēng shūjú (学生书局).

Zhōu, Zǔmó (周祖谟). 1984. Hàndài zhúshū yǔ bóshū zhōng de tōngjiǎzì yǔ gǔyīn kǎo dìng (汉代竹书与帛书中的通假字与古音的考订). In: *Yīnyùnxué yánjiū* (音韵学研究), Vol. I. Beijing: Zhōnghuá shūjú (中华书局).

Цзинь Лисинь, Хо Вэньвэнь. Древнекитайские истоки среднекитайских звонких сибилянтов *z*- и ʑ-

В статье анализируются древнекитайские истоки звонких сибилянтов *z*- и ʑ- в среднекитайском языке. Сперва, исходя из тщательного анализа словаря рифм *Гуанъюнь* (广韵) и других текстов, авторы утверждают, что несмотря на то, что различие между звонкими аффрикатами и фрикативными в целом сохранялось в системе *Гуанъюнь*, в отдельных словах звонкие аффрикаты *dz*-/dʑ- уже совпали с фрикативными *z*-/ʑ- из-за диалектного влияния других текстов и/или в результате фонетических изменений. Далее утверждается, что инициали *z*- и ʑ- в среднекитайском языке находятся в отношении либо дополнительной дистрибуции, либо (в отдельных случаях) свободного варьирования; следовательно, для них можно предполагать единое развитие из исходной фонемы /z/. Основываясь на этих предположениях, авторы проводят детальный анализ фонетических серий китайской иероглифики и выявляют четыре различных источника происхождения фонемы /z/, а именно, древнекитайские инициали и начальные кластеры **r-j-*, **s-ɢ-*, **z-* и **s-d-*.

Ключевые слова: древнекитайская фонология, историческая реконструкция, звуковые изменения, переднеязычные аффрикаты.

Mᴀ Kun

Department of Chinese Language and Literature, Sun Yat-sen University; 9filippo@gmail.com

A study of phonological issues in the text variants
of *Xiaochu* and *Dachu* hexagrams, *Zhouyi*

This paper discusses some phonological issues with the *Xiaochu* (小畜) and *Dachu* (大畜) hexagrams by using text variants from three versions of the unearthed *Zhouyi* (周易). We point out that the reading of the character 輻 *fú* 'spoke', which rhymes with 目 *mù* 'eye' in the current version of *Xiaochu*, is the result of phonetic assimilation. We also propose that the phonetic 逐 *zhú* 'pursue' be assigned to the **-uk* group instead of **-iwk* in the Baxter-Sagart (2014) system.

Keywords: Old Chinese language; *Zhouyi* hexagrams; Chinese rhymes; phonological reconstruction.

The third line statement (爻辭) of the current version of the *Xiaochu* hexagram in *Zhouyi* (周易·小畜) has the following passage: "九三, 輿說輻, 夫妻反目"[1]. Commentators from Han to Wei periods proposed diverse interpretations for the word 輻 *fú*: both Wang Bi's *Zhouyi zhu* (王弼, 周易注) and Yu Fan's *Zhouyi zhu* (虞翻, 周易注) noted that 輻 *fú* represents the wheel spoke of a chariot, while Ma Rong (馬融) and Zheng Xuan (鄭玄) had another version which used the character 輹 *fù* instead. According to Ma and Zheng, 輹 *fù* is a component of the chariot which connects the axle with the compartment (車伏兔)[2]. The textual variation between 輻 *fú* and 輹 *fù* also triggers certain phonological issues which will be discussed in this paper.

During the Song and Ming dynasties, philologists usually followed Wang Bi's version. For instance, in his *Zhouyi benyi* (周易本義), Zhu Xi (朱熹) explained "輿說輻" as "spokes detach from the wheel"[3]. During the Qing dynasty, however, while using Old Chinese rhyming analysis, scholars became suspicious about the character 輻 *fú* in Wang Bi's version. In *Yi Yin, Yinxue Wushu* (音學五書·易音), Gu Yanwu (顧炎武) examined all of the rhyming passages in the *Zhouyi* and pointed out that 輻 *fú* does not rhyme regularly in the *Xiaochu* (小畜) line statements. He noted:

> The character 輻 rhymes with 側, 直, 億, 特, 食 in *Fa tan, Shi Jing* (詩 • 伐檀, 112.2A), with 載 and 意 in *Zhengyue* (正月, 192.10A), with 塞 and 息 from quotations of lost *Shi* (逸詩) in *Yunzi* (荀子); whereas 目 *mù* rhymes with 腹 and 復 in *The second year of duke Xuan, Zuozhuan* (左傳 • 宣二年), with 蹙 in *The sixteenth year of duke Cheng* (成公十六年). One cannot force 輻 to rhyme with 目, hence 輹 should be the correct version in the *Xiaochu* hexagram (The first volume of *Yi Yin, Yinxue Shishu*)[4].

[1] "The third NINE, undivided, suggests the idea of a carriage, the strap beneath which has been removed, or of a husband and wife looking on each other with averted eyes."

[2] This paper deals only with phonological issues; we will address the controversial statements in the commentaries elsewhere.

[3] 有輿說輻之象。

[4] "輻"字《詩•伐檀》與"側""直""億""特""食"韻，《正月》與"載""意"韻，《荀子》引逸《詩》與"塞""息"韻；"目"字《左•宣二年》與"腹""復"韻，《成十六年》與"蹙"韻。 不可強合也，作"輹"為是(顧炎武《音學十書·易音·卷一》)。

Gu believed that 目 *mù* in the *Xiaochu* line statement originally rhymed with 輹 *fù*, and that the character 輻 *fú* in the current version was due to text corruption. Gu's opinion was later adopted by many other scholars, such as Duan Yucai (段玉裁), Jiang Yougao (江有誥), Wang Niansun (王念孫), etc. Their detailed rhyming analyses are as follows[5]:

Table 1. Rhyming patterns of *Xiaochu*

Xiaochu(小畜) line statements				
	Gu	Duan	Jiang	Wang
初九：復自道A，何其咎A，吉。	○	○	○	○
九二：牽復B，吉。		○		○
九三：輿說輻[輹]B，夫妻反目B。	○		○	
六四：有孚，血去惕出，无咎。				
九五：有孚攣如，富以其鄰。				
上九：既雨C既處C，尚德載，婦貞厲。月幾望，君子征凶。	○	○		○

Duan, Jiang and Wang agree with Gu Yanwu that 輻 *fú* in the current version should be revised to 輹 *fù*. Jiang Yougao's analysis coincides with Gu in linking the rhyming words 輹 *fù* and 目 *mù*. Apart from that, Duan Yucai and Wang Niansun recognize 復 *fù* 'return' in the second line as a rhyme word along with 輹 *fù* 'wedge' and 目 *mù*. We follow Duan and Wang's opinion.

Qing scholars sometimes quoted commentaries from Ma Rong and Zheng Xuan, in which an alternate text variant 輿說輹 from *Xiaochu* was preserved. *Shuowen* also features the same quotation 輿脫輹, but it is hard to determine whether it is from *Xiaochu* (小畜) or *Dachu* (大畜). The rhyming contexts in the *Dachu* hexagram are as follows:

Table 2. Rhyming patterns of *Dachu*

Dachu (大畜) line statements				
	Gu	Duan	Jiang	Wang
初九：有厲，利己。				
九二：輿說輹A。				
九三：良馬逐A，利艱貞。曰閑輿衛，利有攸往。		○		○
六四：童牛之牿B，元吉。				
六五：豶豕之牙B，吉。		○		○
上九：何天之衢B，亨。				

Neither Gu Yanwu nor Jiang Yougao had identified the rhyming pattern in the second and third line statements. Both Duan Yucai and Wang Niansun pointed out that 輹 *fù* 'wedge' rhymes with 逐 *zhú* 'pursue', and that 牿 *gù* 'livestock' rhymes with 牙 *yá* 'tooth' and 衢 *qú* 'broad street'. Here we follow Duan and Wang's analyses.

The phrase 輿說輹 can further be compared to the *Dazhuang* (大壯) hexagram. The rhyming situations are as follows:

---

[5] Rhyme words are marked with underlining; A, B, C represent different rhyme groups.

Table 3. Rhyming patterns of *Dazhuang*.

Dazhuang (大壯) line statements				
	Gu	Duan	Jiang	Wang
初九：壯于趾，征凶，有孚。				
九二：貞吉。				
九三：小人用<u>壯</u>A，君子用<u>罔</u>A，貞厲。羝羊觸藩，羸其角。	○	○		○
九四：貞吉悔亡，藩決不羸，壯于大輿之輹。				
六五：喪羊于易，无悔。				
上六：羝羊觸藩，不能<u>退</u>B，不能<u>遂</u>B，无攸<u>利</u>B，艱則吉。	○	○		○

Jiang did not identify the rhyming patterns in *Dazhuang* (大壯). Duan believed that 壯 *zhuàng* rhymed with 罔 *wǎng* in the third line and that 退 *tuì* rhymed with 遂 *suì* and 利 *lì* in the sixth line statement; Gu and Wang agreed with Duan, except that they believed that 利 *lì* did not rhyme, due to certain inherent features of *Zhouyi* (周易): divinatory idioms such as "good fortune" (吉), "bad fortune" (凶), "disappearance of repentance" (悔亡) and so on should be excluded from the line statements, hence they do not rhyme under any circumstances. As for the character 輹 *fù*, it is commonly agreed that it does not rhyme in this hexagram.

There are now several unearthed pieces of evidence supporting the revision of 輹 *fù*. Three versions of unearthed *Zhouyi* are available to us: 1) Shangbo version (上博本), which is written on bamboo strips, dated to the late Warring States period (312–221 BC); 2) Fuyang version (阜陽本), also written on bamboo strips, dated to the late Western Han period (206 BC – 9 AD); 3) Boshu version (帛書本), which is written on silk, also dated to the late Western Han. Their preservation situation varies greatly: Boshu version contains all sixty-four hexagrams, including the hexagram's name (卦名), the hexagram statement (彖辭), and six line statements (爻辭), as well as parts of "Ten Wings" (十翼), such as *Xici* (繫辭) and even texts that were lost in history, including *Er san zi* (二三子), *Yi zhi yi* (易之義), *Yao* (要), *Miuhe* (繆和), *Zhaoli* (昭力); Shangbo version contains about half of the sixty-four hexagrams with no Ten Wings attached; Fuyang version is severely damaged, with hexagram fragments found only occasionally[6].

The characters 輹 *fù* or 輻 *fú* in the current versions indeed find their correlations in these unearthed versions of the *Zhouyi*[7]:

Boshu *Xiaochu* (少(小)蕶(畜), upper line 84)：

(1) 九二：堅(牽)復，吉。
九三：車說緮，夫妻反目。
'The second NINE, undivided, shows its subject, by the attraction (of the former line), returning (to the proper course). There will be good fortune.[8]
The third NINE, undivided, suggests the idea of a carriage, the <strap> (wedge) beneath which has been removed, or of a husband and wife looking on each other with averted eyes.'

Here 緮 *fù* corresponds to 輻 *fú* in the current version; in Shangbo and Fuyang *Zhouyi*, however, this hexagram has been lost. The character 緮 *fù* has the same phonetic component as 輹 *fù*, which fits well with Qing scholars' prediction. According to the principle of "*tong sheng*

---

[6] For details, see Han Ziqiang (韓自強) 2004, Ma Chengyuan (馬承源) 2003, and Qiu Xigui (裴錫圭) 2014.

[7] For now, we only deal with text variants which are associated with rhyming activities.

[8] This paper uses James Legge (2013)'s English translation of *Zhouyi*, with several modifications.

bi tong bu" (同聲必同部), since 輹 *fù* and 縷 *fù* share the same phonetic they must have had an identical or close pronunciation in the Old Chinese period.

Shangbo *Dachu* (上博本大壴(畜), third volume, strip 20) has:

(2) 九二：車(敓)說复。
九晶(三)：良馬由(逐)，利董(艱)貞。曰班車衛，利有攸往。
"The second NINE, undivided, shows a carriage with the strap under it removed.
The third NINE, undivided, shows its subject urging his way with good horses."

Whereas the Boshu version of *Dachu* (泰 (大) 蓄 (畜), line 11) has:

(3) 九二：車說縷。
九三：良馬逐，利根(艱)貞。曰闌(閑)車衛，利有攸往。
(refer to the previous translation)

Here 复 *fù* and 縷 *fù* correspond to 輹 *fù* in the current version. Their pronunciation in the Old Chinese period is also close since they share the same phonetic component.

Boshu *Dazhuang* (泰 (大) 壯, line 33) has:

(4) 九四：貞吉悔亡，藩(藩)块不羸，壯于大車之縷。
"The fourth NINE, undivided, shows (a case in which) firm correctness leads to good fortune, and occasion for repentance disappears. (We see) the fence opened without the horns being entangled. The strength is like that in the wheel-spokes of a large wagon."

Here 縷 *fù* corresponds to 輹 *fù* in the current version. This hexagram has been lost in both Shangbo and Fuyang *Zhouyi*.

Connecting the rhyming words above, we have the following table:

Table 4. Rhyming behaviors of 輹 *fù* in various versions of *Zhouyi*.

	Rhyme group	OC	MC	Notes
復 *fù*	覺	*m-puk	*bjuwk*	
目 *mù*	覺	*C.muk	*mjuwk*	
輹 *fù*	覺	*m-puk	*bjuwk*	
縷 *fù*	覺	——	——	Boshu *Xiaochu* and *Dachu*
复 *fù*	覺	*m-puk	*bjuwk*	Shangbo *Dachu*
輻 *fú*	職	*pək	*pjuwk*	The current version of *Xiaochu*
逐 *zhú*	覺	*lriwk	*drjuwk*	The current version and Boshu *Dachu*
由 *yóu*	幽	*lu	*yuw*	Shangbo *Dachu*

There are two irregular rhyming behaviors in the table above, all of which are related to the Middle Chinese rhyme *-juwk* (屋韻三等). According to the reconstruction of Baxter, Sagart (2014) as well as other "six vowel" systems, *-juwk* has several Old Chinese origins, and they merged mainly under the influence of initial consonants (see below).

The first issue is that 輻 *fú* (*pək) in the current version of *Xiaochu* (小畜) should not rhyme with characters from group *-uk (覺部) or group *-u (幽部).

Boshu version   縷 *buk > *bjuwk > fù* 'wedge, a borrowed form of 輹'
         目 *muk > *mjuwk > mù* 'eye'
Current version   輻 *pək > *pjuwk > fú* 'spoke'

緮 *fù* rhymes with 目 *mù* in the Boshu version, and they have the same final in both Old Chinese and Middle Chinese. This situation is parallel with *Dachu* (大畜) in Shangbo and the current version (see table 4). Whereas 輻 *fú* and 目 *mù* in the current version do rhyme together in Middle Chinese but have different main vowels in Old Chinese, which indicates that *-ək* had already assimilated into *-juwk* at the end of the Western Han period. The *-ə* (之) group and the *-əŋ* (蒸) group have a paralleled development under the same condition, for instance:

(5) 有 *[ɢ]ʷəʔ > hjuwX > yǒu* 'have, exist'
    弓 *kʷəŋ > kjuwng > gōng* 'bow'

Such assimilation could be blocked by pharyngealization or medial *-r-, for instance:

(6) 怪 *kʷˤrə-s > kweajH > guài* 'strange'
    洈 *[ɢ]ʷrəʔ > hwijX > wěi* 'name of a river'

The rhyming behavior of 輻 *fú* in the current version of *Xiaochu* (小畜) can be explained by the above phonological development. Namely, *-ək* was assimilated into *-juwk* under the influence of a non-pharyngealized labial initial. After this happened, 輻 *fú* (MC *pjuwk*) and 目 *mù* (MC *mjuwk*) could still rhyme together, since they have the same Middle Chinese final, which led to the change of rhyme word from 緮 *fù* to 輻 *fú* at that time. It may be therefore concluded that the rhyme word 輻 *fú* in the current version of *Zhouyi* reflects a phonological phenomenon of Early Middle Chinese. In the rhyming passage of *Minming, Jiutan* by Liu Xiang (劉向, 九嘆·愍命) "卻騏驥以轉運兮, 騰驢贏以馳逐; 蔡女黜而出帷兮, 戎婦入而綵繡服", 逐 *zhú* rhymes with 服 *fú*, which also reflects assimilation *-ək > -juwk* in late Western Han times:

(7) 逐 *lriwk > drjuwk > zhú* 'pursue'
    服 *[b]ək > bjuwk > fú* 'cloth'

The second issue is the assignment of *-juwk* (屋韻三等) characters after splitting up the traditional Jué (覺) group. This rhyme group includes Middle Chinese *-owk* (divison-I), *-aewk* (divison-Ⅱ), *-juwk* (divison-Ⅲ) and *-ek* (divison-IV). According to the "front-vowel hypothesis" (Baxter 1992), two different main vowels must be reconstructed with front and non-front contrasts, namely, *-iwk* and *-uk*. Middle Chinese words in the *-juwk* rhyme group should be assigned to *-iwk* or *-uk* according to their rhyming or *xiesheng* behaviors. For instance:

(8) *C(r)uk > Cjuwk*
    畜 *qʰuk > xjuwk > xù* 'nourish'; it rhymes with *-uk* category words in Ode 188.2A, 202.4A; 蓄 *xù*, 稸 *xù*, and 慉 *xù*, which share the same phonetic 畜 *xù* all belong to the *-uk* category.

(9) *C(r)iwk > Cjuwk*
    穆 *mriwk > mjuwk > mù* 'harmonious'; it rhymes with *-iwk* category words in Ode 282.1B.

In table 4, according to Baxter (1992) and Baxter-Sagart (2014), 緮 *fù*, 复 *fù*, 緮 *fù*, 復 *fù*, 目 *mù* and 由 *yóu* all belong to the *-uk* category, while 逐 *zhú* belongs to the *-iwk* category. Apparently, 逐 *zhú* is abnormal in this table. In both current and Boshu versions of *Zhouyi*, this abnormality repeats itself:

(10) 初九：悔亡，喪馬勿逐，自復。見惡人无咎。(The current version of *Kuí* (睽) hexagram) "The first NINE, undivided, shows that (to its subject) occasion for repentance will disappear. He has lost his horses, but let him not seek for them — they will return of themselves. Should he meet with bad men, he will not err (in communicating with them)."

It is well agreed among Qing scholars that 逐 *zhú* rhymes with 復 *fù*. However, 逐 *zhú* is replaced by 由 and 遂 <逐> respectively in Shangbo and Boshu version, which is similar to the above situation in *Dachu* (大畜):

(11)初九：悉(悔)亡=(亡，亡)馬，勿由(逐)，自邊(復)。見亞(惡)人无咎。　　(Shangbo　*Kuí* (楑(睽)，　strip 32, refer to the previous translation).

(12)初九：悉(悔)亡=(亡。亡)馬勿遂<逐>，自復。見亞(惡)人，无咎。(Boshu　*Guaī*　乖(睽), upper line 75, refer to the previous translation.)

The character 逐 *zhú* does not appear in the *Shijing*. However, there does exist a rhyme word 蓫 *zhú* with 逐 *zhú* as phonetic:

(13)我行其野，言采其蓫。
昏姻之故，言就爾宿。
爾不我畜，言歸思復。(小雅·我行其野 二章, 188.2A)
"I travelled through the open country, luxuriant are the Rumex; with a view to (our) marriage, I went and lodged with you; but you did not cherish me, I go back and return (home)."[9]

蓫 *zhú* rhymes with 宿 *sù*, 畜 *xù*, and 復 *fù* in this stanza (188.2A); all rhyme characters belong to the traditional Jué (覺) group. Baxter-Sagart (2014)'s reconstructions are as follows:

Table 5. Rhyming situation of 蓫 in the *Shijing* (188.2A).

	Traditional group	OC	MC
蓫 *zhú*[10]	Jué (覺)	*hlrjiwk (*lriwk)	*drjuwk*
宿 *sù*	Jué (覺)	*s[u]k	*sjuwk*
畜 *xù*	Jué (覺)	*qʰuk-s	*xjuwX*
復 *fù*	Jué (覺)	*m-puk	*bjuwk*

According to their reconstruction, 蓫 *zhú* is abnormal in the above table, parallel with the situation with 逐 *zhú* in Table 4, where *-uk and *-iwk categories were also mixed up. It is thus clear that both 逐 *zhú* and 蓫 *zhú* are rhyming with *-uk words in the *Shijing* and *Zhouyi*. Hence, we propose that the phonetic 逐 *zhú* should be assigned to the *-uk group in the Baxter-Sagart (2014) system:

(14)逐 *lriwk → *lruk > *drjuwk* > *zhú* 'pursue'
蓫 *hlrjiwk (*l̥riwk) → *l̥ruk > *drjuwk* > *zhú* 'sheep's-foot (name of grass)'

However, there is a case where 逐 *zhú* in the current version of *Zhouyi* became entangled with the *-iwk category:

(15)六四:顛頤，吉，虎視眈眈，其欲逐逐，无咎。(The current version of *Yí* (頤) hexagram)
"The fourth SIX, divided, shows one looking downwards for (the power to) nourish. There will be good fortune. Looking with a tiger's downward unwavering glare, and with his desire that impels him to spring after spring, he will fall into no error."

---

[9] This paper uses Karlgren (1974)'s English translation of *Shijing*.
[10] Baxter and Sagart (2014) do not include the item 蓫 *zhú*; Baxter (1992) reconstructs it as *hlrjiwk, which is equivalent to *lriwk in their new system.

We adopt the conclusion of Qing scholars' analyses that the *Yí* (頤) hexagram does not rhyme, apart from its first line statement[11]. The character 逐 *zhú* is replaced by 攸 *yōu* and 笛 *dí* in Shangbo and Boshu versions respectively:

(16) 六四：顛頤，吉。虎見(視)≣=(眈眈)，其猶(欲)攸=(逐逐)，亡(无)咎。(Shangbo *Yí* (頤), strip 25, refer to the previous translation.)

(17) 六四:顛頤，吉。虎視沇【=】(沇【沇】—眈【眈】)，其容(欲)笛=(笛笛—逐逐)，无咎。(Boshu *Yí* (頤), upper line 19, refer to the previous translation.)

We can conveniently join all the text variants of 逐 *zhú* in the following table:

Table 6. Text variants of 逐 *zhú* in the *Yí* (頤) hexagram.

	Traditional group	Old Chinese	Middle Chinese
逐 *zhú*	Jué (覺)	*lriwk	*drjuwk*
攸 *yōu*	Yōu (幽)	*liw	*yuw*
笛 *dí*	Jué (覺)	*liwk[12]	*dek*

It seems that the phonological status of the three versions of the *Yí* (頤) hexagram is somewhat identical: all the variants have a front vowel *-i. Note that this passage does not intend to rhyme, which might lower the standards for correspondences. Moreover, the characters in question actually form binomes, such as 逐逐, 攸攸 and 笛笛, which do not always faithfully reflect the phonological phenomena[13]. For instance, in the received version of *Laozi* (also called Wangbi version (王弼本老子)), we have 繩繩不可名; however, 繩繩 is replaced by the loan binome 蠅蠅 in the Xiang'er version (想爾注本), by 台微台微 in the Hanjian version (漢簡本) and by 尋尋 in the Boshu version (帛書本). Although all these binomes have similar syllabic structures as well as initials and main vowels, they do have quite different onsets or final consonants:

(18) 繩繩 *Cə.ləŋ Cə.ləŋ ~ *mə.ləŋ mə.ləŋ > *shéngshéng* "abundant"
蠅蠅 *m.rəŋ m.rəŋ > *yíngyíng* "abundant"
台微台微 *ləməj ləməj → *ləm ləm > *yíwēi yíwēi* "abundant"
尋尋 *sə-ləm sə-ləm > *xúnxún* "abundant"

There are nine characters in the *xiesheng* series of 逐 *zhú*. According to the *Guangyun shengxi* (廣韻聲系), they could be divided into four groups: 1) *drjuwk* (直六切): 逐鯈鯈蓫𧾷篴; 2) *trhjuwk* (丑六切): 蓫; 3) *xjuwk* (許竹切): 蓫; 4) *dek* (徒歷切): 篴. In *Grammata Serica*, Karlgren

---

[11] "初九：舍爾靈龜，觀我朵頤。""The first NINE, undivided, (seems to be thus addressed), 'You leave your efficacious tortoise, and look at me till your lower jaw hangs down.' There will be evil." Gu Yanwu (顧炎武), Duan Yucai (段玉裁) and Wang Niansun propose that 龜 *guī* rhymes with 頤 *yí*, while Jiang Yougao (江有誥) believes there is no rhyming. Here we follow Gu, Duan and Wang's analyses.

[12] Baxter and Sagart (2014) have not reconstructed the item 笛 *dí*; however, the character 迪 *dí* with the same Middle Chinese reading as笛*dí* is reconstructed as *liwk in Baxter (1992). We provisionally assume that 笛*dí* and 迪 *dí* had the same pronunciation in Old Chinese.

[13] 迪 *dí* rhymes with 復 *m-puk > *fù*, 毒 *m-[d]ˤuk > *dú* in 257.11A (大雅·桑柔); 迪 *dí*, 復 *fù* and 毒 *dú* are assigned to the traditional Jué (覺) group. However, after the split of Jué (覺) by Baxter and Sagart (2014), 迪 *dí* is assumed to have had a different main vowel from 復 *fù* and 毒 *dú*. A probable explanation is that 迪 *dí* might come from a late version of the *Shijing*.

recorded three characters (GSR 1022): 逐 *drjuwk > zhú*, 篴 *dek > zhú*, and 蓫 *trhjuwk > zhú*. How-ever, according to the *Guangyun* (廣韻), 篴 *zhú* is equivalent to 笛 *dí* 'Chinese flute', which was invented in the middle of the Western Han period (around 157–87 BC)[14]. Hence, the pronun-ciation of division-IV word 篴 *dek > zhú* might be of late origin and should not be used as evi-dence to reconstruct Old Chinese front vowel *i.

References

Baxter, William. 1992. *A Handbook of Old Chinese Phonology*. Mouton de Gruyter.

Baxter, William, Laurent Sagart. 2014. *Old Chinese: A New Reconstruction.* Oxford University Press.

Duan, Yucai (段玉裁). 1988. *Liushu Yinyun Biao* 六書音均表 [*Phonological Tables of the Liushu*], in *Shuowen Jiezi Zhu* 说文解字注 [*Annotations of Analytical Dictionary of Characters*]. Shanghai: Shanghai guji chubanshe 上海:上海古籍出版社.

Gu, Yanwu (顧炎武). 2012. *Ri Zhi Lu* 日知錄 [*Records of Daily Knowledge*]. Shanghai: Shanghai guji chubanshe 上海:上海古籍出版社.

Han, Ziqiang (韓自強). 2004. *Fuyang Hanjian Zhouyi Yanjiu* 阜陽漢簡周易研究 [*Study of the Fuyang Zhouyi*]. Shanghai: Shanghai guji chubanshe 上海: 上海古籍出版社.

Jiang, Yougao (江有誥). 1993. *Yinxue shi shu* 音學十書 [*Ten Books on Old Chinese Phonology*]. Beijing: Zhonghua shuju 北京: 中華書局.

Karlgren, Bernhard. 1974. *Book of Odes (Chinese text, transcription and translation).* The Museum of Far Eastern Antiquities.

Legge, James. 2013. *I Ching (Book of Changes).* Kessinger Publishing.

Ma, Chengyuan (馬承源). 2003. *Shanghai Bowuguan Cang Zhan'guo Chu Zhushu* 上海博物館藏戰國楚竹書(三). [*Shanghai Museum Collection of Chu Bamboo Books from the Warring States Period* (III)]. Shanghai: Shanghai guji chubanshe 上海:上海古籍出版社.

Qiu, Xigui (裘錫圭). 2014. *Mawangdui Hanmu Jianbo Jicheng* 馬王堆漢墓簡帛集成[*Integration of Bamboo Slip and Silk Documents in Mawangdui Han Tomb*]. Beijing: Zhonghua shuju.

Shen, Jianshi (沈兼士). 1985. *Guangyun shengxi* 廣韻聲系 [*The Xiesheng system of Guangyun*]. Beijing: Zhonghua shuju.

Ма Кунь. Исследование фонологических проблем в вариантах текста гексаграмм *сяочу* и *дачу* («Книга Перемен»)

В статье обсуждаются вопросы фонологической реконструкции ряда древнекитайских слов, встречающихся в гексаграммах *Сяочу* (小畜) и *Дачу* (大畜) «Книги перемен» с при-влечением новых данных, полученных по трем альтернативным версиям памятника, найденным в результате археологических раскопок. Обосновывается гипотеза, что не-обычная рифмовка между знаками 輻 *fú* 'колесничная спица' и 目 *mù* 'глаз', наблюдае-мая в канонической версии *Сяочу*, вызвана фонетической ассимиляцией. Также выска-зано предположение, что фонетик 逐 *zhú* 'преследовать' необходимо отнести к группе рифм *-uk вместо *-iwk, как это предложено в системе Бэкстера и Сагара (2014).

Ключевые слова: древнекитайский язык; гексаграммы Ицзина; китайская система рифм; фонологическая реконструкция.

---

[14] 《廣韻·錫韻》：“笛，樂器。《風俗通》云武帝時丘仲所作也。晉協律中郎列和善吹笛也。出《周禮》。”

Johann-Mattis L*IST*[†], *Nathan W.* H*ILL*[‡], *Christopher J.* F*OSTER*[§]

[†] Department of Linguistic and Cultural Evolution, Max Planck Institute for the Science of Human History, Jena, list@shh.mpg.de

[‡] School of African and Oriental Sciences, University of London, London, nh36@soas.ac.uk

[§] Pembroke College, University of Oxford, Oxford, christopher.foster@pmb.ox.ac.uk

Towards a standardized annotation of rhyme judgments in Chinese historical phonology (and beyond)[1]

Although rhyme analysis plays a crucial role in the reconstruction of Old Chinese phonology, the field has not yet developed a standardized annotation framework for rhyme judgments applied to Ancient Chinese texts. Building on initial attempts to standardize cross-linguistic data for the purpose of historical and typological language comparison (as part of the Cross-Linguistic Data Formats initiative), we present a proposal for consistent and transparent rhyme annotation. This proposal allows scholars to annotate the rhymes they identify in historical texts in such a way that the judgments can be analyzed with computational tools as well as conveniently inspected by scholars. Our framework is accompanied by software tools and exemplary datasets, which were annotated by various scholars, and reflect not only Chinese, but also contemporary poetry in different languages. In the paper, we present the framework and also point to caveats and current insufficiencies in annotation. In doing so, we hope to inspire more scholars working on Old Chinese reconstruction to share their judgments, allowing others working in the field to improve, revise, and analyze them.

Keywords: Old Chinese language, Chinese rhymes, data annotation, cross-linguistic data formats.

1. Introduction

Rhyme analysis plays a crucial role for the reconstruction of Old Chinese phonology, but the field has not yet developed a standardized framework for annotating rhyme judgments. In this paper, we want to present a new annotation framework for rhyme judgments, which builds on the general idea of increasing the comparability of data in historical linguistics and language typology, and has the goal of being not only applicable to Chinese texts, but to the poetic traditions of any language that uses rhyme as a device.

In the following, we introduce our framework in detail, by first pointing to the importance of rhyme analysis for Chinese historical phonology (1.1), discussing the typical practice of rhyme annotation in Chinese linguistics (1.2), and presenting some general thoughts on the importance of annotation in philology and linguistics (1.3). We then present our framework in detail, by introducing the Cross-Linguistic Data Formats initiative (2.1), presenting the main ideas for rhyme annotation (2.2), and providing several examples of rhyme annotation in practice (2.3). We conclude by articulating the hope that our example can inspire scholars in our field to improve the transparency of our research by providing data underlying analyses in generally comparable formats.

1.1. Rhyme analysis in Chinese historical phonology

Due to phonetic change, the rhymes of ancient Chinese texts often cease to rhyme in more modern pronunciations. Already in the sixth century of our era Shěn Zhòng 沈重 noticed failures of expected rhymes in the Shījīng 詩經; he suggested adjusting one's pronunciation to make the rhymes read smoothly. The Míng 明 dynasty scholar Chén Dì 陳弟 (1541–1617) explained that sound change had altered the original pronunciation of at least some words, and that these words normally had a single pronunciation in the mouths of the ancients (Baxter 1992, 154). The scholar Gù Yánwǔ 顧炎武 (1613–1682) was first to undertake a reconstruction of the rime categories of Old Chinese; he elaborated ten rime categories (*yùnbù* 韻部) in the *Shījīng*, which split into the more elaborate categories of Middle Chinese rimes (Baxter 1992: 155–57). Subsequent scholars distinguished categories that the *Shījīng* keeps apart in its rhyming practices, which Gù Yánwǔ had failed to notice. The categories recognized by scholars working within the Chinese philological tradition steadily rose over time to 22 (Baxter 1992: 157–71). In the late 20th century, armed with the six vowel hypothesis of Old Chinese, and motivated by the internal reconstruction of Middle Chinese, the three scholars Zhèngzhāng Shàngfāng 鄭張尚芳 (Zhengzhang 2000), Sergei Starostin (Starostin 1989), and William Baxter (Baxter 1992) independently recognized many more rime categories. For example, Schuessler (Schuessler 2009), who also operates in the six-vowel tradition, puts the total number of Old Chinese rime categories at 38 and we count 45 in Baxter & Sagart's most recent Old Chinese reconstruction (Baxter and Sagart 2014).

The rime category of an Old Chinese word is only directly knowable if that word happens to occur as a rhyme word in the *Shījīng*. Except for in those few cases where the Middle Chinese pronunciation of a word may, according to one's overall theory, develop only from a single Old Chinese rime category, in order to speak of the rime category of words that do not appear as rhyme words in the *Shījīng*, one must turn to the phonetic information inherent in the Chinese writing system.

1.2. Rhyme annotation in Chinese historical phonology

The ways in which scholars share their respective rhyme judgments in the literature is very diverse and makes a formal comparison of different rhyme analyses difficult. The problem here lies only to some degree in missing digital versions of important contributions, which would be merely a problem for pure computational approaches. A more significant problem is that many authors report their rhyme judgments in a form that is insufficiently explicit to infer the individual judgments made on individual poems and stanzas. Apart from scholars who presented only the *results* of their analyses, without providing the evidence (Zhèngzhāng 2003; Pān 2000), we also often find analyses that are extremely difficult to inspect, due to the way they present their judgments. In this sense, only a small amount of rhyme analyses is truly *explicit*.

An example for the problem of insufficient explicitness in the way rhyme judgments are reported is the otherwise excellent study of Old Chinese phonology by Sergei Starostin (Starostin 1989: 458–674): Instead of providing a full version of the *Shījīng* that he used for his reconstruction, Starostin's data starts from rhyme groups and then lists all rhyme words per stanza that he judges to reflect this rhyme group. For example, for the rhyme group *zhī* 之 *-ə*, we find the rhyme words *ćə:* 哉, *gə* 其, *tə* 之, and *sə* 思 (p. 448), which directly corresponds to the classical analysis of stanza 2 in Ode 109, for which Wáng Lì gives the following rhyme judgments (Wáng 1980):

彼人是哉(tzə)! 子曰何其(giə)!
心之憂矣,其誰知之(ʈjiə)?
其誰知之(ʈjiə)?
蓋亦勿思(siə)! (*Shījīng*, 109.2)

Starostin's analysis is more explicit than other attested analyses, in that it makes a formal representation, in which each rhyme word in the text of the *Shījīng* is marked as such along with the proposed reconstruction. Nonetheless, any attempt to digitize or reverse-engineer individual judgments from the data in the book would require a full digitization and numerous hours of identifying each character's occurrence in the original source. In contrast, Wáng Lì's format is very transparent, insofar as it marks exactly where each rhyme word occurs in context.

Explicit analyses of *Shījīng* rhymes — apart from Wáng Lì (Wáng 1980) — also include Karlgren (Karlgren 1950), Baxter (Baxter 1992: 583–743), and Wáng Xiǎ'an (Wáng 2011). In all these analyses, the original text of the *Shījīng* that was taken as the basis for the rhyme judgments is accompanied by a note indicating which lines in each stanza rhyme and how the authors reconstruct the rhyme words in those lines. Here again, however, we can find differences in the degree of explicitness by which authors report their actual rhyme judgments. While Wáng Lì, for example, adopts an annotation that marks rhymes that recur across stanzas, Baxter only shows rhymes inside each stanza. Furthermore, it is rare for any of the authors to point to instances of internal rhyme, probably also due to the fact that their general rhyme annotation schema is built in such a way that it describes the relation between lines in the *Shījīng* (as opposed to the relation between words inside a stanza or a poem).

Apart from the obvious problem of explicitly showing what scholars think should rhyme in a given analysis of the *Book of Odes* or other rhyme collections, we also face many less obvious problems when dealing with rhyme judgments. In many cases, for example, scholars may themselves be uncertain if a given instance reflects an actual rhyme or not. So far, however, we have not found any example in the literature where scholars would try to express their uncertainty in any form. A further problem lies in the inclusion of supporting data that would allow to contrast a given scholar's decisions with external evidence. While Wáng Lì's rhyme judgments, for example, only provide one reading in his Old Chinese reconstruction per rhyme word, Baxter's 1992 version also provides the Middle Chinese readings which are similarly important for the evaluation of his judgments, and while both Baxter and Wáng only provide one possible reading per rhyme word, there are quite a few instances in the *Book of Odes* where several readings would be possible.

1.3 Annotation in linguistics and philology

Annotation is crucial for scientific research involving language and texts. The major idea of annotation is to provide some kind of *added value* for a given resource (Milà–Garcia 2018), i.e., some information that could not — or only with great efforts — be extracted from the original resource without resorting to intensive search or complex computational algorithms. What value we add when annotating a resource depends on our research question. In *interlinear-glossed text* (MPI EVA 2008), for example, linguists try to provide some kind of a meta-language for disentangling grammatical particles from content words, in order to help other linguists to understand how the general meaning of a phrase or sentence is constructed. In *morphological annotation,* as introduced by Hill and List (Hill and List 2017), the same idea is applied to multi-morphemic forms in cross-linguistic word lists.

One can roughly distinguish two basic types of annotation: *inline* and *stand-off* annotation (Eckart 2012). While inline annotation manipulates the original data directly, for example, by adding tags, stand-off annotation only references the original data, without directly modifying it. Most annotation frameworks, however, typically use a mixture between the two types, although it is clear that stand-off annotation has the advantage of allowing for far more flexibility, especially if adding multiple layers of annotation to a given resource.

As an example illustrating the difference between the two annotation styles, consider the rhyme annotation employed by Baxter (Baxter 1992) as compared to the one by Wáng Lì (Wáng 1980), shown above, for poem 109 (second part of stanza 2 in the *Book of Odes*). While Wáng Lì provides the rhyme judgements inline, Baxter (p. 625) basically uses a stand-off annotation by listing all relevant data in tabular form:

Table 1. Rhyme annotation in Baxter (1992), Ode 109, Stanza 2

Character	Pīnyīn	MCH	OCH	Rhyme
哉	zāi	tsoj	*tsɨ	B
其	jī	ki	*k(r)jɨ	B
之	zhī	tsyi	*tjɨ	B
之	zhī	tsyi	*tjɨ	B
思	sī	si	*sjɨ	B

Both types of annotation have advantages and disadvantages. Wáng presents the whole text, so we know exactly which words he judges to rhyme and where he locates the relevant rhyme words. Since Baxter does not provide an index to the words in the original *Shījīng* text, we cannot know exactly where the rhyme words occur in the lines (it is, for example, possible that a character is repeated throughout the same line), and we can also not see the poem as a whole, along with its structure of rhyming and non-rhyming lines. The advantage of Baxter's system, however, is that it allows him to list more data related to each word, including the Pīnyīn transliteration, Middle Chinese and Old Chinese readings, and even his assessments as to which lines rhyme with each other. Thus, while Baxter loses explicitness with respect to the underlying *Shījīng* text, Wáng loses the flexibility of annotation. Ideally, an advanced annotation framework for rhyme judgments should allow for the advantages of both approaches.

2. Towards a standard of rhyme annotation in Chinese historical phonology

As we have seen in the foregoing discussion, the annotation of rhymes — be it in Chinese historical phonology or in general — is not trivial, in particular since there are considerable desiderata for common rhyme annotation frameworks. Thus, we would first like to be able to annotate large collections of poems, like the *Shījīng*, where we retain the original text, but could also indicate character readings, as proposed by different authors in the literature. We may also want to indicate details of rhyming, for example, pointing to impure rhymes or indicating internal rhymes, which we know occasionally occur in the *Shījīng*.

In order to advance our understanding of rhyming in China, we will in the long run require a more comparative, typolological perspective that could tell us to which degree the rhyme practice that we observe in ancient Chinese texts is peculiar or expected. For this reason, it would also be desirable if our rhyme annotation framework could be used for all kinds of rhyming poetry, stemming from different genres, languages, and cultures. Judging from

our knowledge of different genres, both in the history of Chinese poetry, but also of poetry world-wide, we may occasionally want to add a lot more information, for example on meter, syllables, word boundaries, or tonal patterns.

While all these aspects need to be taken into consideration when proposing a first format for rhyme annotation, it is also important to be pragmatic to some degree, since we know from experience that very complex format prescriptions will intimidate users rather than encouraging them to take part. Finding the right balance between pragmatism and perfectionism is thus crucial for our endeavor.

2.1. The Cross-Linguistic Data Formats initiative

The Cross-Linguistic Data Formats initiative (https://cldf.clld.org) is an attempt to standardize different types of data which are frequently used in the context of historical linguistics and linguistic typology (Forkel et al. 2018). While the current version mainly focuses on standardized formats for wordlists and structural data, the specifications are intended to be expandable in future versions, and draft proposals for dictionaries and parallel texts are underway.

The common procedure of adding new format specifications to the CLDF initiative is by testing the ideas on sufficiently large amounts of data first, before an official discussion of whether and how to integrate a new data format into the CLDF framework should be undertaken. The attempts described here are a first effort at presenting our basic ideas to a broader public, in the hope that after sufficient testing and discussion we can include rhyme annotation frameworks in future versions of the CLDF. Although rhyme analyses of the depth as we propose here are — at least to our knowledge — a rather new enterprise, we are confident that our format proposals are sufficiently useful for inclusion in the CLDF initiative, because they would allow focus on new, fascinating, and largely unexplored cross-linguistic data.

2.2. Main ideas for rhyme annotation

The main ideas for our proposed format of rhyme annotation follow largely the ideas that drove the development of the CLDF format, and although our current proposal has to be seen as independent of CLDF, we hope that the ideas can later be included into a new release of CLDF that would include poems and rhyme annotations as an additional component. The major criteria for the choice of our format proposal follow to a large degree the — among programmers well-known — "Zen of Python", which claims that "Simple things should be simple, complex things should be possible".

Our basic ideas thus require: (1) simplicity, (2) exhaustiveness, (3) flexibility. Simplicity means that people should be able to apply our format prescriptions with a minimal amount of work, using standard off-the-shelf tools, like text or spreadsheet editors, rather than complex new tools that would have to be created specifically for rhyme analysis. Exhaustiveness means that we wish to be able to reflect all knowledge that can be formalized in a given rhyme analysis. While we would always allow adding ad-hoc information in note-fields, we want to offer a high degree of granularity in annotations, allowing, for example, the inclusion of phonetic transcriptions and phonetic alignments (List 2014). Flexibility allows for a quick extension of the data when needed, using mechanisms already offered by the framework.

In order to achieve all these goals, we draw largely from our experience with the enhanced annotation and computer-assisted manipulation of *wordlists* in historical linguistics (Hill and List 2017) and their subsequent inclusion into the CLDF specifications.

2.2.1. Representing rhyme collections in spreadsheets

Following the basic idea of CLDF to represent most of the data in the form of spreadsheets, we propose a very straightforward way to represent rhyme annotations in spreadsheet format. While CLDF proper would require that the data is delivered in form of comma-separated or tab-separated value (CSV or TSV), the data can be easily annotated with widely used spreadsheet editors, such as Excel or LibreOffice. The key component of a spreadsheet is a header line that indicates the values that we find in the sheet, and the rows, that add values for each column as it is described by the header.

Based on the discussions of the desiderata and past experiments which proved the particular insufficiency of certain annotation forms, our core annotation of a poem or a poem collection now contains the following main components:

- **ID:** the identifier, which is a numerical ID.
- **POEM:** a name for the given poem.
- **STANZA:** the stanza of the poem (usually a numeric value, preceded by the name of the poem).
- **LINE_IN_SOURCE:** the line of the poem as we find it in the source from which the data is taken (especially containing original punctuation etc.).
- **LINE:** a double-segmented version of the line, in which words are separated with help of + as a separator, and spaces can be used to represent segments of phonetic values (similar to the format adopted by the LingPy software package to represent phonetic sequences and alignments).
- **LINE_ORDER:** A numerical value that provides the order of the lines of a poem in a given stanza.
- **RHYMEIDS:** A list of numerical identifiers, indicating which words in a the LINE rhyme by assigning the same ID to different words, using 0 to indicate that a given word does not rhyme.
- **ALIGNMENT:** A double-segmented version of the line that can, however, store aligned content, differing from the data in LINE, as well. This data comes in handy when trying to check questions of phonetic similarity of rhyme words, or of vowel purity, which would greatly facilitate automatic analyses as the one presented in List et al. (2017).

With these eight columns provided, poems can be annotated in a very straightforward way, regardless of the language in which they were written. One can, of course, add many more columns, depending on specific characteristics of the datasets, but for the general rhyme annotation, we think that these fields will be sufficient for most of the cases; it substantially exceeds rhyme annotation frameworks that have been proposed so far in terms of detail.

As an example, consider (again) ode 109, stanza 2, in the rhyme judgments of Wáng Lì (Wáng 1980), shown in the table below. Note that the entry for LINE_IN_SOURCE is not shown in Table 2, as the excess length of each row would run beyond the width of this paper, thereby disorienting readers; however it is still a crucial component for this annotation standard, and readers can see the full analysis by Wáng Lì in the supplementary data accompanying this paper.

While this representation may look complicated at first, it offers a degree of explicitness we have not found in any of the transparent rhyme annotations proposed in the past. On the one hand, we manage to avoid a complex inline annotation, while on the other hand we can express in a very detailed way which words (or characters) in the stanza rhyme, and how they should be pronounced.

Table 2. Rhyme annotation format (excerpt) with alignments and identifiers for rhyme words.

ID	POEM	ST.	LO	LINE	ALIGNMENT	RHYMEIDS
1733	園有桃	109.2	1	園 + 有 + 棘	園 + 有 + kiək	0 0 467
1734	園有桃	109.2	2	其 + 實 + 之 + 食	其 + 實 + 之 + djiək	0 0 0 467
1735	園有桃	109.2	3	心 + 之 + 憂 + 矣	心 + 之 + 憂 + 矣	0 0 0 0
1736	園有桃	109.2	4	聊 + 以 + 行 + 國	聊 + 以 + 行 + kuək	0 0 0 467
1737	園有桃	109.2	5	不 + 我 + 知 + 者	不 + 我 + 知 + 者	0 0 0 0
1738	園有桃	109.2	6	謂 + 我 + 士 + 也 + 罔 + 極	謂 + 我 + 士 + 也 + 罔 + qiək	0 0 0 0 0 467
1739	園有桃	109.2	7	彼 + 人 + 是 + 哉	彼 + 人 + 是 + tzə	0 0 0 468
1740	園有桃	109.2	8	子 + 曰 + 何 + 其	子 + 曰 + 何 + giə	0 0 0 468
1741	園有桃	109.2	10	其 + 誰 + 知 + 之	其 + 誰 + 知 + tjiə	0 0 0 468
1742	園有桃	109.2	10	其 + 誰 + 知 + 之	其 + 誰 + 知 + tjiə	0 0 0 468
1744	園有桃	109.2	12	蓋 + 亦 + 勿 + 思	蓋 + 亦 + 勿 + siə	0 0 0 468

In addition, the ALIGNMENT column allows us an even greater detail of the representation of our rhyme analysis, since we can use the column to share explicit phonetic alignments of our data, allowing for a much more fine-grained analysis of questions regarding impure rhymes.

Table 3. Illustrating the power of alignments in our rhyme annotation format.

ID	ALIGNMENT	RHYMEIDS
1733	(k) i ə k	467
1734	(dʲ) i ə k	467
1735		
1736	(kʷ) - ə k	467
1737		
1738	(q) i ə k	467
1739	(tz) - ə	468
1740	(g) i ə	468
1741	(tʲ) i ə	468
1742	(tʲ) i ə	468
1744	(s) i ə	468

Comparing this new format proposal with previous annotation frameworks, we can easily see that the possibility of annotating the similarity of rhyme words in the form of *phonetic alignments* offers a multitude of future possibilities, especially when more datasets are annotated in this form. Alignments would allow us not only to access automatically or formally the similarity between two or more rhyme words, they would also allow us to investigate cases of impure rhyming on a large scale, drawing statistics not only across poems that appeared in different epochs of the same language, but also across languages and cultures.

2.2.2. Software API for curation and analysis of rhyme datasets

We have developed a software API, called PoePy (https://github.com/lingpy/poepy), that allows one to parse, manipulate, and convert files following our new rhyme annotation schema in a convenient way, with help of the Python language. The framework builds heavily

on LingPy, a Python library for quantitative tasks in historical linguistics (List, Greenhill, and Forkel 2017), as well as SinoPy, a Python library for specialized tasks in Chinese historical linguistics (List 2018b). The GitHub site of our API offers additional information for installing and using our software library.

PoePy can read datasets in our general format mentioned above, it can also be used to align rhyme words, provided they are readily assigned to the data, and it can convert the data to different formats, that ease rhyme pattern inspection. Our stanza 2 from Ode 109 of the *Shījīng*, for example, can be rendered directly in the following tabular form, that greatly facilitates seeing the rhyme structure of the poem.

Table 4. Tabular representation of the rhyme schema underlying stanza 2 in Ode 109

ID	STANZA	LINE	R:467	R:468
1733	109.2	園 有 棘	kiək	
1734	109.2	其 實 之 食	djiək	
1735	109.2	心 之 憂 矣		
1736	109.2	聊 以 行 國	kuək	
1737	109.2	不 我 知 者		
1738	109.2	謂 我 士 也 罔 極	qiək	
1739	109.2	彼 人 是 哉		tzə
1740	109.2	子 曰 何 其		giə
1742	109.2	其 誰 知 之		tjiə
1744	109.2	蓋 亦 勿 思		siə

Figure 1. Colored HTML-output. Colors of the alignments in Wáng Lì's reconstruction indicate the basic sound class to which the sounds belong (alveolars, affricates and velars, vowels).

PoePy can also be used to output the data to HTML format, which allows for a convenient color-coding of rhyme patterns. This format can both be useful for inspection of datasets, or for sharing annotated rhyme data online. An example for our stanza 2 from Ode 109 from the *Shījīng* is given in Figure 1 below.

Given that our current format is rather tedious to produce, PoePy also offers a convenient parser from a much simpler format specification that uses inline-annotation of rhymes. In this format, the same Ode 109, stanza 2, would be rendered as follows:

```
@title: Ode 109
@annotator: Wáng Lì

園有[a/kiək]棘
其實之[a/djiək]食
心之憂矣
聊以行[a/kuək]國
不我知者
謂我士也罔[a/qiək]極
彼人是[b/tzə]哉
子曰何[b/giə]其
其誰知[b/tjiə]之
蓋亦勿[b/siə]思
```

Example 1. Inline format for Wáng Lì's analysis of Ode 109, Stanza 2.

Thus, one can see that the annotation can be easily achieved by using minimal inline markup, namely square brackets to indicate the rhyme (which is represented by alphabet letters here), along with the option to mark the reading. In a similar way, this format can also be used for a quick annotation of poetry in general. As an example, consider the following excerpt from Mike Naumenko's song "Leto, Pesnja dlja Tsoja" (*Summer, a song for Tsoj*, 1982).

```
@title: Leto. Pesnja dlja Tsoja
@author: Mike Naumenko
@year: 1982
@publisher: ЁRIO
@collection: LV
@editor: Mike Naumenko

[a]Лето!
Я изжарен, как кот[a]лета.
Время есть, а денег нету,
Но мне на это напле[b]вать.

[a]Лето!
Я купил себе га[c]зету.
Газета есть, а пива [c]нету.
И я иду его ис[b]кать.
```

Example 2. Inline format for Mike Naumenko's song *Leto* ("summer")

The first line is used to store the metadata, which is provided as a pair of a keyword and a value, while the following lines list the poem, separating different stanzas by adding a blank line. Once loading this file in text format with the PoePy library, the data can again be directly queried by printing a table illustrating the rhyme structure, or by querying general statistics about the data. These statistics would, for example, tell us that the song has 119 words in total, 32 lines, 8 stanzas, and 29 rhyme words. From this raw text form based on inline annotation, the data can, of course, also be directly converted to our more refined and flexible format, from where it can be further annotated.

Table 5. The first two stanzas of the song *Leto*. Since rhyme markers were placed in the middle of the rhyming words, they are now used to split the words into rhyming and non-rhyming parts

ID	STANZA	LINE	R:1	R:2	R:3
1	1.1	*Лето!*	лето		
2	1.1	Я изжарен, как *котлета.*	кот лета		
3	1.1	Время есть, а денег нету,			
4	1.1	Но мне на это *наплевать.*		напле вать	
5	1.2	*Лето!*	лето		
6	1.2	Я купил себе *газету.*			га зету
7	1.2	Газета есть, а пива *нету.*			нету
8	1.2	И я иду его *искать.*		ис кать	

2.3. Examples

2.3.1 Sample datasets

We have started to collect a number of sample datasets that we use for the illustration of our new formats. The largest collection includes the rhyme judgments by Baxter (1992) and Wáng (1980) for the *Shījīng*. In addition, we have started to annotate many small pieces of literature, especially poems, but also popular songs in different languages, which we use to illustrate the usefulness of our annotation system. In the future, we hope to be able to add more datasets in a more consistent manner, digitizing specifically alternative rhyme judgments of the *Shījīng* (such as the those of Karlgren 1950 and Starostin 1989), but also less frequently analyzed rhyme collections, especially from Hàn times.

2.3.2. Rhymes across languages and genres

In the following, we quickly illustrate how our format can be used to annotate rhymes in a much more consistent way than has been done before. Our collection is not bound to a particular language or a particular culture. On the contrary, since the goal of our annotation framework is to provide a much more profound way of annotating formed speech, we have tried to illustrate its usefulness by collecting small examples from different languages and genres.

As a first example, consider Joseph von Eichendorff's (1788–1857) poem *Zwielicht*, which was published as part of a novel in 1815. This poem contains four stanzas of four lines each, all written in form of an "envelope rhyme" (with the general schema "abba"). Our annotation example of stanza 1.1, in which we render the rhyme words in IPA and align them, putting non-rhyming parts of the words in brackets, makes it easy to quickly identify the impure rhyming of the first and the fourth line, which reflects a general peculiarity of German rhyming, in that the diphthongs [ai] and [ɔi] can rhyme freely (Table 6).

As another example, consider the first stanza of Bob Dylan's song "I want you" (from the album *Blonde on Blonde*, 1966). Here the rhyme patterns are more complex than in Eichendorff's poem, but rhyming is in parts also more lax, with more imperfect rhymes, reflecting the typical style of Dylan's poetry (Table 7).

Table 6. Eichendorff's *Zwielicht* (first stanza) in aligned form

ID	ST	LINE	R:1	R:2
1	1.1	Dämmrung will die Flügel *spreiten*	(ʃ p - r) ai t ə n	
2	1.1	Schaurig rühren sich die *Bäume*		(- b) ɔi m ə
3	1.1	Wolken ziehn wie schwere *Träume* -		(t r) ɔi m ə
4	1.1	Was will dieses Graun *bedeuten?*	(- b ə d) ɔi t ə n	

Table 7. Bob Dylan's *I want you* in aligned form

ID	ST	LINE	R:1	R:2	R:3
1	1.1	The guilty undertaker *sighs*	s - ai s		
2	1.1	The lonesome organ grinder *cries*	k r ai s		
3	1.1	The silver saxophones *say*	s - æi -		
4	1.1	I should *refuse_you*		r i f j u: s j u:	
5	1.1	The cracked bells and washed-out *horns*			h - ɔ r n s
6	1.1	Blow into my face with *scorn,*			s k ɔ r n -
7	1.1	but it's not that way, I wasn't *born*			b - ɔ r n -
8	1.1	to *lose_you*		- - - l u: s j u:	

As a further example, the following table presents the first and the third stanza from the famous Chinese song "Yuèliàng dàibiǎo wǒ de xīn", which was popularized in the 1977 version by Teresa Teng. In our analysis of this song, lines 5 and 12 are believed to rhyme with rhyme group R:1, which may be problematic, as it seems that not all native speakers of Mandarin Chinese accept rhymes of *-en* [ən] and *-in* [in]. However, since our analysis will make the overall rhyme schema of the song appear much more harmonic, we think that this reflects the intention of the song writer.

Table 8. Rhyme annotation for *The moon expresses my heart*

ID	ST	LINE	R:1	R:3
1	1.1	你 問 我 愛 你 有 多 **深**	sh ēn	
2	1.1	我 愛 你 有 幾 **分**	f ēn	
3	1.1	我 的 情 也 **真**	zh ēn	
4	1.1	我 的 愛 也 **真**	zh ēn	
5	1.1	月 亮 代 表 我 的 **心**	x īn	
11	1.3	輕 輕 的 一 個 **吻**	w ěn	
12	1.3	已 經 打 動 我 的 **心**	x īn	
13	1.3	深 深 的 一 段 **情**		q íng
14	1.3	叫 我 思 念 到 如 **今**		l ìng

This case shows that the question of whether a given rhyme is indeed intended by a poet or not, may not always be easily solved, and precisely for this reason it is necessary to have frameworks in which the analyses of different readers can be compared. A further example is the song *Te doy una canción* by Silvio Rodriguez (from the album *Mujeres*, 1978), in which none of the three rhyme pairs which we have annotated in stanza 1.2 rhymes perfectly. One might

thus assume that rhyming was generally not intended in this song, but we find a very similar pattern in stanza 1.4., and songs in which the words *tú* "you" and *luz* "light" co-occur in potential rhyming position are very frequent in Spanish songs. Our hope is, that with a growing body of datasets in this form, we may learn more about the difference between rhymes which are intended and rhymes which might occur simply by chance.

Table 9. Silvio Rodriguez' "Te doy una canción": are the rhymes intended?

ID	ST	LINE	R:1	R:2	R:3
7	1.2	Te doy una canción si abro una *puerta*	puer ta		
8	1.2	Y de las sombras sales *tú*		tú	
9	1.2	Te doy una canción de *madrugada,*	madruga da		
10	1.2	Cuando más quiero tu *luz*		luz	
11	1.2	Te doy una canción cuando apareces			
12	1.2	El misterio del *amor*			a mor
13	1.2	Y si no apareces, no me importa:			
14	1.2	Yo te doy una *canción*			can ción

As two final examples in this section, let us get back to rhyming in Classical Chinese. In Weingarten (2016), rhyming maxims supposedly spoken by Confucius, quoted in the Han period *Shuoyuan* 說苑, are presented and analyzed. Such examples potentially provide valuable evidence for the reconstruction of Old Chinese phonology, in addition to its later development into the Han dynasty. It would be desirable if a general corpus could be constructed in which all pieces of evidence that can be found throughout different epochs of Chinese language history could be assembled. If we compare the original annotation provided in the text by Weingarten with our extended schema, we think it is obvious how much standardized representations of rhyme judgments, collected collaboratively by all experts in the field, could advance our knowledge about the history of Chinese phonology.

Table 10. Rhymes in Confucius' work (as detected by Weingarten 2016)

ID	ST	LINE	R:1	R:2
1	1	夫 人 君 無 諫 臣 則 失 **政**	t e ŋ h	
2	1	士 無 教 **友** 則 失 聽	lh ê ŋ h	
3	1	狂 馬 不 釋 其 **策**		tsh r ê k
4	1	操 弓 不 **反** 於 **繁**	g r e ŋ	
5	1	木 受 繩 則 **直**		d r ə k
6	1	人 受 諫 則 **聖**	lh e ŋ h	
7	1	受 學 重 問 孰 不 順 **成**	d e ŋ	
8	1	毀 人 惡 士 且 近 於 **刑**	g ê ŋ	
9	1	君 子 不 可 以 不 學		

In addition to the received corpus of Chinese texts, recently unearthed manuscript sources are now also providing a rich new data set for the study of rhyming in early China. Working with these sources however often requires exhaustive notations about the witnesses consulted, the condition of the physical material carrier, the presence of textual variants, and so forth. Our proposed schema for standardizing the presentation of rhyming judgments is flexible

enough to accommodate a more extensive critical apparatus. Consider for instance the following content from the *Cāng Jié piān* 蒼頡篇, a long-lost scribal primer of great importance during the Han dynasty, which has been rediscovered among the manuscript finds of the past century. The table below presents the first twelve lines to its "opening chapter," which establishes a *zhī* 之/ *zhí* 職 cross rhyme every other four-character line. Close attention will be paid only to variants in rhyming positions.

Table 11: Rhymes with variants in the *Cang Jie pian* "opening chapter"

ID	LINE	LO	RHYMEIDS	RW	PROTOFORM	SOURCE
1	蒼 + 頡 + 作 + 書	1	0 0 0 0			JYX EPT50.1
2	以 + 教 + 後 + 嗣	2	0 0 0 1	嗣	*ziəᶜ	JYX EPT50.1
3	以 + 教 + 後 + 子	2	0 0 0 1	子	*tsiəᶜ	JYX EPT56.40
4	以 + 教 + 後 + 生	2	0 0 0 1	生	*ṣeŋ	Cang Jie Mirror
5	幼 + 子 + 承 + 昭	3	0 0 0 0			JYX EPT50.1
6	謹 + 慎 + 敬 + 戒	4	0 0 0 1	戒	*kɛᶜ	JYX EPT50.1
7	謹 + 慎 + 敬 + 式	4	0 0 0 1	式	*śɨk	DHHJ 1459
8	勉 + 力 + 風 + 誦	5	0 0 0 0			JYX EPT50.1
9	晝 + 夜 + 勿 + 置	6	0 0 0 1	置	*ṭiəᶜ	JYX EPT50.1
10	苟 + 務 + 成 + 史	7	0 0 0 0			JYX EPT50.1
11	計 + 會 + 辨 + 治	8	0 0 0 1	治	*ḍiəᶜ	JYX EPT50.1
12	超 + 等 + 軼 + 羣	9	0 0 0 0			JYX EPT50.1
13	出 + 尤 + 別 + 異	10	0 0 0 1	異	*jəᶜ	JYX EPT50.1
14	□ + □ + □ + 夜	10	0 0 0 1	夜	*jaᶜ	JY 260.18
15	初 + 雖 + 勞 + 苦	11	0 0 0 0			JYX EPT50.1
16	卒 + 必 + 有 + 意	12	0 0 0 1	意	*ʔiəᶜ	JYX EPT50.1
17	卒 + 必 + 有 + 憙	12	0 0 0 1	憙	*hɨəᴮ	YT 3380

In this table, a bamboo strip found among the Juyan II cache (JYX EPT50.1) is used as the base text. This strip carries an almost complete version of the "opening chapter" to the *Cāng Jié piān*, running from its recto to verso. There are, however, eighty-seven manuscript fragments altogether with content potentially related to this section of the *Cāng Jié piān*, and some include variants in rhyming positions (Foster 2017: 272). To reflect this, the table above adds rows for lines where variants are found. Given that all variants are assigned the same number in the Line Order (LO) column, but different sources in the new source (SOURCE) column to cite which fragment carries the variant in question, they can be automatically detected and contrasted with one another.

In this way, the table above is able to quickly communicate where variants exist among our manuscript sources, highlighting those which could impact our understanding of rhyming in the text. Thus, we find three variants for the second line in the stanza (rows 2, 3, and 4, all given a 2 in the LO column). Our base text writes *ziə* 嗣 "descendants" (following Axel Schuessler's 2009 reconstruction of Later Han Chinese), but in the variant in row 3, we find the similar-sounding synonym *tsiə* 子 "children" in JYX ETP56.40 instead, while a later bronze mirror inscription given in row 4 (Cang Jie Mirror) bears a potential variant of *ṣeŋ* 生 "offspring", which is phonetically incongruent. Similarly, our base text has *kɛ* 戒 "instructions" at the end of line 4 (row 6), whereas strip DHHJ 1459 (and also 1460 and 1461, not shown here) appear to write *śɨk* 式 "models" in row 7. The rhyme word *jə* 異 "extraordinary" in line

number 10 is replaced by *jɐ̌^{C} "night" in JY 260.18 (row 14). Finally, our base text on JYX EPT50.1 concludes line 12 with *ʔiə 意 "think of" (line 16), but YT 3380 has the variant * hiə^{B} 憙 "desire" (line 17).

Of course, to better understand these variants, we must examine the manuscripts on which they are found. Because the *Cāng Jié piān* was a scribal primer, it was often used to practice writing. Uncertainty surrounds a number of the variants given above. For example, a novice hand was responsible for *śik 式 on DHHJ 1459, and scholars have debated the appropriateness of this transcription considering its odd orthography (Foster 2017: 267-268). Whether or not * ja$_{C}$ 夜 truly belongs to line 10 on JY 260.18 is also ambiguous. Damage to the material carrier has removed the text above it, destroying valuable context. Moreover, the writing is not always aligned consistently on the board, with some characters repeated or brushed on in different sizes, leading one to question if these are random scribbles, without any line coherency (Foster 2017: 274-275).

A more interesting case is presented with * tsiə^{C} 子 at line 2 on JYX ETP56.40 (row 3). Following this word, the text on JYX ETP56.40 continues to differ dramatically from most of our other wood and bamboo-strip witnesses. Although this too may be garbled practice writing, a parallel with another strip, YT 1855, perhaps betrays that this is an altogether different edition of the *Cāng Jié piān*, or even another text (Foster 2017: 119f and 122). If we wanted to reflect this uncertainty and include edition-level variance in our table, we could add them as separate rows in our file, specifying to which line they would pertain. Thus, underneath our current row 5 (LO 3), which reads 幼子承昭, we could add additional rows for JYX EPT56.40 (為史□□) and YT1855 (為史知[莫?]), still labeling them LO 3.

The bronze mirror inscription included in the table above offers a more radical example (Foster 2017: 111). The opening two lines of the *Cāng Jié piān* are either quoted or coincidentally incorporated into other material, which on the whole is a different text entirely: "I have cast this luminous mirror, (in imitation of how) the three kings (of yore) invented decorum. Kingfishers' feathers (make for) a marvelous canopy, and a numinous turtle (serves as) support for the umbrella post. Cang Jie created writing, and taught it to later offspring. Suiren made fire, and the five flavors [ripened]. 余造明鏡，三王作容，翠羽秘盖，靈鳩（龜）臺杠，倉頡作書，以教後生，遂（燧）人造火，五味[熟成]". The word *ʂeŋ 生 is adopted because it fits better into the rhyming of this new text, where it is paired with the words *joŋ 容, *kɔŋ 杠, and potentially *dźeŋ 成 instead (Péng Yǔ 2014, with additional comments by Wáng Níng 王寧, et al.). If desired, all of this could be re)flected in the table as well, in the same manner in which we demonstrated how variants can be presented in the text.

2.3.4. Comparing differences in rhyme annotations

In List et al. (List et al. 2017), rhyme networks were used to test to which degree different reconstruction systems conform to what Ho (2016) calls "vowel purity", namely the hypothesis that rhyming practice in Old Chinese (and probably also later) was very strict in adhering to identical vowels in rhyming. The test by List et al. (2017) revealed that the system of Baxter and Sagart (2014) (and of six-vowel theories of Old Chinese in general) reflects the principle of vowel purity much more closely than do systems with more vowels (Karlgren 1950) or fewer vowels (Wáng 1980; Lǐ 1971).

In this context, it is important to recall that — what was also mentioned in the paper by List et al. (2017), but might easily be misunderstood by readers — the adherence to vowel purity cannot be used to prove or disprove a given reconstruction system, since the adherence to vowel purity is a hypothesis about Old Chinese rhyming practice itself, and we know well that vowel

purity in rhyming can be easily abandoned or disregarded across rhyming traditions in different cultures. Apart from the problem that studies on vowel purity do not bear any diagnostic value with respect to the accuracy of reconstruction systems, one additional problem in the study by List, et al. (2017) is the fact that vowel purity itself was only tested by comparing the rhyme judgments of one source Baxter (1992) with different reconstruction systems. Given that Baxter himself is reconstructing a six-vowel system on the basis of rhyme evidence, it is quite likely that the rhyme decisions proposed by Baxter (1992) could have influenced the analysis.

While alternative rhyme judgments were not available when drafting the original study on vowel purity, we have now, thanks to our new format for rhyme annotations, also had the time to digitize the rhyme judgments reported in Wáng (1980). Given that two different rhyme analyses have been digitized now, it is interesting and also important for the reconstruction of Old Chinese Phonology to check to which degree different scholars differ in what they judge to rhyme and what not.

We can think of different measures to compare the difference in the actual rhyme judgments of the two versions. A simple measure is to compare how many stanzas differ. From 1070 common stanzas, 175 are different between Wáng and Baxter, which amounts to 15.9%. A far more interesting aspect is to check *how much* different stanzas differ. Similar to a common partitioning task by which we compare to which degree two partitions of the same data differ, we can do this with help of the B-Cubed scores (Amigó et al. 2009; List, Greenhill, and Gray 2017), since the assessment for a given stanza, whether two words rhyme or not, can also be thought of as a clustering task (authors decide which words belong to the same rhyme partition in a given cluster). Applying B-Cubed scores to compare the rhyme judgments, with help of the PoePy library, to which we added a function to compare different rhyme judgments (implementing the code presented in List 2018a), we find 97% of similarity between Baxter's and Wáng's rhyme judgments. This means that the internal difference between the rhyme judgments by Baxter and Wáng is less pronounced than one might think when only checking whether a given stanza is interpreted differently in *any* way.

Table 11. Comparing Wáng's and Baxter's rhyme analysis of Ode 71, Stanza 1. For Baxter's analysis, our current digitized version does not have the original reconstructions, which is why the software only shows the rhyming characters instead

ID	ST	LINE	R:331
1208	71.1	綿 綿 葛 藟	
1209	71.1	在 河 之 滸	xa
1210	71.1	終 遠 兄 弟	
1211	71.1	謂 他 人 父	biua
1213	71.1	亦 莫 我 顧	ka

(a) Wáng's rhyme analysis.

ID	ST	LINE	R:319	R:320
1229	71.1	綿 綿 葛 藟	藟	
1230	71.1	在 河 之 滸		滸
1231	71.1	終 遠 兄 弟	弟	
1232	71.1	謂 他 人 父		父
1234	71.1	亦 莫 我 顧		顧

(b) Baxter's analysis

As an example for differences in Baxter's and Wáng's rhyme annotations, compare stanza 1 in Ode 71, which is given in the version of both authors below. As can be seen from this example, both authors agree regarding the rhyming of *xǔ* 許, *fù* 父, and *gù* 顧, but while in Wáng's analyses these three characters are the only ones that write rhyming words, Baxter's analysis assumes in addition, that *lěi* 藟 and *dì* 弟 rhyme as well.

3. Summary and Outlook

In this paper, we have proposed a new framework for rhyme annotation that can be used for a more consistent rendering of the rhyme judgments proposed by different scholars. The framework is inspired by general attempts to standardize cross-linguistic data within the Cross-Linguistic Data Formats initiative, and offers a software library that can be used to check, curate, and analyze rhyme data which has been annotated according to our format specifications. We have illustrated the usefulness of the framework by providing examples of how different cases can be handled. Thanks to the format, we can furthermore easily compare different rhyme annotations in a consistent way. In the future, we hope to expand the so far rather small database of rhyme annotations we have assembled so far. We hope, however, also that our annotation framework will convince our fellow colleagues to help increase the evidence for Old Chinese reconstruction by publishing their future rhyme analyses in a transparent form. Given the multitude of open problems related to the history of the Chinese language from its origins until today, we will only be able to advance our field when working in collaboration and sharing our data in a transparent form.

Source Code and Data

The data discussed in this paper is available along with the PoePy library, which can be accessed on GitHub at https://github.com/lingpy/poepy, and will be officially released in case this paper gets accepted. The code to run the experiments discussed in this paper (especially the comparison of two rhyme datasets) is also available from this repository: https://doi.org/10.5281/zenodo.3252141.

References

Amigó, Enrique, J. Gonzalo, J. Artiles, F. Verdejo, 2009. A Comparison of Extrinsic Clustering Evaluation Metrics Based on Formal Constraints. In: *Information Retrieval* 12(4): 461–86.

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: De Gruyter.

Baxter, William H., Laurent Sagart. 2014. *Old Chinese. A New Reconstruction*. Oxford University Press.

Eckart, Kerstin. 2012. Resource Annotations. In: *CLARIN-d User Guide*, edited by AP 5 Clarin-D: 30–42. Berlin: DWDS.

Forkel, Robert, Johann-Mattis List, Simon J. Greenhill, C. Rzymski, S. Bank, M. Cysouw, H. Hammarström, M. Haspelmath, G. A. Kaiping, R. D. Gray. 2018. Cross-Linguistic Data Formats, Advancing Data Sharing and Re-Use in Comparative Linguistics. *Scientific Data* 5(180205): 1–10.

Foster, C. J. 2017. *Study of the* Cang Jie pian: *Past and Present.* Ph.D. Harvard University.

Hill, Nathan W., Johann-Mattis List. 2017. Challenges of Annotation and Analysis in Computer-Assisted Language Comparison: A Case Study on Burmish Languages. *Yearbook of the Poznań Linguistic Meeting* 3(1): 47–76.

Ho, Dah-an. 2016. Such Errors Could Have Been Avoided. Review of "Old Chinese: A New Reconstruction" by William H. Baxter and Laurent Sagart. *Journal of Chinese Linguistics* 44(1): 175–230.

Karlgren, Bernard. 1950. *The Book of Odes: Chinese Text, Transcription and Translation*. Stockholm: Museum of Far Eastern Antiquities.

Li, Fang-kuei (李方桂). 1971. Shànggǔyīn Yánjiū [Investigations of the phonological system of Old Chinese]. *Qīnghuá Xuébào* (清華學報) 9(1-2): 1–60.

List, Johann-Mattis. 2014. *Sequence Comparison in Historical Linguistics*. Düsseldorf: Düsseldorf University Press.

List, Johann-Mattis. 2018a. More on Network Approaches in Historical Chinese Phonology (音韵学). In: *The 2nd Li Fang-Kuei Society Young Scholars Symposium*: 157–74. Taipei: Li Fang-Kuei Society for Chinese Linguistics.

List, Johann-Mattis. 2018b. *SinoPy: A Python Library for Quantitative Tasks in Chinese Historical Linguistics*. Jena: Max Planck Institute for the Science of Human History. Available at: https://github.com/lingpy/sinopy.

List, Johann-Mattis, Simon J. Greenhill, and Russell D. Gray. 2017. The Potential of Automatic Word Comparison for Historical Linguistics. *PLOS ONE* 12(1): 1–18.

List, Johann-Mattis, Simon J. Greenhill, Robert Forkel. 2017. *LingPy. A Python Library for Quantitative Tasks in Historical Linguistics* (version 2.6). Jena: Max Planck Institute for the Science of Human History. doi:https://doi.org/10.5281/zenodo.1065403.

List, Johann-Mattis, J. S. Pathmanathan, Nathan W. Hill, E. Bapteste, P. Lopez. 2017. Vowel Purity and Rhyme Evidence in Old Chinese Reconstruction. *Lingua Sinica* 3.1: 1–17.

Milà–Garcia, Alba. 2018. Pragmatic Annotation for a Multi-Layered Analysis of Speech Acts: A Methodological Proposal. *Corpus Pragmatics* 2(1): 265–87.

MPI EVA. 2008. *The Leipzig Glossing Rules: Conventions for Interlinear Morpheme-by-Morpheme Glosses*. Leipzig: Max Planck Institute for the Science of Human History. Available at: http://www.eva.mpg.de/lingua/resources/glossing-rules.php.

Pān, Wùyún (潘悟云). 2000. *Hànyǔ lìshǐ yīnyùnxué* [*Chinese historical linguistics*]. Shànghǎi: Shànghǎi jiàoyù (上海教育).

Péng, Yǔ 鵬宇. 2014. 'Cāng Jíe Zuòshū' Míng Sānduànshì Tóngjìng Jiǎnshì (倉頡作書銘三段式銅鏡束釋) [*Assorted Notes on the Interpretation of the "Cang Jie Created Writing" Three-Register Bronze Mirror Inscription*]. Fùdàn Dàxué Chūtǔ Wénxiàn yǔ Gǔwénzì Yánjiū Zhōngxīn. October 18th, 2014. Available at: http://www.gwz.fudan.edu.cn/Web/Show/2351 (accessed: May 6, 2019).

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

Starostin, Sergei A. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy (Reconstruction of the Phonological System of Old Chinese)*. Moscow: Nauka.

Wáng, Lì (王力). 1980. *Shījīng Yùndú* [*Rhyme annotations on the Book of Odes*]. Shànghǎi: Shànghǎi gǔjí (上海古籍).

Wáng, Xiǎ (王显). 2011. *Shījīng Yùnpǔ* (诗经韵谱) [*Rhymes in the Book of Odes*]. Běijīng: Shāngwù (商務).

Weingarten, Oliver. 2016. The Singing Sage: Rhymes in Confucius Dialogues. *Bulletin of the School of Oriental and African Studies* 79(3): 581–607.

Zhengzhang, Shangfang. 2000. *The Phonological System of Old Chinese*. Translated by Laurent Sagart. Paris: École des Hautes Études en Sciences Sociales.

Zhèngzhāng, Shàngfāng (郑张尚芳). 2003. *Shànggǔ Yīnxì* [*The phonological system of Old Chinese*]. Shànghǎi: Shànghǎi jiàoyù (上海教育).

Й.-М. Лист, Н. Хилл, К. Фостер. К вопросу стандартизированной аннотации рифмовки в древнекитайских текстах (и не только)

Несмотря на то, что анализ рифм играет ключевую роль в реконструкции древнекитайской фонологии, в этой области до сих пор отсутствует сколь-либо стандартизированная формальная система аннотаций рифмовки в древнекитайских текстах. Опираясь на предыдущий опыт стандартизирования кросс-лингвистических данных в целях исторического и типологического сравнения (в рамках инициативы Cross-Linguistic Data Formats), мы в данной статье предлагаем прозрачный и последовательный формат для такого рода аннотаций, который позволил бы исследователям организовывать выявленные рифмы в древних текстах таким образом, что их можно было бы, с одной стороны, анализировать компьютерными методами, с другой, удобно использовать в

ручном режиме. Формат разработан вместе с соответствующим программным обеспечением и с образцовыми таблицами данных, аннотированными различными исследователями, и предназначен не только для китайской, но и для современной иноязычной поэзии. В статье подробно описывается как сам формат, так и связанные с ним возможные проблемы и недостатки текущих аннотаций. Мы надеемся, что эта работа вдохновит других исследователей, занимающихся вопросами древнекитайской реконструкции, на собственные предложения в этой области, которые в дальнейшем приведут к пересмотру и улучшению тех или иных его параметров.

Ключевые слова: древнекитайский язык, китайская система рифм, аннотация данных, кросс-лингвистические форматы данных

Shen Ruiqing
National University of Singapore; shenruiqings@hotmail.com

The monosyllabicization of Old Chinese and the birth of Chinese Writing: A hypothesis on the co-evolution of the Chinese language and its writing system[1]

The invention of the Ancient Chinese Writing System (henceforth ACWS) is a significant event in world history. In this paper I put forward a hypothesis on the co-evolution of the Old Chinese language and its writing system (ACWS). I argue that the invention of ACWS bears a strong correlation with the linguistic evolution, more specifically, the monosyllabicization, of Old Chinese. In other words, ACWS might never be invented if monosyllabicization had not occured in Chinese.

The paper is organized in the following way. First, we discuss the reason why a subsyllabic writing system was not invented for Old Chinese (section 2). Next, we discuss the nature of the rebus principle in ACWS (section 3), and its correlation with morphological alternations (section 4). Then I argue that monosyllabicization of Old Chinese is a precondition for the rebus principle, which is crucial for the birth of ACWS (section 5). Lastly, I discuss the implication of the hypothesis for the study of Old Chinese (section 6).

Keywords: Old Chinese language, Chinese writing, monosyllabicization, rebus principle, morphological alternations.

Introduction

The invention of the Ancient Chinese Writing System (henceforth ACWS) is a significant event in world history. The nature and origin of ACWS, however, is still not well understood. In this paper I put forward a hypothesis on the co-evolution of the Old Chinese language and its writing system, ACWS. I argue that the invention of ACWS bears a strong correlation with the linguistic evolution, more specifically, the monosyllabicization of Old Chinese. It is not accidental that Old Chinese became the first language that developed an independent writing system in Southeast Asia.

First, the definition of ACWS and Old Chinese needs some clarification. I use the term ACWS to refer to the earliest systematic writing in Ancient China. In this paper, examples mostly come from oracle bone writing from the later Shang dynasty in the late 2nd millennium BCE, as they are the earliest known form of Chinese writing so far (Keightley 1985). Academics still debate on the question of when ACWS was first invented. Some scholars suggest that

writing systems have been invented as early as the 3rd millennium BCE (Qiu 2013: 34). However, this does not affect the hypothesis put forward here. My argument is that regardless of when the ACWS was invented, certain linguistic innovations must have occurred before this invention.

Old Chinese is defined in this paper as the Chinese language reflected in the earliest known texts, i.e. in the late 2nd millennium BCE. This definition is different from Baxter (1992: 24) where Old Chinese is loosely defined as "Chinese of early and middle Zhou dynasty" (roughly the early and middle 1st millennium BCE), as well as Baxter & Sagart (2014: 1) where Old Chinese is defined in a broad sense to refer to "varieties of Chinese used before the unification of China under the Qin dynasty in 221 BCE". We do not know yet how much the language of Shang differs from that of Zhou. However, as far as the topic of this paper is concerned, the language of Shang is more closely related to the invention of ACWS than that of Zhou.

Why not a sub-syllabic writing system?

Unlike modern alphabets, ACWS is a writing system that does not break down syllables into sub-syllabic units. It is natural for a person who lives in our modern, global world to wonder why an alphabetic writing system was never independently invented in China or even Asia in general. The answer to this question is related to the nature of speech production and perception. Many phoneticians and phonologists think of sub-syllabic units like consonants and vowels as universal phonological concepts. However, Ladefoged (2005) disagrees:

> Talking involves pulling stored forms of words out of some part of the brain, but words are not stored as sequences of sounds. They are stored as wholes, or at least as whole syllables, in which the consonants and vowels are not separate items. [...] The symbols of the alphabet represent segments of speech, and it is probably from thinking in terms of these symbolized segments that we get the idea that there are separate sounds (Ladefoged 2005: 186–187).

For Ladefoged, sub-syllabic units like consonants and vowels are not universal primitives. Therefore, the history of alphabetic writing should be viewed as a unique invention:

> Breaking syllables up into vowels and consonants was an enormous scientific achievement. Speakers of other languages saw what could be done and started using alphabetic characters. But the original notion that syllables could be split into vowels and consonants occurred only once in human history. [...] We also lose out in that our thinking about words and sounds is strongly influenced by writing. We imagine that the letters of the alphabet represent separate sounds instead of being just clever ways of artificially breaking up syllables. Alphabetic writing has almost certainly been invented only once, whereas there are many independent inventions of systems for writing down syllables (Ladefoged 2005: 189–190).

Phonological knowledge of sub-syllabic units was likely non-existent in native Chinese tradition; it only developed at a later stage through intensive communication with other languages[2]. Therefore, a sub-syllabic writing system was not possible during the time that ACWS was invented. Even for the contemporary Chinese language, O'Seaghdha, Chen & Chen (2010: 297) argue based on their psycholinguistic studies that "the original or proximate phonological encoding units in Mandarin are syllables whereas in English and related languages they are phonemic segments".

---

[2] For the development of native Chinese phonological theories, see Halliday (1981).

The rebus principle in ACWS

The invention of writing systems was an important milestone in human history. Historically, all ancient writing systems began with the invention of pictograms, semantic symbols for words with concrete meanings. The next stage of development involved figuring out how to represent abstract concepts and functional meaning using only a limited set of symbols, so as to cover the enormous mental lexicon stored in a speaker's brain. One of the most common strategies to solve this problem, developed independently in many ancient writing systems, is the rebus principle. The rebus principle refers to the use of existing symbols purely for their sounds, regardless of their meaning, to represent new words. The application of the rebus principle, which involves the activation of sound at a certain phonological level, is crucial in the invention of all mature writing systems[3].

In ACWS, the rebus principle was frequently used in the earliest documented oracle bone writing.[4] A well-known example of rebus principle in Late Shang oracle bone writing is shown in Figure 1.

Figure 1. An example of rebus principle in oracle bone writing
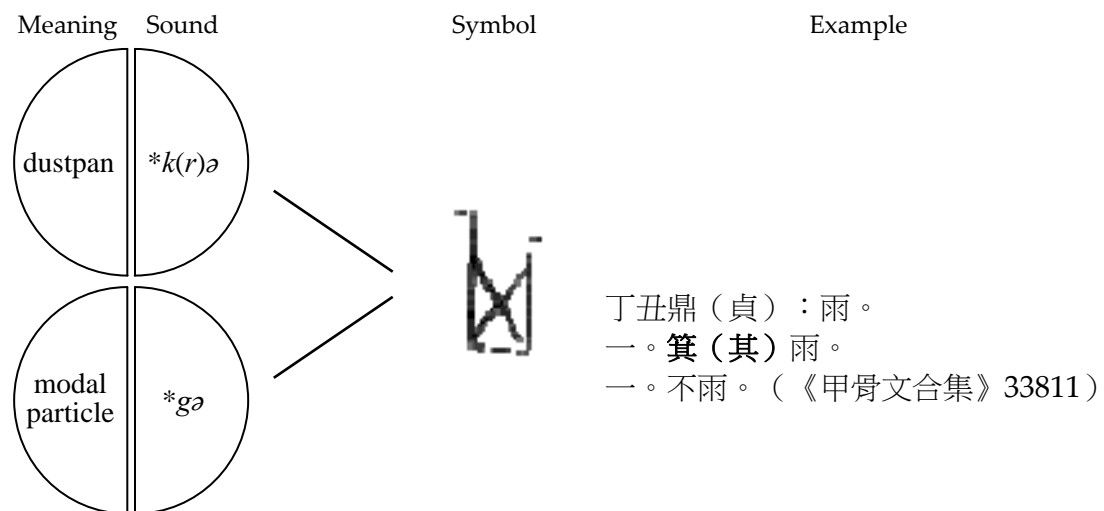
In Figure 1, the symbol invented for the word 'dustpan' 箕 was used for the similar-sounding word 'modal particle' 其. This example shows that the rebus principle was often applied to words with abstract meaning. In addition, when applying rebus principle, scribes from different ages or different groups may have used different symbols to write the same word with abstract meaning. For example, words such as 'disaster', 'difficult', 'finish', 'lose', 'faint', 'morning', 'all', 'raise', and 'strange' were written with different symbols by different groups of scribes (Chen 2007).

According to Xia (2014: 72), the rebus usage of characters in oracle bone writing is over 70%[5]. Although the rebus principle was frequently used, it was not always used whenever

---

[3] For the development of ancient writing systems, see Robertson (2004) and references therein.

[4] This process is related to the term *jiǎjiè* (literally meaning "borrowing") in the traditional *liùshū* categorization of Chinese characters developed during the Han period (see Boltz 1994: 143–155 for more details about *liùshū*). However, there is a significant difference between the two terms. While *jiǎjiè* refers to a certain type of method to invent new characters, the rebus principle refers to the usage of characters regardless of their origins.

[5] The rebus principle is much broader than the *jiǎjiè* method in traditional *liùshū* categorization. Li (1974) found that 129 out of 1156 (about 11%) characters were invented by the *jiǎjiè* method.

possible. On the contrary, in numerous cases when rebus usage was possible, it was not applied. Figure 2 shows one such example.

Figure 2. An example of the limitation of rebus principle in oracle bone writing

Meaning	Sound	Symbol	Example

In Figure 2, the word 'next (day)' can be represented by two kinds of graphs, one with a pictograph of 'wing', and the other with a complex character composed of a pictograph of 'wing' plus an additional pictograph of 'stand'. However, the symbol for 'stand' is used only for 'stand' and never appears as a rebus for 'next (day)'. The reason why there is no rebus between 'tomorrow' and 'stand' is not because they are not phonetically close enough; otherwise 'stand' would not be added as a symbol indicating the sound in the complex character. Instead, it is because the rebus usage of 'wing' for 'next (day)' is well established, and 'stand' is frequently used for its original meaning 'stand'; thus the two symbols were kept apart to signal different semantic meanings.

The limited application shown here and other numerous cases indicate that the rebus principle is only a backup option when the semantic meaning of a word is difficult to convey directly by pictograph. Unlike a syllabic writing system where most symbols have lost their semantic meanings and are only used to represent sounds, most concrete words are represented by pictographs in ACWS even when rebus usage of other symbols is possible. Therefore, the distinction of meanings is much more important than economy of symbols in ACWS. In fact, this is one of the main reasons why most scholars consider ACWS as a logographic writing system rather than a syllabic writing system.

Rebus and Old Chinese morphological alternations

In ACWS, rebus principle was applied to both homophonic and near-homophonic monosyllabic words. For near-homophonic pairs, certain phonological contrasts did not block the usage of rebus principle, while others frequently did. Three phonological contrasts in Old

Chinese which almost never blocked rebus usage are *qīngzhuó*[6], *sìshēng*[7], and *děng*[8] respectively. Table 1 contains some examples.

Table 1. Three examples of rebus usage across certain phonological contrasts

	Gloss	Middle Chinese phonological contrasts[9]			Example
		qīngzhuó	*sìshēng*	*děng*	
箕	dustpan	*qúanqīng*			丁丑鼎（貞）：雨。一。**箕（其）**雨。
其	modal particle	*qúanzhuó*			一。不雨。（《合集》33811）
皿	utensil	*cìzhuó*	*shǎng*		乙丑卜，設鼎（貞）：甲子**皿（向）**乙丑王夢牧石麋，不隹
向	to face	*qúanqīng*	*qù*		（唯）囚（憂），隹（唯）又（祐）。（《合集》376）
擐	put on clothing	*cìzhuó*	*qù*	2nd	□其田于□，其**擐（遠）**，〔湄〕日亡※（災）。
遠	far	*qúanzhuó*	*shǎng*	3rd	

In Table 1, the first rebus usage ('dustpan' for 'modal particle', see Figure 1 for details) involves *qīngzhuó* contrast, the second rebus usage ('utensil' for 'to face', see Qiu 1993) involves both *qīngzhuó* contrast and *sìshēng* contrast, and the last rebus usage ('put on clothes' for 'far', see Qiu 1985) involves all three contrasts. In these cases and many others, rebus is not blocked by these three phonological contrasts. Notably, all three also appear in morphological alternations in Old Chinese[10]. Table 2 gives some examples.

Table 2. Examples of morphological alternations in Old Chinese

	Gloss	Middle Chinese phonological contrasts		
		qīngzhuó	*sìshēng*	*děng*
敗	defeat (v.t.)	*qúanqīng*		
敗	suffer defeat	*qúanzhuó*		
受	receive		*shǎng*	
授	give		*qù*	
封	enfeoff			3rd
邦	country			2nd
入	enter		*rù*	3rd
内	inside		*qù*	1st
學	learn	*qúanzhuó*	*rù*	
教	teach	*qúanqīng*	*qù*	

---

[6] *Qīngzhuó* (清浊) is a traditional term related to Middle Chinese consonantal initials. There are four subsets: *qúanqīng* (全清), *cìqīng* (次清), *qúanzhuó* (全浊), and *cìzhuó* (次浊).

[7] *Sisheng* (四聲) is a traditional term related to Middle Chinese tones. There are four subsets: *pīng* (平), *shǎng* (上), *qù* (去), *rù* (入).

[8] *Děng* (等) is a traditional term related to Middle Chinese Medial and/or vowel qualities. There are four subsets: 1st (一等), 2nd (二等), 3rd (三等), and 4th (四等).

[9] These are traditional Chinese phonological terms invented for Middle Chinese. I intentionally use these abstract terms in order to avoid the controversy over their phonetic details in Old Chinese. See the previous footnotes for details.

[10] Old Chinese shows no evidence of reconstructing inflectional morphology; only derivational morphology can be reconstructed (LaPolla 2003).

In Table 2, there are five pairs of morphological alternations. The first three pairs differ in *qīngzhuó*, *sìshēng*, and *děng* respectively. The fourth pair differs in both *sìshēng* and *děng*, while the last pair differs in both *qīngzhuó* and *sìshēng*. The *qīngzhuó* alternation is one of the most frequent morphological alternations in Old Chinese[11]. The *sìshēng* alternation is also an active morphological device that has received much attention[12]. The *děng* alternation has received relatively less attention[13].

Why are the phonological contrasts omitted in the writing system and those used as morphological devices the same — is this just a coincidence?

According to classic phonological theories (Dresher, Piggott, & Rice 1994), a phonological relationship is either contrastive or non-contrastive. In other words, contrastive phonological features are contrastive to the same extent. However, Hall (2009) proposes a new understanding of phonological relationships in terms of a continuum with contrastive and non-contrastive at either end.

Figure 3. A continuous set of phonological relationships (Hall 2009:16)

| Non-overlapping distributions | Overlapping distributions |
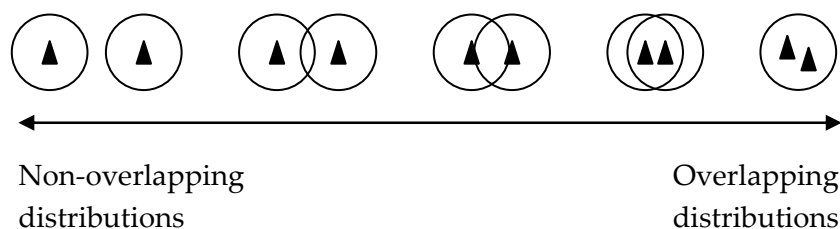
According to the proposal illustrated in Figure 3, phonological relationships are gradient rather than categorical. Besides the contrastive and non-contrastive relationships, there could be something in between. Multiple factors could lead to an intermediate status of this phonological relationship, one of them being morphology. When a set of phonological features is used to differentiate words, it is contrastive. However, its differentiating function could be contracted to some extent when the same set of features is applied to a word pair with the same morphological root. As a result, those phonological features used in morphological alternations have an intermediate status between contrastive and non-contrastive.

Therefore, when the rebus principle extends its application from homophones to near-homophones, those features used in morphological alternations are the most suitable candidates because they are less contrastive than other phonological contrasts which are not used in morphological alternations.

The monosyllabicization of Old Chinese as a precondition for *rebus*

In this chapter, we will discuss the precondition for rebus in ACWS, focusing on the case of *qīngzhuó*, one of the most popular morphological devices and also a phonological contrast that does not block rebus. Table 3 lists two pairs of words which differ only in terms of *qīngzhuó* categories.

---

[11] See Sagart (2003), Phua (2004), Handel (2012), Jacques (2018), among others.

[12] This alternation is also known as the "*qù* tone alternation" since the derived forms are almost always Middle Chinese *qù* tone. See Downer (1959), Mei (1980), Jacques (2016), among others.

[13] See Yu (1999) for the 3rd-*deng* and non-3rd-*deng* alternation, and Sagart (1999) for the reconstructed *-r-* infix for the 2nd-*deng* derivation from non-2nd-*deng*.

Table 3. Reconstruction of the qīngzhuó contrast in Old Chinese

	敗	敗	伯	白
	defeat(v.t.)	suffer defeat	eldest	white
Middle Chinese	paejH	baejH	paek	baek
Old Chinese (Zhengzhang 2003)	praads	braads	praag	braag
Old Chinese (Baxter & Sagart 2014)	pˤa[t]-s	N-pˤa[t]-s	pˤrak	bˤrak

As shown in Table 3, while *qīngzhuó* categories are reconstructed as voicing contrasts in Middle Chinese, scholars have different opinions on their reconstructed forms in Old Chinese. Some scholars, represented by Zhengzhang (2003), project the voicing contrast in Middle Chinese directly back to Old Chinese. Other scholars, represented by Baxter & Sagart (2014), propose two origins for the Middle Chinese voicing contrast (henceforth called two-origins proposal). For those pairs involved in morphological alternations, they reconstruct a prefix *N- as the source of voiced consonants, while for those pairs that were not involve in such alternations, they project the voicing contrast back to Old Chinese[14].

In terms of writing, the *rebus* method could be applied to both types of pairs regardless of whether they are morphologically related. For example, the first two words meaning ˈdefeat (v.t.)ˈ and ˈsuffer defeatˈ in Table 3 are represented by the same character 敗, while the second two, meaning ˈeldestˈ and ˈwhiteˈ, are both represented by the same character 白 in ACWS. If the proposal about phonological relationships outlined in Chapter 4 is on the right track, then, when reconstructing Old Chinese, we must maintain a similar phonological relationship as reflected in Middle Chinese regardless of their morphological relations. Therefore, our hypothesis favors the first approach represented by Zhengzhang (2003).

However, this does not mean that the two-origins proposal by Baxter & Sagart (2014) is wrong, as it is supported by comparative evidence[15]. In fact, both proposals in Table 3 could be right, if we regard them as reflecting different evolutionary stages of the Chinese language: the two-origins proposal could be true for Proto-Chinese, when the ancestor of the Chinese languages first split from other Tibeto-Burman languages, while the two origins could have already merged into one in Old Chinese, when the writing system was invented. In other words, although the ultimate origin of the *qīngzhuó* alternation was prefixation, it had already become a stem alternation at the stage of Old Chinese.

This brings us to the problem of reconstructing the word-template in Old Chinese. Scholars have different views on this issue: many Chinese scholars reconstruct a monosyllabic word-template for Old Chinese (see Ding 2002 for a review), while Baxter & Sagart (2014) reconstruct the word-template shown in Figure 4.

Baxter & Sagart (2014) regard the word template in Figure 4 as iambic-disyllable[16]. Following Pittayaporn's (2015) definition of sesquisyllable,[17] we classify it as sesquisyllable rather

---

[14] An alternative hypothesis which reconstructs an *s- prefix as the source of voiceless consonants has also been proposed (see Handel 2003 for a review). Nevertheless, the details of reconstruction are not relevant here. The alternative hypothesis may be viewed as a variant of the "two-origins proposal", since its proponents reconstruct different phonological contrasts for pairs that were involved in morphological alternation and for those that were not.

[15] See Sagart (2003) for details. Also, see Phua (2004) for typological concerns on reconstructing prefixation rather than stem alternation.

[16] See Brunelle & Pittayaporn (2012) for further details of this term.

[17] Pittayaporn (2015) defines "sesquisyllable" as "a prosodic word consisting of a full stressed syllable preceded by a consonant or a sequence of consonants. The consonant or consonant sequence must not contain a phonemically contrastive vowel."

than iambic-disyllable. From a diachronic perspective of *monosyllabicization*, a word-template evolution widely attested in Asian languages (Michaud 2012), sesquisyllable is a common intermediate stage from iambic-disyllable to monosyllable, as discussed in detail in Brunelle & Pittayaporn (2012), shown in Figure 5.

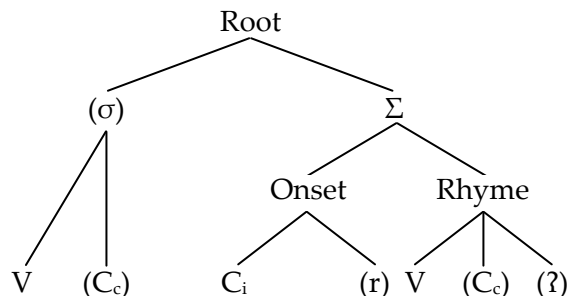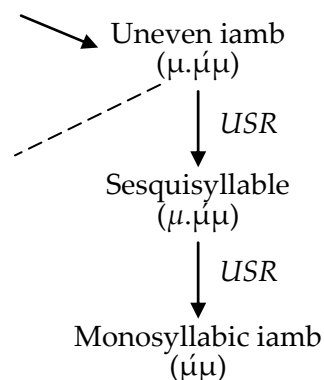*Figure 4.* Hypothetical Old Chinese word-template (Baxter & Sagart 2014: 53)

Figure 5. Word-type shifts (uneven iamb>sesquisyllable>monosyllabic iamb)[18]

If we accept that Chinese also went through the monosyllabicization process,[19] then, the question of reconstructing the word-template in Old Chinese is not a matter of right and wrong. Rather, the question should be reframed as: At which stage in the evolution process was Old Chinese?

We have argued that the application of the *rebus* method in ACWS could be better explained in terms of a language system where prefixation had already changed to stem alternation. This is very likely to be a byproduct of the monosyllabicization process outline in Figure 5, where the minor syllables in the iambic-disyllable template became pre-initials in the sesquisyllabic template, or consonant clusters in the monosyllabic template. As a result of these word-type shifts, morphological alternations originally derived from prefixation gradually became less and less productive, and finally remained only in fossilized forms[20].

This hypothesis is also supported by a recent proposal on the role that language contact must have played in the formation of the Chinese language. It has long been recognized that

---

[18] This figure is a part of the original figure in Brunelle & Pittayaporn (2012: 424), where they emphasize the important role of rhythmic effects on word-type shifts.

[19] Salmons & Zhuang (2018: 556) proposed a hypothetical evolution cycle from Proto-Chinese to Modern Chinese. They regard the reconstructed iambic-disyllable template as the initial stage of Proto-Chinese, and argue that the Chinese language changed to a monosyllabic template through an intermediate stage of C(C)V(C), before returning to the disyllabic template again in Modern Chinese.

[20] A similar case could be found in the fossilization of the suffix *-n* in modern Wu dialects, see Fang (1993) and references therein.

the division between Northern and Southern Chinese dialects and the typological shift from Old Chinese to modern Chinese is largely due to language contact between Chinese and various neighboring language families (Hashimoto 1978, Mei 1997). Recent studies suggest that the origin of the Chinese language may also be a result of language contact. Delancey (2013) argues that "the language of Shang was a highly-creolized lingua franca based on languages of the Southeast Asian type". Although the exact nature and chronology of the language contacts that shaped Old Chinese require further study, the scenario is very much compatible with our hypothesis that the linguistic innovation of Old Chinese provided the basis for the wide application of *rebus* (the rebus principle), which ultimately gave rise to the birth of the Ancient Chinese Writing System.

Implication and future direction

In this paper I put forward a hypothesis on the co-evolution of the Old Chinese language and its writing system, ACWS. I have argued that the invention of ACWS bears a strong correlation with the linguistic evolution, more specifically monosyllabicization, of Old Chinese. In other words, ACWS may have never been invented if Chinese had not gone through monosyllabicization.

If the hypothesis we have proposed is convincing, it will have a significant consequence for the study of Old Chinese. To understand better the evolution of Chinese languages, it seems necessary to distinguish two different stages. One is Proto-Chinese or Pre-Old-Chinese, which represents a stage that predates the invention of Chinese writing. The other is Old Chinese, which represents a stage when the writing system has already been invented. The evidence for reconstructing Proto-Chinese will mainly come from comparison with other Sino-Tibetan languages, while the evidence for reconstructing Old Chinese will mainly be based on excavated texts, including writing practice, rhyming patterns, and the formation of word families.

Regarding the direction of possible future research, we would like to explore the similarities and dissimilarities between independently invented writing systems and seek for potential parallels that might show a similar nature to the origin of the Ancient Chinese Writing System.

References

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

Baxter, William H., Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford: Oxford University Press.

Boltz, William G. 1994. *The origin and early development of the Chinese writing system*. Indiana: Eisenbrauns.

Brunelle, Marc, Pittayawat Pittayaporn. 2012. Phonologically-constrained change: The role of the foot in monosyllabization and rhythmic shifts in Mainland Southeast Asia. *Diachronica* 29(4): 411–433.

Chen, Jian (陳劍). 2007. Yinxu buci de fenqi fenlei dui jiagu wenzi kaoshi de zhongyaoxing (殷墟卜辭的分期分類對甲骨文字考釋的重要性). *Jiagu jinwen kaoshi lunji* (甲骨金文考釋論集): 317–453. Beijing: Xianzhuang shuju (綫裝書局).

DeLancey, Scott. 2013. The origins of Sinitic. In: Zhuo Jing-Schmidt (ed.). *Increased Empiricism: New Advances in Chinese Linguistics*: 73–99. Amsterdam: John Benjamins Publishing.

Ding, Bangxin (丁邦新). 2002. Shangguhanyu de gouci wenti shuping (上古汉语的构词问题述评). *Yuyanxue luncong* (语言学论丛) 26: 1–11.

Downer, Gordon B. 1959. Derivation by tone-change in Classical Chinese. *Bulletin of the School of Oriental and African Studies* 22(2): 258–290.

Dresher, B. Elan, Glyne Piggott, Keren Rice. 1994. Contrast in phonology: Overview. *Toronto Working Papers in Linguistics* 13(1): iii–xvii.

Fang, Songxi (方松熹). 1993. Zhejiang wufangyan li de erwei (浙江吴方言里的儿尾). *Zhongguo yuwen* (中国语文) 2: 134–139.

Guo, Moruo (郭沫若), Hu Houxuan (胡厚宣) (eds.). 1982. *Jiaguwen heji* (甲骨文合集). Beijing: Zhonghua shuju (中华书局).

Hall, Kathleen C. 2009. *A probabilistic model of phonological relationships from contrast to allophony*. Doctoral dissertation. The Ohio State University.

Halliday, Michael A. 1981. The origin and early development of Chinese phonological theory. In: R. E. Asher, Eugenie J. A. Henderson (eds.). *Towards a history of Phonetics*: 123-140. Edinburgh University Press.

Handel, Zev. 2012. Valence-changing prefixes and voicing alternation in Old Chinese and Proto-Sino-Tibetan: Reconstructing *s- and *N-prefixes. *Language and Linguistics 13*: 61–82.

Hashimoto, Mantaro J. 1978. Language Typology and Geography. *Tokyo: Kobundo*.

Jacques, Guillaume. 2016. How many *-s suffixes in Old Chinese? *Bulletin of Chinese linguistics* 9(2): 205–217.

Jacques, Guillaume. 2018. *Fossil normalization prefixes in Tibetan and Chinese*. Paper presented at the 2nd Li Fang-Kuei Society Young Scholars Symposium, Institute of Linguistics, Academia Sinica, Taiwan, 12–15, Jul. 2018.

Keightley, David N. 1985. *Sources of Shang history: The oracle-bone inscriptions of Bronze Age China*. Oakland: University of California Press.

Ladefoged, Peter. 2005. *Vowels and consonants: An Introduction to the Sounds of Languages (second edition)*. Oxford: Blackwell.

LaPolla, Randy J. 2003. Overview of Sino-Tibetan morphosyntax. In: Randy J. LaPolla, Graham Thurgood (eds.). *The Sino-Tibetan Languages*: 22–42. London / New York: Routledge.

Li, Xiaoding (李孝定). 1974. Jiaguwen de liushufenxi (甲骨文的六書分析). *Bulletin of the Institute of History and Philology* (中央研究院歷史語言研究所集刊) 45(2): 374–380.

List, Johann-Mattis, Anselm Terhalle, Daniel Schulzek. 2016. Traces of embodiment in Chinese character formation. A frame approach to the interaction of writing, speaking, and meaning. In: Ströbel, L. (ed.). *Sensory-motor concepts. At the crossroad between language & cognition*: 45-62. Düsseldorf University Press.

Mei, Zulin (梅祖麟). 1980. Sishengbieyi zhong de shijian cengci (四声别义中的时间层次). *Zhongguo yuwen* (中国语文) 6: 427–443.

Mei, Zulin (梅祖麟). 1997. Hanyu qige leixingtezheng de laiyuan (汉语七个类型特征的来源). *Zhongguo jingnei yuyan ji yuyanxue* (中国境内语言暨语言学) 4: 81–103.

Michaud, Alexis. 2012. Monosyllabicization: Patterns of evolution in Asian languages. In: Thomas Stolz, Nicole Nau, Cornelia Stroh (eds.). *Monosyllables: from phonology to typology*: 115–130. Berlin: De Gruyter Mouton.

O'Seaghdha, Padraig G., Jenn-Yeu Chen, Train-Min Chen. 2010. Proximate units in word production: Phonological encoding begins with syllables in Mandarin Chinese but with segments in English. *Cognition* 115(2): 282–302.

Pittayawat Pittayaporn. 2015. Typologizing sesquisyllabicity. In: N. J. Enfield, Bernard Comrie (eds.). *Languages of Mainland Southeast Asia: The state of the art*: 500-528. Berlin: De Gruyter Mouton.

Phua, Chiew Pheng (潘秋平). 2004. Shangguhanyu zhong de qingzhuobieyi jiqi xiangguan de wenti (上古漢語中的清濁別義及其相關問題). In: *Dingbangxin jiaoshou rongxiu jinian lunwenji* (丁邦新教授榮休紀念論文集): 1–63. Hong Kong: Center for Chinese Linguistics, Hong Kong University of Science and Technology.

Qiu, Xigui (裘锡圭). 1985. Shi yinxujiaguwen li de yuan er ji youguanzhuzi (释殷墟甲骨文里的远迩及有关诸字). Beijing: Zhonghua shuju (中华书局).

Qiu, Xigui (裘锡圭). 1993. Shi yinxubuci zhong de ming dengzi (释殷虚卜辞中的 皿等字). In: *Di erjie guoji zhongguo guwenzixue yantaohui lunwenji* (第二届国际中国古文字学研讨会论文集). Hong Kong: Xianggang zhongwendaxue zhongguoyuyanjiwenxuexi (香港中文大学中国语言及文学系).

Qiu, Xigui (裘锡圭). 2013. *Wenzixue gaiyao* (文字学概要). Beijing: Shangwu yinshuguan (商务印书馆).

Robertson, John S. 2004. The possibility and actuality of writing. In: Stephen D. Houston (ed.). *The first writing: Script invention as history and process*: 16–38. Cambridge University Press.

Rogers, Henry. 2004. *Writing systems: A linguistic approach*. Blackwell Textbooks In Linguistics.

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

Sagart, Laurent. 2003. Sources of Middle Chinese manner types: Old Chinese prenasalized initials in Hmong-Mien and Sino-Tibetan perspective. *Language and Linguistics* 4(4): 757–768.

Salmons, Joseph, Zhuang Huibin. 2018. The diachrony of East Asian prosodic templates. *Linguistics 56*(3): 549–580.

Xia, Dazhao (夏大兆). 2014. *Jiaguwen ziyong yanjiu* (甲骨文字用研究). Doctoral dissertation. Anhui daxue (安徽大学).

Yu, Min (俞敏). 1999. Guhanyu paisheng xinci de moshi (古汉语派生新词的模式). *Yumin yuyanxue lunwenji* (俞敏语言学论文集). Beijing: Shangwu yinshuguan (商务印书馆).

Zhengzhang, Shangfang (郑张尚芳). 2003. *Shanggu yinxi* (上古音系). Shanghai: Shanghai Educational Publishing House (上海教育出版社).

Шэнь Жуйцин. Моносиллабизация в древнекитайском языке и генезис китайской иероглифики: гипотеза о совместной эволюции китайского языка и китайской письменной системы

В статье выдвигается гипотеза относительно совместной эволюции древнекитайского языка и сложившейся на его основе системы письменности. Утверждается, что изобретениее китайской письменности тесно коррелирует с лингвистической эволюцией китайского языка, в особенности его моносиллабизацией, и что без перехода китайского к односложной структуре слова изобретение такого рода письменности было бы невозможно.

Работа начинается с обсуждения возможных причин того, что для древнекитайского языка не была изобретена алфавитная система письма. Вслед за этим анализируются особенности ребусного принципа в устройстве древнекитайской письменности и его связь с морфологическими чередованиями; приводятся аргументы в пользу того, что обязательным условием для появления ребусного принципа (и вместе с ним собственно древнекитайского письма) является моносиллабизация. В последней части работы обсуждается значимость данной гипотезы в общем контексте изучения древнекитайского языка.

Ключевые слова: древнекитайский язык, китайская письменность, моносиллабизация, ребусный принцип, морфологические чередования.

GONG Xun

School of Oriental and African Studies, University of London; minus273cn@gmail.com

Chinese loans in Old Vietnamese
with a sesquisyllabic phonology [1]

While consonant clusters, taken broadly to include presyllables, are commonly hypothesized for Old Chinese, little direct evidence is available for establishing the early forms of specific words. This essay examines a hitherto overlooked source: Old Vietnamese, a language substantially attested in a single document, which writes certain words, monosyllabic in modern Vietnamese, in an orthography suggesting sesquisyllabic phonology. For a number of words loaned from Chinese, Old Vietnamese provides the only testimony of the form of the Vietic borrowing. The small list of currently known sesquisyllabic words of Chinese origin attested in this document includes examples of both words with a secure initial Chinese cluster and words with plausible Vietic-internal prefixation.

Keywords: Old Chinese language, Old Vietnamese language, historical reconstruction, sesquisyllabic words, prefixal morphology.

1. Context

Old Vietnamese[2] is a Vietic language of which substantial attestation is limited to a single document, the 佛說大報父母恩重經 *Phật thuyết Đại báo phụ mẫu ân trọng kinh* (*Fóshuō Dàbào fùmǔ ēnzhòngjīng*, "Sūtra explained by the Buddha on the Great Repayment of the Heavy Debt to Parents", henceforth *Đại báo*). The language of this document preserves Proto-Vietic sesquisyllabic phonology. In Old Vietnamese, we find words of Chinese origin like 𩜹 *s-kương* 'mirror', borrowed from Chinese 鏡 *kiæŋH*.[3] This paper examines whether words like

---

[2] On the term *Old Vietnamese* see §1.3. In this essay, Old Vietnamese is transcribed with reconstructed initial, in roman type, and projected Modern Vietnamese rime in italic type. For example, *plời* 'heaven' denotes that I reconstruct Old Vietnamese initial *pl-, and that the Modern Vietnamese cognate of the word is *trời*.

[3] Chinese (and Sino-Vietnamese) are transcribed in several different ways depending on the context. In proper names and terms of cultural nature, the Chinese is rendered in pinyin, in Sino-Vietnamese or both, depending on the specific context to which the term pertains.

Chinese sources of loanwords under consideration are transcribed in a modified version of Baxter's Middle Chinese transcription (1992). Most notably, the transcription of this paper uses medial *-i-* instead of *-j-*, and the vowel *o* is written *ʌ*. Other sounds are transcribed in an IPA-like fashion: among Baxter's alternative orthographies for vowels, *æ, ɛ, ɨ* are preferred to *ae, ea, +*. Retroflex stops are written *ṭ, ṭh, ḍ, ŋ*, retroflex sibilants *tʂ, tʂh, dʐ, ʂ, ʐ*, palatal sibilants *tɕ, tɕh, dʑ, ɲ, ɕ, ʑ*. Finally, we have *j* for *y*, *ŋ* for *ng*, *ʔ* for *'* and *ɣ* for *h*.

Finally, Chữ Nôm orthography is transcribed first in Sino-Vietnamese followed by Middle Chinese transcription: 破散 (*phá tán* < *phaH sanX*). Unencoded Chữ Nôm characters are represented with Unicode ideographic

this support the reconstruction of initial consonant clusters in Old Chinese. More specifically, it investigates whether Chinese loans in Old Vietnamese that exhibit sesquisyllabic phonology confirm or challenge Baxter and Sagart's reconstruction of Old Chinese (2014), which employs data from Vietic languages (notably Vietnamese and Rục) to support Old Chinese clusters.

This section introduces the context of the paper, beginning with a general introduction (§1.1) to the difficulties of reconstructing Old Chinese initial consonant clusters. §1.2 discusses the importance of Vietic languages, which offer an important source of evidence for them in cases where they might have disappeared without a trace. The source *Đại báo* is introduced in §1.3, and the phonetic interpretation of its sesquisyllabic orthography in §1.4.

1.1. Initial consonant clusters in Old Chinese

While Middle Chinese and contemporary Chinese dialects have a simple syllabic canon, scholars have long hypothesized that Old Chinese has a more complex phonotactics, especially on the left periphery of the syllable. Scholars from Henri Maspero (1930) reconstruct initial consonant clusters like *pr- or *sm-, while recent reconstructions (Pan Wuyun 2000, Baxter and Sagart 2014) also hypothesize a sesquisyllabic phonology with initial minor syllables. Baxter and Sagart (2014), for example, reconstruct Old Chinese forms such as *k.teʔ for 紙 *tɕeX "paper" and *mə.lat for 舌 *ʑet "tongue".

In this paper, *par abus de langage*, we call both initial clusters like *kt- and the sequence of an initial minor syllable followed by the initial consonant like *məl- 'consonant clusters'. I follow the convention of Baxter and Sagart (2014) in referring to the first components, such as *k- or *mə-, as *preinitials*, and distinguish them as consonantal (*k-) and syllabic (*mə-).

While various kinds of evidence point to consonant clusters in Old Chinese, it is very difficult for scholars to agree on the clusters to reconstruct for precise etymons. One of the reasons for this uncertainty is that, under the type of evolution that languages of China and Mainland Southeast Asia usually undergo, different phonotactic types of clusters have almost intrinsically different rates of survival in descendants. A case in point is Old Tibetan, in which most consonants can take preinitials *s-* and *d-*. However, the existence and identity of preinitials is much better preserved before sonorants than stops: in a typical modern Kham dialect (for example the dialects of Derge and Batang, cf. Skal·bzang 'Gyur·med and Skal·bzang Dbyangs·can, 2002), Old Tibetan *ng-*, *sng-* and *dng-* remain distinct as _ŋ-[4], ‾ŋ̊- and ‾ŋ-, but *k-*, *sk-* and *dk-* all merged into ‾k-. If one were reconstructing Old Tibetan without the benefit of orthographic forms, it is likely that *ŋ-, *sŋ- and *dŋ- would be reconstructed correctly, but *k-, *sk- and *dk- would be very hard to distinguish.

This asymmetry in the preservation of clusters before obstruents and sonorants explains the curious situation of Old Chinese reconstructions, where there is a clear gradient to the levels of consensus among different reconstruction of different phonotactic types of OC clusters. For example, recent reconstructions of Old Chinese agree on either one of two theories about *sm-* type initial consonants (see Mei 2012, Sagart & Baxter 2012). Given that there is an implicational hierarchy to the effect that languages with *sm-* type initial consonants usually have *sk-* type initial consonants (Goad 2011), Old Chinese likely had *sk-* type initial consonants as well. However, authors cannot agree on which particular OC words have *sk-* type consonants (see Gong & Lai 2017 for a brief account).

---

description characters such as ⿰ or ⿱. For example, the sequence "⿰目它" represents a character composed horizontally of 目 and 它.

    [4] ‾ designates the high tone, and _ the low tone, in tonal modern dialects of Tibetan.

In order to better understand both the phonological structure of Old Chinese and the reconstruction of individual Old Chinese words, we need to search for more direct evidence of consonant clusters.

1.2. Does Vietic evidence support Old Chinese clusters?

From the very beginning of research on Vietnamese words of Chinese origin, it has been noted that certain words borrowed from Chinese exhibit lenition of initial consonants (Maspero 1912: 19–39, Wáng Lì 1948: 71). An example is Chinese 劍 *kiæmH 'sword', which is borrowed as Vietnamese 劍 gươm [ɣ-]. This contrasts with the absence of initial lenition in the Sino-Vietnamese pronunciation of the word, kiếm.

Vietnamese belongs to the Vietic branch of Austroasiatic languages. Modern Vietic language exhibit the whole typological spectrum from monosyllabic Chinese-like Việt-Mường languages[5] to completely sesquisyllabic languages such as the Chứt (Rục-Sách) cluster. Haudricourt (1965) first explained Vietnamese spirant initials like v-, d- [z- < ð-], g- [ɣ-] and r- as reflecting a process of lenition caused by lost Proto-Vietic presyllables. According to this theory (cf. also Thompson 1976: 1131–1133, Ferlus 1976, 1982), Proto-Vietic presyllables, still preserved almost intact in conservative Vietic languages like Rục, disappeared in all modern Việt-Mường languages. Their former presence triggered lenition in Mainstream (Hanoi-Saigon) Vietnamese, but not in Mường varieties nor the "Haut-Annam" (Maspero 1912) or "heterodox" dialects of Vietnamese, such as Vinh (Ferlus 1991) or Quảng Bình (Michaud, Ferlus and Nguyễn 2015).

Table 1: Lenition in Vietnamese

Proto-Vietic	Vietnamese	Mường[4]	Rục[6]
*p- 'four'	bốn ([ɓ-])	pon˧˧	pó:n
*CVp- 'lime'	vôi	pol˧˧	kəpu:l
*s- 'hand'	tay	thăj˧˧	si:
*CVs- 'snake'	rắn	thăɲ˧˧	pəsí:ɲ

As predicted by this theory of the origin of Vietnamese lenition, for Chinese words just as for native words, conservative Vietic languages have a cognate with a prefix: Vietnamese 劍 gươm [ɣ-] 'sword' is cognate to Rục təkɨəm. Pulleyblank (1981: 281–286), the first focused treatment of Chinese loans in Vietnamese showing lenition, considers the possibility that the preinitial was present in the original Chinese form 'of great interest' but 'not easy to pursue further without additional information to enable one to determine the kind of clusters involved'.[7]

---

[5] This essay follows the terminology of Hayes (1992), now in general usage. *Vietic* designates the larger group, which includes notably Chứt (Rục-Sách). *Việt-Mường*, a term which formerly often comprises all Vietic languages, designates only the languages spoken by the ethnic Vietnamese and Mường groups. For the terminological questions, cf. Michaud, Ferlus and Nguyễn (2015: 126).

[6] Rục forms are cited from Nguyễn, Trần and Ferlus (1988). 'Mường' designate Khen Mường (Hòa Bình province), transcribed by Milton and Muriel Barker, and cited from Thompson (1976).

[7] Maspero (1912: 21-23 et passim) deems Vietnamese words with lenition to be a 'dérivé récemment formé' of their counterparts without lenition in Mường and Vinh-type Vietnamese dialects. Wáng Lì (1948) includes lenition in what he calls the process of *Vietnamization* (越化, Việt hoá), which Nguyễn Tài Cẩn (1979) defines as "development under a different path than that of Sino-Vietnamese readings" (diễn biến theo một con đường khác với cách đọc Hán Việt).

Baxter and Sagart (2014) were the first to systematically use loanword material in Vietic languages to reconstruct consonant clusters in Old Chinese. By comparing the Vietic data with Chinese loans in the Kra-Dai language Lakkia, Baxter and Sagart (2014: 36–37, 93–97) give a compelling argument that preinitials in these words do come from the Chinese source. Their examples are 紙 *tɕeX 'paper', borrowed as Vietnamese *giấy* [z-], Rục *kəcáj*, Lakkia *khjei*[3], and 賊 *dzʌk* 'bandit', borrowed as Vietnamese *giặc* [z-], Rục *kəcʌ́k*, Lakkia *kjak*[8]. The oldest sizeable set of Chinese loans in Vietic and Lakkia, just as their counterparts in Tai and Kam-Sui, all result from the same event of massive borrowing which follows the Qín-Hàn conquest of the region corresponding to modern Guǎngdōng, Guǎngxī and Vietnam. For a Chinese word to receive a Vietic prefix and be then borrowed into Lakkia would be extremely implausible. Even if the number of correspondences is limited, we have here an important argument for the Chinese origin of the preinitials.

In the system of Baxter and Sagart (2014), Vietic preinitials are understood as reflecting genuine Old Chinese preinitials, thus 紙 *tɕeX 'paper', Vietic *k- < *k.teʔ; 賊 *dzʌk 'bandit', Vietic *k- < *k.dzˤək.

While there are a great number of words of Chinese origin with a softened initial in Vietnamese, only a few of them can be found in conservative Vietic languages as Rục. In order to further evaluate Baxter and Sagart's use of Vietic data in Old Chinese reconstruction, we need additional sources that attest to Chinese loanwords with preinitials.

1.3. Old Vietnamese and *Đại báo*

We find many cases where a word of Chinese origin shows a lenited modern reflex in Vietnamese, but is not attested in conservative Vietic languages such as Rục and Thavưng. We know that the Proto-Việt-Mường form must have contained a preinitial, but it is unclear which.

The preponderance of older Chinese loans in Vietnamese and Viêt-Mường but not other Vietic languages is not surprising given the linguistic history of Vietic. As shown by Ferlus' work (2010) on the Khmer names of the duodecimal year cycle (地支 *dìzhī*, *Địa Chi* 'earthly branches'), the Vietic branch of Austroasiatic had already split into subbranches by the time the early layer of Chinese words arrived. Chinese words entered the ancestor of today's Việt-Mường, and were then borrowed into other languages. It is not surprising that lexical Sinicization was followed by phonological Sinicization: the group that originally received the Chinese loans, which also contains the largest number of them, was then monosyllabicized under the weight of Chinese influence. Middle Vietnamese, as attested in Alexandre de Rhodes' dictionary (1651), had already lost all sesquisyllabic preinitials, and retained only a few relictual initial clusters such as *bl-*.

This paper continues Baxter and Sagart's work (2014) on Vietic borrowings in Old Chinese by examining a remarkable document, which greatly improves our knowledge of Việt-Mường historical phonology. The document, 佛說大報父母恩重經 *Phật thuyết Đại báo phụ mẫu ân trọng kinh* ("Sūtra explained by the Buddha on the Great Repayment of the Heavy Debt to Parents", henceforth *Đại báo*), is held in the *Société asiatique*, Paris. It is a version of a popular Chinese apochyphon more commonly known under the title 父母恩重難報經 *Fùmǔ Ēnzhòng Nánbàojīng*, *Phụ mẫu ân trọng nan báo kinh* ("Sūtra on the Difficulty of Repaying the Heavy Debt to Parents"), in which the Chinese text is accompanied by a vernacular translation (called 解音

Maspero, Wáng Lì and Nguyễn Tài Cẩn essentially postulate a non-Neogrammarian unconditional split, the condition of which is convincingly explained by Haudricourt, Thompson and Ferlus as Proto-Vietic presyllables. As a terminological convenience, we can understand *Vietnamization* as morphological prefixation within Proto-Vietic or Proto-Việt-Mường, which developed into Vietnamese lenition.

giải âm in Vietnam) in a rudimentary form of Chữ Nôm, where vernacular words are written with Chinese characters and modified versions thereof.

The language of the vernacular translation in *Đại báo* is clearly Việt-Mường. I propose to call this language Old Vietnamese, as the language reflected in *Đại báo* shows an important archaism: certain words that are monosyllabic in Modern Vietnamese are written with two Chinese characters (*digraphic orthography*), the first of which corresponding to a preinitial in conservative Vietic languages and proto-Vietic: *rắn* 'snake', Middle Vietnamese *rắn* '*cobra; anguis*' (Rhodes 1651: 636), is written 破散 (*phá tản < *phaH sanX*).

The traditional periodization schemes of Vietnamese distinguish a period *Archaic Vietnamese* (Nguyễn Dình-Hoà, 2009) or *Old Vietnamese* (*vietnamien ancien*, Ferlus, 2010). I believe that calling the language of the *Đại báo* is justified by the fact that there is a fundamental divide between, on the one hand, Archaic Vietnamese as traditionally defined as well as the language of *Đại báo*, and, on the other hand, Middle Vietnamese of Ānnán Yìyǔ (安南譯語, *An Nam Dịch Ngũ*) and Rhodes (1651). For example, 'snake' was borrowed from Archaic Vietnamese to Old Khmer as *masāñ* (Ferlus 2010: 9). The Archaic Vietnamese form clearly sides with the form in the *Đái Báo* **p-sắn* 破散, and is distinguished from Middle and Modern Vietnamese *rắn*. Old Vietnamese, as defined in this essay, can be regarded as the last stage, uniquely attested, of Archaic/Old Vietnamese as hitherto defined in the literature.

Apart from the *Đại báo*, an extensive text, Old Vietnamese is attested in a much smaller scale in two sources: remnants in Modern Chữ Nôm usage (cf. §3.1) and proper names attested in Chinese-language inscriptions, most importantly the Hộ Thành Mountain (護城山, Ninh Bình province) inscription (Shimizu, Lê & Momoki 2005).

Đại báo was first brought to scholarly attention when a copy of it was sent to the Hán Nôm Institute in 1979; Nguyễn Ngọc San (1982) wrote about the digraphic orthography of *Đại báo*, which he interpreted as clusters and preglottalized consonants. Shimizu Masaaki (1996) made the first systematic study of *Đại báo*'s digraphic orthography, in which he collected a corpus of 24 glyph-word pairs exhibiting cluster orthography, as well as 47 glyph-word pairs exhibiting sesquisyllabic orthography. Hoàng Thị Ngọ's candidate thesis (1996), later published as Hoàng (1999), is the first and only transcription of the text into modern Vietnamese.

Đại báo shows a great number of cases of digraphic orthography, where a word, monosyllabic in Modern Vietnamese, is written with two Chinese characters (*two-character orthography*), or in a composite character made of two different characters (*composite-character orthography*). Following Shimizu (1996), we classify Old Vietnamese digraphic orthography into two phonotactic classes:

- Consonant-liquid cluster orthography: The word transcribed is strictly monosyllabic and has a consonant-liquid initial cluster: CRVC; the first character transcribes the initial consonant C; the second character transcribes the medial RVC.

  For example, the word *trời* 'heaven', MViet *blời* '*ceo; cælum*' (Rhodes 1651: 45), is written (巴+例 *ba lệ < *pa lieiH*) in *Đại báo*. I reconstruct Old Vietnamese **plời*, the first character 巴 *ba < pa* transcribes the initial consonant **p*, the character 例 *lệ < *lieiH* transcribes the remaining part of the syllable **lời*.

- Sesquisyllabic orthography[8]: the word transcribed is sesquisyllabic C-CVC. The first character transcribes the presyllable C-; the second character transcribes the initial and the rime CVC.

---

[8] This type is called *disyllabic construction* (双音節構造 *sō-onsetsu kōzō*) in Shimizu (1996).

> For example, the word *rắn* 'snake', MViet *rắn* '*cobra; anguis*' (Rhodes 1651: 636), is written 破散 (*phá tản* < **pha*H *san*X). I reconstruct Old Vietnamese **p-sắn*, cf. Rục *pəsí:ɲ*. The first character 破 *phá* < **pha*H transcribes the preinitial **p-*, and the second character 散 *tản* < **san*X transcribes the remaining part **sắn*.

In this study, we focus our attention on the second kind of digraphic orthography, which relies on Vietic and Old Vietnamese sources. The first kind of digraphic orthography, which transcribe consonant clusters of the TR- type, will be deferred to another discussion along with other sources of evidence.

Concerning the dating of the *Đại Báo*, the document held in the *Société asiatique* itself was printed at the initiative of Trịnh Quán 鄭檊 around 1730 (Shimizu 1996: 84). The text conspicuously avoids the character 利 *lợi*, the name of the first emperor (c. 1384–1433) of the Lê dynasty (1428–1788). Mainly based on this taboo, Shimizu (1996, 2015) dates the text to the beginning of the Lê dynasty (XVth century). On the other hand, Nguyễn Tài Cẩn (2008) suggests that the text is likely to be copied from a Trần-dynasty (1225–1400) precedent, by pointing to a number of less conspicuous taboo practices in the text[9].

A number of texts are dated, with more or less certainty, to the Trần dynasty. Trần Trọng Dương (2011) pointed out that, compared to known texts dated to the Trần dynasty, *Đại Báo* contains at least an order of magnitude more cases of two-character orthography. For example, compared with the text 指南玉音解義 *Chỉ nam ngọc âm giải nghĩa*, which shows the most cases of orthography in two characters apart from the *Đại Báo*, there are 16 cases of two-character orthography for a total length of 15000 characters; in the *Đại Báo*, on the other hand, Trần Trọng Dương counts 103 cases of two-character orthography for a total length of 4942 characters. Independently of Trần Trọng Dương's work, I checked the list of words showing sesquisyllabic orthography in Early Vietnamese texts given in Trần and Nguyễn (2007). Among the 62 examples, 50 are attested only in *Đại báo*. The result also shows that there is a qualitative difference between the language of *Đại báo* and that of later texts. As Nguyễn Quang Hồng (2008: 127–144) and Trần Trọng Dương (2011) observe, the text should be dated before Trần dynasty, likely to XIIth century.

1.4. The reconstruction of Old Vietnamese preinitials in this paper

The reconstruction of Old Vietnamese preinitials in this paper roughly follows that of Shimizu (1996). I reconstruct the following Old Vietnamese preinitials: **p-*, **t-*, **k-*, **ɕ-*, **s-*, *ʔ-*. We note that the list of Old Vietnamese preinitials coincides with the more numerous preinitials in Michel Ferlus's reconstruction of Proto-Vietic (2007): **p-*, **t-*, **k-*, **c-*, **s-*, **ʔa*. In the following list, I provide the spellers and corresponding Proto-Vietic initial for each Old Vietnamese preinitial.

Old Vietnamese **p-* < Proto-Vietic **p-*
   Spellers: 波 (*ba* < **pa*), 巴 (*ba* < **pæ*), 破 (*phá* < **pha*H)
   Example: **p-sắn* 'snake' 破散 (*phá tản* < **pha*H *san*X) = *rắn*, Middle Vietnamese *rắn* '*cobra; anguis*' (Rhodes 1651: 636)
   Vietic cognates: Rục *pəsí:ɲ* 'snake'

Old Vietnamese **t-*, cf. Proto-Vietic **t-*
   Speller: 多 (*đa* < **ta*)

---

9   Trần Trọng Dương (2010) disagrees with Nguyễn Tài Cẩn's analysis, and takes the taboo characters under question to be normal graph variants at the period.

Example: *t-m*ai* 'shoulder' 多埋 (*đa mai* < *ta mɛj*) = *vai*, Middle Vietnamese *ƀai* 'ombros: humerus' (Rhodes 1651: 65)

Vietic cognates: Thavưng *ʔapî:ɲ / ʔapî:ŋ* 'shoulder'

Old Vietnamese *k- < Proto-Vietic *k-

Spellers: 可 (*khả* < *khaX*), 丁 (*cá* < *ka*H)[10]

Example: *k-ɲớ* 'remember' 可汝 (*khả nhũ* < *khaX ɲʌX*) = *nhớ*, Middle Vietnamese *đớ* or *nhớ* 'lembrarse; recordor' (Rhodes 1651: 175)

Vietic cognates: Rục *kəɲớ:* 'to remember'

Old Vietnamese *ɕ- < Proto-Vietic *c-

Spellers: 車 (*xa* < *tɕhæ*), 舍 (*xả* < *ɕæX*)

Example: *ɕ-mãng* 'hear' 車莽 (*xa mãng* < *tɕhæ maŋX*), 舍莽 (*xả mãng* < *ɕæX maŋX*) = *mắng*, obsolete in Modern Vietnamese, but cf. Kiều l.535 嗻信掣浽驚惶 *Mắng tin xiết nỗi kinh hoàng* 'How he was scared after hearing the news!' Middle Vietnamese cf. *mắng tin* 'fama; fama' (Rhodes 1651: 450)

Vietic cognates: Rục *camắɲ* 'to hear, listen'

Note: The Chinese character 車 has two alternative readings: MC *kiʌ* = Mandarin *jū*, Sino-Vietnamese *cư*; MC *tɕhæ* = Mandarin *chē*, Sino-Vietnamese *xa*. Nguyễn Hữu Vinh et al. (2009: 764) and Trần Trọng Dương (2012) read 車 as *cư* < *kiʌ*, reconstructing *k-mắng*. This reading is to be rejected considering the alternative orthography and Vietic cognates.

Old Vietnamese *s-, cf. Proto-Vietic *s-

Speller: 司 (*tư* < *si*)[11]

Example: *s-p*ui* 'merry, joyful' (司+盃 *tư bôi* < *si pwʌi*) = Vietnamese *vui*, Middle Vietnamese *ƀui* 'allegre; hilaris' (Rhodes 1651: 74)

Vietic cognates: Rục *tupu:j* 'merry, happy'

Old Vietnamese *ʔ-, comparable to Proto-Vietic *ʔa

Speller: 阿 (*a* < *a*)

Example: *a-p*ội* 'early' 阿盃 (*a bôi* < *a pwʌi*) = Vietnamese *vội* 'hasty', Middle Vietnamese *ƀội* 'cousa apressada; properus' (Rhodes 1651: 71).

(No known cognates in conservative Vietic languages)

Shimizu (1996) reconstructs *ɓ- (or *p-), *ph-, *ɗ- (or *t-), *l-, *ś-, *kh-, *s- and *ʔ-. My reconstruction of Old Vietnamese preinitials differs from his treatment in the following respects:

- Voiceless unaspirated stops, akin to other Vietic languages and reconstructed Proto-Vietic, are preferred to implosives.
- I do not admit the difference between *p- and *ph-. Although two different sets of spellers are used with bilabial stops, with unaspirated stops (波 *ba* < *pa* and 巴 *ba* < *pæ*) and with aspirated stops (破 *phá* < *pha*H), the same behaviour is seen in *k-, with

---

[10] This character, a simplified form of 箇 (*cá* < *ka*H) 'one', is a graphical variant of other simplified forms of 箇 in currency in East Asia, such as Simplified Chinese 个 (*gè*) or Japanese ヶ (*ka*). All derive from one half of the bamboo component ⺮ in 箇.

[11] Nguyễn Ngọc San (1982) and Trần Trọng Dương (p.c.) propose that the speller 司 (*tư* < *si*) should be read as *t-* instead of *s-*. This is unlikely, since the sound change *t* < *s* in syllable initial did not happen before lenition, as *đ* < *t* gives Middle Vietnamese *d* [ð] in lenition, but *t* < *s* gives *r*.

unaspirated speller 丁 (*cá* < **ka*H) as well as aspirated speller 可 (*khả* < **kha*X). In both cases, there is no pressing evidence to support the distinction, which does not conform to usual Austroasiatic phonological patterns, as in those of Khmer or reconstructed proto-Vietic.

- I do not consider the word 扎打 (*la đả* < **la tæŋ*X) 'rock', which would be **l-tá* according to Shimizu (1996) and other studies, as a case of an Old Vietnamese sesquisyllabic word. It does not cause lenition in Modern Vietnamese: *đá* not **dá*; it survived in forms of Vietnamese later than Old Vietnamese: we find 羅硰 (first character *la* < **la*, second character Nôm with 多 *đa* < **ta* as phonetic component) in 國音詩集 *Quốc âm thi tập*; more strikingly, we find Middle Vietnamese *là đá* '*lagea de pedra*, lapis planus' (Rhodes 1651: 390).

2. Some sesquisyllabic Old Vietnamese words borrowed from Chinese

In this section, I study some words of Chinese origin from *Đại báo*. I collect Old Vietnamese words in two-character orthography from two studies of the document (Shimizu 1996, Trần and Nguyễn 2007), and select the words given in the two studies for which the Chinese origin appears secure to me.

The *quốc ngữ* reading of the text of *Đại báo* is based on Hoàng Thị Ngọ (1999), who provides the only complete transcription so far available. I report all cases where my judgment differs from that of Hoàng Thị Ngọ (1999), Shimizu (1996) or Trần and Nguyễn (2007).

Three examples will be discussed not in this section, but in subsequent sections. Section 3 discusses 䲷 **s-kɯɔŋ* 'mirror', which bears important consequences on several aspects concerning Old Chinese reconstruction in general and Baxter and Sagart's reconstruction (2014) in particular. Section 4 treats two words with Old Vietnamese preinitial **ʔ-*, which could be a case of Vietic-internal prefixation.

2.1. 巴拭 **p-ɕức* 'to spread, to apply' < 拭 **ɕik* 'to wipe'

(1) 巴拭　　仍　　粉
    *p-ɕức　*nhũng*　*phấn*
    spread　PL　　powder
    'apply different kinds of powder' (Đại báo 8a-1)
    Original: （女人在世）濃塗脂粉 '(Women live in this world,) apply themselves thickly with rouge and powder...'

I follow Trần and Nguyễn (2007, #37) in reading 巴拭 (*ba-thức* < **pa-ɕik*) as *xức* 'to anoint, to rub, to apply', Middle Vietnamese *xức* '*untar, ungir; ungo*' (Rhodes 1651: 896). Hoàng Thị Ngọ (1999: 147) has *sức* 'force', likely merely an alternative spelling of the same word.

The Old Vietnamese reconstructs to **p-ɕức*. The Old Vietnamese preinitial **p-* fails to cause lenition in **ɕ* and yields *x-* ([s]) in Modern Vietnamese. This behaviour is probably regular, and observed in **k-ɕa* > *xa* 'far' (§ 2.3) and **ʔ-ɕɯớng* > *xướng* 'chant' (§ 4).

The Chinese original of 巴拭 **p-ɕức* is 拭 **ɕik* 'to wipe' < OC (Baxter-Sagart) **l̥ək*. The Old Vietnamese form supports an initial cluster **pl̥-* or **pəl̥-* in Old Chinese.

Under the Baxter-Sagart system, a syllabic preinitial is lost before **l-* in pre-Middle Chinese, while a non-syllabic preinitial prevails over **l-*: cf. 脫 **thwat* < **l̥ˤot* < **mə-l̥ˤot* 'peel off'; 粤 **pheŋ* < **ph̥ˤeŋ* < **[p.l̥]ˤeŋ* 'frank words'. The preinitial needs to be syllabic **pə-* under the

Baxter-Sagart system. Old Vietnamese data support a revised Baxter-Sagart reconstruction of 拭 *ɕik 'to wipe' as *pə.l̥ək.

2.2. 多边 *t-pen 'edge' < 邊 *pen 'edge, side'

(2) 庄　　故　　多边　　挨
    *chẳng có *t-pen cõi*
    NEG   have edge   border
    'does not have edges or borders' (Đại báo 44b-3)
    Original: （所有）無邊（塵刹佛）'(all the Buddhas of lands as countless as specks of dust) without borders'

I read 多边 (*đa-biên < ta-pen*) as Modern Vietnamese *ven* '(river)bank; near', cf. Middle Vietnamese *uen soũ* '*borda do rio; ripa fluminis, melius*' (Rhodes 1651: 865)[12]. My reading differs from Trần and Nguyễn (2007, #79) and Hoàng Thị Ngọ (1999: 147), both reading *biên*, the Sino-Vietnamese reading of the same etymon 邊 *pen 'edge, side'. Compared with *biên*, or other words deriving from the Chinese etymon, such as *bên* 'side, edge', the reading *ven* is preferable in order to account for the presyllable, which triggers lenition: Old Vietnamese *t- should cause lenition of initial *p into Vietnamese *v*. Also note the locution *ven cõi* survived in the glossing register of Modern Vietnamese: in Thiều Chửu's *Hán-Việt tự điển* (1942) the Chinese character 陲 *dʐwe 'frontier, border' is glossed as *ven cõi*.

The Old Vietnamese is reconstructed as *t-pen. As is explained above, Old Vietnamese *t-p- yields a lenited initial in Modern Vietnamese *v-*.

The Chinese original is 邊 *pen 'edge, side', Baxter-Sagart *pˤe[n]. The rime correspondence, where Chinese *e*-like vowel is rendered as Vietnamese *e* [ɛ] puts the word among the oldest borrowings from Chinese to Vietnamese, cf. forms preserving Old Chinese r-: *sen* 'lotus' < *kr-, borrowed from Chinese 蓮 *len, Baxter-Sagart *k.[r]ˤe[n]; *rèm* 'curtain', borrowed from Chinese 簾 *liem, Baxter-Sagart *rem.

Chiang Chia-lu (2011: 106) proposes another etymology for *ven*, namely from Chinese 緣/沿 'edge' *jwen < Baxter-Sagart *lon. As she herself recognizes, it is highly improbable for Middle Chinese *j-* < Old Chinese *l- to be rendered in Vietnamese by *v-*, an anomaly she explains as a late borrowing from an unspecified Southern Chinese dialect. Analysing the word as coming from 邊 *pen 'edge, side', which is furthermore supported by the Chữ Nôm orthography 边, 邊 and ⿰口边, avoids these formal problems.

The Old Vietnamese form *t-pen for the Chinese word 邊 *pen 'edge, side' supports an initial cluster *tp- or *təp- in Old Chinese.

Under the Baxter-Sagart system, syllabic and non-syllabic preinitials exist with the same consonantism, which has different treatments in Late Old and Middle Chinese. A syllabic preinitial is lost before a voiceless stop, cf. 九 *kiuwX < *[k]uʔ 'nine', which could derive from *tə.kuʔ (Baxter and Sagart 2014: 155). On the other hand, a non-syllabic alveolar preinitial prevails over a grave initial: cf. 帚 *tɕuwX < *tuʔ < *[t.p]əʔ 'peel off'. The preinitial in 邊 *pen needs to be syllabic *tə- under the Baxter-Sagart system. Old Vietnamese supports a revised reconstruction under Baxter-Sagart system as 邊 *tə.pˤe[n].

---

[12] Guillaume Jacques (p.c.) points to a difficulty in this etymology: the word is spellt *uen* in Rhodes, with the letter *v-* (probably [w-]) instead of ȸ ([β-]), the predicted outcome of a *p-* with lenition. However, as Haudricourt (1974) notes, there is a fluctuation between letters ȸ and *v* in Rhodes. For example, *ȸai* '*ombros; humerus*' (Rhodes 1651: 65), annotated 'alij *vaĩ*'.

2.3. 可耶, 可賒, 可車 *k-ɕa < 賒 *ɕæ 'far'

(3) 多　可耶
đi　**k-ɕa*
go　far
'go far' (Đại báo 15b-1)
Original: 遠行（憶念恩）'(the moral debt of parents missing you when you) travel far'

(4) 盎　　那　　強　　可賒
áng　*nạ*　*càng*　**k-ɕa*
father　mother　more　far
'the parents become still more distant' (Đại báo 36b-5)
Original: （及長大爲子索妻得他女子）父母轉踈 '(Parents look for a wife for their son; after they obtained another woman,) the parents become on the contrary distant from their son.'

(5) 當　　召　　庄　　可車
đáng　*chịu*　*chẳng*　**k-ɕa*
will　suffer　NEG　far
'will suffer it; it's not far' (Đại báo 43a-2)
Original: （三塗苦報）將受非遙 '(We will suffer the bitter repayment in the three ways;) we will suffer it not far away'

The orthographies 可耶 (*khả da* < **khaX jæ*), 可賒 (*khả xa* < **khaX ɕæ*) and 可車 (*khả xa* < **khaX tɕhæ*) transparently spell the Modern Vietnamese word *xa* 'far', Middle Vietnamese *xa* '*longe*; *distans*' (Rhodes 1651: 879), as Shimizu (1996), Trần and Nguyễn (2007, #90, #91) and Hoàng Thị Ngọ (1999: 153 et passim) observe.

The Old Vietnamese form of this word reconstructs to **k-ɕa*. The Old Vietnamese preinitial **k-* fails to cause lenition in **ɕ-* and yields *x-* [s] in Modern Vietnamese. This behaviour is probably regular and observed in **p-ɕức* 'to spread, to apply' > *xức* 'to spread, to apply' (§ 2.1) and **ʔ-ɕuóng* > *xuóng* 'chant' (§ 4). The orthography 可耶 (*khả da* < **khaX jæ*) might indicate an alternative pronunciation with lenition **k-ʑa*, which did not survive into Modern Vietnamese.

The Chinese original is 賒 **ɕæ* 'far'[13]. The Chinese character, which contains the component 貝 'cowrie shell, wealth', is generally used to write a word **ɕæ* that has the sense 'buy or sell on credit, defer payment'. This economic sense must be original, as it is the only one attested in pre-Hàn and Hàn texts such as *Rites of Zhou* 周禮 *Zhōulǐ* and *Book of Han* 漢書 *Hànshū* (fascicles 24, 91 and 99), as well as the definition given in the *Shuōwén* 說文.

The sense 'far' is attested much later. Its first occurrence according to *Hànyǔ Dà Cídiǎn* (Luo Zhufeng et al. 1993) is in *Bàopǔzǐ* 抱樸子 (Book of the Master Who Embraces Simplicity), a text ascribed to 葛洪 Gě Hóng (283–343), who led his whole life in 句容 Jùróng, in the 江左 Jiāngzuǒ (Lower Yangtse Basin) region. The subsequent attestations are from Southern Dynasties literature.

The word 賒 **ɕæ* 'far' is likely to derive from 賒 **ɕæ* 'to defer payment', via the following semantic development 'defer payment' > 'delay, postpone'[14] > 'late'[15] > 'far'. The word is likely

---

[13] The Southern provenance of the Chinese word suggests another possibility, namely that the Chinese word is borrowed from Vietnamese. This possibility can be dismissed, as the original Vietic word for 'far' is Vietnamese *ngái* - reduced to a secondary synonym of *xa*, as in the synonymic locution *xa ngái* - and Rục *cəɲá:j*.

[14] Cf. Táo Qián 陶潛's poem 和胡西曹示顧賊曹 (~ 400 CE): 悠悠待秋稼，寥落將賒遲 "I'm slowly waiting for the autumn harvest, but the crops look meagre and poor; the harvest will be delayed long."

a Southern dialect word which entered the literary language during the period of the cultural dominance of the Jiāngzuǒ region over the rest of China (roughly 317–619).

We finally note that the word 賖 *ɕæ has a third sense in Literary Chinese, namely 'many', which might have been borrowed into Vietnamese as *xe* (Trần Trọng Dương, 2012b).

The late date and derived sense suggest that Old Vietnamese *k-ɕa is borrowed from a Southern dialect form that corresponds with Middle Chinese ɕæ. Barring unexpected later prefixation, however, the *k in the source dialect to the Vietnamese form can be projected back in Old Chinese, and supports an initial cluster *kl̥- or *kəl̥- in Old Chinese. Under the Baxter-Sagart system, the Old Vietnamese data suggest that the reconstruction of 賖 *ɕæ, currently *l̥A, should be revised to *kə.l̥A.

2.4. 宁正 *k-ciếng < 井 *tsieŋX 'well'

(6) 工　　宁正　　　工　　法　　准　　▢目它
trong *k-ciếng* *trong* *bếp* *chốn* *đá*
in　　well　　　in　　stove　place　stone
'to wells, to stoves or the milling-places' (Đại báo 35a-3,4)
Original: （行來東西鄰里）井竈碓磨 '(when the children) go to neighbours east and west, or to wells, stoves, pestles or mills'

I follow Trần and Nguyễn (2007: #45) in reading 宁正 (*cá-chính* < *kaH-tɕieŋH) as *giếng* 'well', Middle Vietnamese *gyếng* '*poço de agoa pera beber*; *puteus aquae ad potum*' (Rhodes 1651: 283). Hoàng Thị Ngọ (1999: 165) reads *chiếng*, with the voiceless initial *ch-* based on the spelling 正 *chính* < *tɕieŋH. This reading should be rejected for two reasons. First, no word *chiếng* 'well' exists in Modern Vietnamese, a fact indirectly admitted by Hoàng herself, who explicitly glosses this form as '*giếng*'. Second, Old Vietnamese *k-ciếng regularly results in modern *giếng* by lenition.

The Old Vietnamese reconstructs to *k-ciếng. Old Vietnamese *k-c- yields lenited initial in Modern Vietnamese *gi-*. Compare for example 賊 *dzʌk 'bandit', where Vietnamese *giặc* derives by lenition from a form akin to Rục *kəcák*.

The Chinese original is clearly 井 *tsieŋX 'well'. The Old Vietnamese form supports an initial cluster *kts- or *kəts- in Old Chinese. Specifically, under the Baxter-Sagart system, the current reconstruction *C.tseŋʔ can be safely refined to *k.tseŋʔ.

2.5. ▢宁針 *k-kăm?? <? 箴, 鍼, 針 *tɕim 'pin, needle'

(7) 怛　　肝　　肆　　⺤　　▢宁針
dứt *gan* *ruột* *làm* *k-kăm??*
break　liver　intestines　make　pin
'breaks liver and intestines into pin-sized pieces' (Đại báo 41b-5)
Original: 寸斷肝腸 '(the work of bringing children up) breaks the liver and intestines (of parents) into inch-sized pieces'

As is the case for 'edge', we have again the situation where we need to choose a reading among different Vietnamese words meaning 'pin, needle', with similar pronunciation, and which can all be written with Chữ Nôm characters based on the Chinese characters 鍼, 針 *tɕim.

---

[15] Compare the locution 賖促 *ɕæ-tshiok 'later or sooner, long or short': 今古既異，賖促不同 "Today is different from long ago; (naturally, whether the period of mourning is) long or short is different" (宋書 Book of Sòng, fascicle 15 禮制二 = 通典 Tōngdiǎn, fascicle 82).

Judging from dictionaries of later Chữ Nôm usage, the major possibilities are *găm*, *kim* and Sino-Vietnamese *châm*. Another word, *ghim*, is usually written 金 in later Chữ Nôm, but the orthography 鍼, 針 is not unimaginable.

Trần Trọng Dương (p.c.) suggests that the character 曰丁針 should be read as *trăm* 'hundred', with *làm trăm* interpreted as 'make into a hundred pieces'. This is an unlikely interpretation, as *tr-* does not show any sign of being confounded with *ch-* until very late. In 'modern' Chữ Nôm, *trăm* is written in as 林 (*lâm* < **lim*) or with characters such as � with the phonogram 林.

The Vietnamese words concerned has been hitherto analysed as deriving from Chinese 箴, 鍼, 針 **tɕim* 'pin, needle', Baxter-Sagart **t.[k]əm*. However, Baxter and Sagart (2017) propose that the words *kim* and *ghim* derive rather from Chinese 金 **kim* 'metal, bronze', Baxter-Sagart **k(r)[ə]m*.

Trần and Nguyễn (2007: #45) interpret the composite character 曰丁針 as a case of digraphic orthography, composed of 丁 (*cá* < **kaH*) and 針 (*châm* < **tɕim*). Trần and Nguyễn's reading *châm* does not conform with their orthographical interpretation. If their orthographical interpretation is correct, the Modern Vietnamese reading would be one showing lenition, namely *găm* (or *ghim*), which would derive from Old Vietnamese **k-kăm*.

However, this reconstruction **k-kăm* is problematic on both internal and external grounds. Internally, sesquisyllabic languages typically do not admit the homorganic stops as the preinitial and as the initial: Ferlus' Proto-Vietic (2007) contains no syllables with **k-k-* or **t-ɗ-*. Externally, the Baxter-Sagart reconstruction **t.[k]əm* is supported by Lakkia *them*[1].

Another possible interpretation of the orthography is an early case of the *dấu cá* (the *cá* 丁 sign, cf. Nguyễn Quang Hồng, 2012), an orthographic device that indicates that the character is a specifically Nôm character, not to be read *à la chinoise*. If 曰丁針 is a case of *dấu cá*, all the non-Sino-Vietnamese readings, namely *găm*, *kim* (and probably *ghim*) are possible, with no implications on Old Chinese preinitials.

3. 麴 **s-kương* < 鏡 **kiæŋ*H 'mirror': Old Vietnamese preinitials and Old Chinese morphology

In this section, I examine the Old Vietnamese word 麴 **s-kương*, borrowed from Chinese 鏡 **kiæŋ*H 'mirror', Baxter-Sagart **C.qraŋʔ-s*, which touches on several questions of interest in the reconstruction of Old Chinese. The Old Vietnamese form is reconstructed in §3.1. The reconstructed Vietnamese form enables the Baxter-Sagart form to be refined to **s.qraŋʔ-s*, which, however, reveals an internal inconsistency in the Baxter-Sagart reconstruction of Old Chinese, discussed in §3.2. The unique value of this loanword is discussed in §3.3, as, regardless of the phonological interpretation, this word offers an example of an Old Chinese **s*-stop cluster which is supported both by morphology with well-known Tibeto-Burman correlates as well as directly transcribed with *s-* in foreign borrowing.

3.1. Reconstruction of the Old Vietnamese form

(8) 麴　　　朗　　　補　　　工　　　架
　　*s-kương *sáng*　*bỏ*　*trong*　*giá*
　　mirror　clear　put　in　　stand
　　'A clear mirror is put on a stand.' (Đại báo 12a-5)
　　Original: 明鏡掩粧臺 'A clear mirror covers the dressing table.'

I agree with Hoàng Thị Ngọ (1999: 151) and Trần and Nguyễn (2007: #115) in reading 麴 (司+姜, *tư-cương* < **si-kiaŋ*) as *gương* 'mirror', Middle Vietnamese *gương* '*espelho*; *speculum*' (Rhodes 1651: 301). This reading is further supported by the fact that the glyph 麴 survived into later Chữ Nôm usage, spelling the word *gương* 'mirror'.

The Old Vietnamese reconstructs to **s-kương*. Old Vietnamese **s-k-* corresponds with lenited Modern Vietnamese *g-* [ɣ-].

This word is borrowed from Chinese 鏡 **kiæŋH* 'mirror'. The correspondence between Chinese *qù* 去 and Vietnamese *bằng* 平 (*ngang* and *huyền*) tones is characteristic of older loans (cf. Alves 2016: 271). The Vietnamese data supports a cluster **s(ə)k-* or **s(ə)q-* in Old Chinese.

3.2. Can we reconstruct Baxter-Sagart *s.qraŋʔ-s?

On account of the Vietnamese initial lenition, Baxter and Sagart (2014) reconstruct **C.qraŋʔ-s* for this word. The Old Vietnamese data enables the refinement of this reconstruction into **s.qraŋʔ-s*.

This emendation, as attractive as it seems, is not compatible with the Baxter-Sagart reconstruction of Old Chinese as it currently stands. Initial **s.q(ˤ)-* is supposed to yield Middle Chinese *s-*, while **s.q(ˤ)r-* should yield Middle Chinese *tʂ-* (Baxter and Sagart 2014: 137–138). The amended reconstruction **s.qraŋʔ-s* would have become Middle Chinese **tʂiæŋH* and not *kiæŋH*.

Further examination reveals that there is a more general problem with **C.q-* clusters in the Baxter-Sagart reconstruction of Old Chinese. Baxter-Sagart **C-* stands for a preinitial that "cannot be identified because it has been lost in all the pronunciations under consideration" (Baxter and Sagart 2014: 168). According to this definition, whenever there is a **C-* in Baxter-Sagart, it can only stand for one of the preinitials reconstructed in the Baxter-Sagart OC phonology: **N-*, **m-*, **s-*, **p-*, **t-*, **k-*:

- **N.q-* and **m.q-* give Middle Chinese *j-* < **ɢ;
- **s.q-* gives Middle Chinese *s-*;
- **p.q-* has no examples but should give Middle Chinese *p-*, cf. **p.k-* which gives *p-* and **p.qʰ-* which gives *ph-*;
- **t.q-* gives Middle Chinese *tɕ-* < **t-*;
- **k.q-* probably did not exist, but would be, indeed, the only one that could give Middle Chinese *k-*.

Baxter-Sagart **C.q-* cannot give Middle Chinese *k-*, except in the unlikely case where **C* is **k*.

On the other hand, Baxter and Sagart (2014) explicitly reconstruct **C.q-* in order to account for the word 鏡 *kiæŋH* 'mirror'. The velar initial in Middle Chinese *kiæŋH* and Vietnamese *gương* is explained by a sound change by which "that uvulars shifted to velars after a tightly attached nonnasal preinitial: **C" (Baxter and Sagart 2014: 101).

Once we become cognizant of the fact that **C stands for any initial in Baxter-Sagart's notation, we notice that **C.q-* clusters are reconstructed twice, with different Middle Chinese reflexes. They are reconstructed once with specific preinitials, for which cases the preinitial prevails, and once in **C, for which cases the Middle Chinese reflex is fronted *k-*. For a hypothetical Old Chinese **s.qraŋʔ-s*, the first reconstruction would yield Middle Chinese *tʂiæŋH*, while the second would give *kiæŋH*. How the two classes should be reconstructed in a Baxter-Sagart-like system remains a problem to be solved.

3.3. 鏡 *kiæŋH 'mirror' as an example of Old Chinese *s-stop cluster

Baxter and Sagart (2014: 101, 168) reconstruct an initial *q- and a *qù*-tone (*-s) derived from *shǎng*-tone (*-ʔ) for 鏡 *kiæŋH < *s.qraŋʔ-s 'mirror', as they consider this word as being in the same family as 景, 影 *ʔiæŋX < *qraŋʔ 'shadow'. The refinement of the reconstruction into *s.qraŋʔ-s enables us to also refine its morphological derivation. Independently of the specific forms reconstructed, 鏡 *kiæŋH 'mirror' is one of the strongest examples of *s-stop clusters in Old Chinese.

A morphological prefix *s- in Old Chinese has been hypothesized since Conrady (1896). As the recent argument between Mei (2012) and Sagart & Baxter (2012) shows, there is broad agreement on the fact that the function of *s- would correspond to that in other Sino-Tibetan languages. However, due to different suppositions on the Middle Chinese reflexes of *s-, there are few words where scholars can actually agree on the reconstruction of *s-.

One of the well-attested functions of *s- is that of deriving an instrumental noun. Rgyal-rongic languages have a sA- prefix[16], that notably derives instrumental nouns from the verb roots. In Zbu Rgyalrong, the prefix *sɐ-* derives from *kɐ-twéʔ* 'to open' the instrumental noun *sɐ-thwéʔ* 'key'. The Zbu form belongs to a relic formation (Sun 2004: 293–294, Gong 2018: 211) which shows the age of prefix. Similarly, in written Tibetan, *s-* can derive instrumental nouns from verb roots: from *nod* 'to receive', *snod* 'container, vessel'; from *nyan* 'to listen', *snyan* 'ear (honorific)' (cf. Zhāng Jìchuān 2009: 246, Gong Xun 2017: 154–155, Jacques to appear). Baxter and Sagart (2014: 101) reconstruct *s- with an instrumental value. A good example (Sagart & Baxter 2012: 50–51) is 鋤 *dʐiɑ 'hoe', reconstructed as *s-l<r>a, derived from 除 *ɖiɑ < *l<r>a 'remove', as a hoe is an instrument of removal.

Based on the Old Vietnamese data and the Sino-Tibetan comparative evidence, we can reconstruct the noun 鏡 *kiæŋH < *s.qraŋʔ-s 'mirror' as derived from the verb 映 *ʔiæŋH < *qraŋʔ-s 'to reflect', which is itself a denominal from 景, 影 *ʔiæŋX < *qraŋʔ 'shadow > reflection'. It is relevant to note that the formation has a parallel in the Late Middle Chinese/Early Mandarin period. During the Sòng dynasty, when the syllable *kiæŋH* was taboo, being the personal name of the grandfather of the first emperor 趙敬 Zhào Jìng (MC *ɖieuX kiæŋH*), mirrors were then called 照子 *zhàozǐ*, MC *tɕieuH-tsiX*. This formation is entirely parallel to what is proposed here for 鏡 *kiæŋH < *s.qraŋʔ-s, as 子 *zǐ*, MC *tsiX* was the dominant instrumental noun suffix, added to 照 zhào, MC *tɕieuH* 'to shine on', later 'to reflect'.

The word 鏡 *kiæŋH for 'mirror' is a Warring States period (475–221 BC) neologism replacing original 鑑 *kæmH, coined inside Warring States period Chinese and not inherited from an older state of the language. It can be regarded as one of the safest examples both of a *s-stop cluster in Old Chinese and an example of the instrumental value of the *s- prefix in Old Chinese.

4. *ʔ-: template-filling dummy prefix?

We find two words of Chinese origin showing the preinitial *ʔ-, written 阿 (*a < a*):

阿路 *ʔ-lò 'kiln, oven' < 爐 *lu 'oven'

(9) 阿路　　堵　　毒　　坡律
　　 *ʔ-lò　 đỏ　 độc　 sốt
　　 kiln　 red　 poison　 burning
　　 'red kiln and burning hot poison' (Đại Báo 29a-3,4)
　　 Original: 炎爐熱毒 'kiln of fire and burning poison'

---

[16] This prefix is, to my knowledge, first described in Lín Xiàngróng (1993: 162), cf. also Sun 1998: 142, Yán-mùchū 2005, Jacques 2008: 332–333, Lai 2017: 158 etc.

Both Trần and Nguyễn (2007, #6) and Hoàng Thị Ngọ (1999: 161) read 阿路 (*a-lộ* < **a-luH*) as *lò* 'kiln, oven', Middle Vietnamese *lò* '*fogão, forno; fornax*' (Rhodes 1651: 417). This word is borrowed from Chinese 爐 **lu* 'oven' < Baxter-Sagart **[r]ˤa*. Note that the word is not from the earliest layer of Chinese loans in Vietnamese, as the initial Middle Chinese **l-* < Old Chinese **r-*, is reflected as *l-* in this word but *r-* in the earliest Chinese loans in Vietnamese.

阿唱 **ʔ-ɕướng* 'to chant' < 唱 **tɕhaŋH* 'to chant'
(10) 省　　　阿唱　　　浪
　　　tiếng　**ʔ-ɕướng*　*rằng*
　　　sound　chant　　this
　　　'They chant as follows' (Đại Báo 22b-4,5)
　　　Original: 高聲唱言 'They loudly chant as follows'

I follow Shimizu (1996) in reading the Sino-Vietnamese *xướng* 'to chant'. Hoàng Thị Ngọ (1999: 158) reads *hét* 'to cry', usually written phonetically as 喝 (*hát* < **hat*) or 嘁 (*hiết* < **hiʌt*). This reading has no argument in its favour except that *xướng* is infrequent in contemporary Vietnamese. Concerning this point, cf. *xướng ten*, '*nomear alguem como por rol; vocare nomen alicuius tanquam è catalogo*' (Rhodes 1651: 900), which should be emended into *xướng tên* 'chant the name-list'. This word is borrowed from Chinese 唱 **tɕhaŋH* 'to chant' < Baxter-Sagart **mə-tʰaŋ-s*.

Both borrowings show the preinitial **ʔ-*, which has not been hypothesized for Old Chinese. We also note that they show signs of rather late borrowing from Chinese: for **ʔ-ɕướng* 'chant', the phonological correspondence is of the newest type, and completely coincides with the Sino-Vietnamese pronunciation. We further notice that in the So dialect of Thavưng (Suwilai 1996), monosyllabic Thai-Lao verbs are borrowed with a prefix *ʔa-*: *lâːk* 'to drag along the ground' is adapted as *ʔalâːk*. For **ʔ-ɕướng* 'to chant', at least, it is attractive to suppose a similar prefix **ʔ-*, a dummy prefix that fills a sesquisyllabic phonological template. The problem of the prefix **ʔ-* in Old Vietnamese and other Vietic languages deserves further study.

5. Conclusion

This paper is a preliminary examination of Old Vietnamese loans from Chinese attested with sesquisyllabic orthography in *Phật Thuyết Đại Báo Phụ Mẫu Ân Trọng Kinh*.

Old Vietnamese data is one of the most eloquent pieces of evidence that argues for the presence and determines the identity of **C₁(ə)C₂-* initial clusters in Old Chinese. It enables the proposal of a few changes to Baxter and Sagart's reconstruction of Old Chinese.

- In two cases, where Baxter and Sagart reconstruct an unknown preinitial **C-* based on lenition in Modern Vietnamese, Old Vietnamese enables the identification of the unknown preinitial: 井 *tsieŋX* 'well', Baxter-Sagart **C.tseŋʔ*, modified to **k.tseŋʔ*; 鏡 *kiæŋH* 'mirror', Baxter-Sagart **C.qraŋʔ-s*, modified to **s.qraŋʔ-s*.
- In two cases, where Baxter and Sagart reconstruct Old Chinese forms without preinitials, Old Vietnamese suggests that a preinitial likely existed in Old Chinese: 拭 *ɕik* 'to wipe', Baxter-Sagart **l̥ək*, modified to **pə.l̥ək*; 邊 *pen* 'edge, side', Baxter-Sagart **pˤe[n]*, modified to **tə.pˤe[n]*.

The Old Vietnamese word 阿路 **ʔ-lò* 'kiln, oven' and 阿唱 **ʔ-ɕướng* 'to chant' are possible examples of a prefix **ʔ-*, internal in Vietic, which is also attested in the conservative Vietic language So Thavưng (Suwilai 1996).

The word 麭 *s-k*ương* < 鏡 *kiæŋH* 'mirror', which permits a revised Baxter-Sagart reconstruction as *s.qraŋʔ-s, is important in two different ways. First, it points to the fact, hitherto ignored, that *C.q- clusters are reconstructed twice in the Baxter-Sagart system, with different reflexes in Middle Chinese and other sources. Second, this loan in Old Vietnamese provides one of the strongest examples both of an Old Chinese initial consonant cluster of the *s-stop type, and of the Old Chinese prefix *s-, which derives here an instrumental noun from a verb.

References

Alves, Mark J. 2016. Identifying Early Sino-Vietnamese Vocabulary via Linguistic, Historical, Archaeological, and Ethnological Data. *Bulletin of Chinese Linguistics* 9: 264–295.

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

Baxter, William H., Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. Oxford University Press.

Baxter, William H., Laurent Sagart. 2017. Old Chinese reconstruction, a response to Schuessler. *Diachronica* 34(4): 559–576.

Chiang, Chia-lu. 2011. *Yuènán hànzìyīn de lìshǐ céngcì yánjiù* (越南漢字音的歷史層次研究). Doctor thesis, National Taiwan Normal University.

Ferlus, Michel. 1976. Du nouveau sur la spirantisation ancienne en vietnamien. *Bulletin de la Société de Linguistique de Paris* 71: 305–312.

Ferlus, Michel. 1982. Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien. *Cahiers de linguistique — Asie orientale* 11: 83–106.

Ferlus, Michel. 1991. Le dialecte vietnamien de Vinh. *24th International Conference on Sino-Tibetan Languages and Linguistics, October 1991, Bangkok*. Available online at: https://halshs.archives-ouvertes.fr/halshs-00922737.

Ferlus, Michel. 2007. *Lexique de racines Proto Viet-Muong (Proto-Vietic Lexicon)*. Ms., accessed from the *SEAlang Mon-Khmer Etymological Dictionary* <http://sealang.net/monkhmer/dictionary/>.

Ferlus, Michel. 2010. Le cycle khmer des douze animaux : histoire d'un contact ancien entre Vietnam et Cambodge. *Cahiers de linguistique - Asie Orientale* 39: 3–19.

Goad, Heather. 2011. The Representation of sC Clusters. In: Marc van Oostendorp, Colin Ewen, Elizabeth Hume, Keren Rice (eds.). *The Blackwell Companion to Phonology*: 898–923. Oxford: Blackwell Publishing.

Gong, Xun, Yunfan Lai. 2017. Consonant clusters. In: Rint Sybesma et al. (ed.). *Encyclopedia of Chinese language and linguistics*: 665–672. Leiden: Brill.

Gong, Xun. 2017. Jiāróng-yǔzǔ yǔyán yǔ Shànggǔ Hànyǔ de Xíngtàixué (嘉絨語組語言與上古漢語的形態學). In: Park Haeree, Chen Shaoxuan (eds.). *Gǔwénzì yǔ hànyǔ lìshǐ bǐjiào yīnyùnxué* (古文字與漢語歷史比較音韻學) [*Paleography and Chinese Historical and Comparative Phonology*]: 132–154. Shanghai: Fudan Daxue Chubanshe.

Gong, Xun. 2018. *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest. Une étude descriptive, typologique et comparative*. Doctoral thesis, INALCO.

Haudricourt, André-Georges. 1965. Les mutations consonantiques des occlusives initiales en môn-khmer. *Bulletin de la Société de Linguistique de Paris* 60(1): 160–172.

Haudricourt, André-Georges. 1974. Hai Chữ B, trong cuốn từ điển của A-lếch-xan-đơ Rốt. *Ngôn Ngữ* 1974(4): 37–38.

Hayes, La Vaughn H. 1992. Vietic and Việt-Mường: a new subgrouping in Mon-Khmer. *Mon-Khmer Studies* 21: 211–228.

Hoàng, Thị Ngọ. 1996. *Nghiên cứu về chữ Nôm và tiếng Việt qua bản giải âm Phật thuyết đại báo phụ mẫu ân trọng kinh*. Candidate Thesis. Hanoi: Viện nghiên cứu Hán Nôm.

Hoàng, Thị Ngọ. 1999. *Chữ Nôm và tiếng Việt qua bản giải âm Phật thuyết Đại báo phụ mẫu ân trọng kinh*. Hanoi: Nxb Khoa học Xã hội.

Jacques, Guillaume (Xiàng, Bólín). 2008. *Jiāróngyǔ yánjiū* (嘉绒语研究). Beijing: Minzu chubanshe.

Jacques, Guillaume. Forthc. Fossil nominalization prefixes in Tibetan and Chinese. To appear in: *Bulletin of Chinese Linguistics*.

Lai, Yunfan. 2017. *Grammaire du khroskyabs de Wobzi*. PhD thesis, Université Sorbonne Nouvelle Paris 3.

Lin, Xiangrong. 1993. *Jiāróngyǔ yánjiū* (嘉戎语研究). Chengdu: Sichuan minzu chubanshe.

Luo, Zhufeng et al. 1993. *Hànyǔ dà cídiǎn* (汉语大词典). Shanghai: Shanghai cishu chubanshe.

Maspero, Henri. 1912. Études sur la phonétique historique de la langue annamite. *Bulletin de l'Ecole française d'Extrême-Orient* 12: 1–124.

Maspero, Henri. 1930. Préfixes et dérivation en chinois archaïque. *Mémoires de la Société de Linguistique de Paris* 23(5): 313–327.

Mei, Tsu-lin. 2012. The Causative *s- and Nominalizing *-s in Old Chinese and Related Matters in Proto-Sino-Tibetan. *Language and Linguistics* 13(1): 1–28.

Michaud, Alexis, Michel Ferlus, Minh-Châu Nguyễn. 2015. Strata of standardization: the Phong Nha dialect of Vietnamese (Quảng Bình Province) in historical perspective. *Linguistics of the Tibeto-Burman Area* 38(1): 124–162.

Nguyễn, Ngọc San. 1982. Góp vài ý kiến về âm đầu tiếng Việt cổ qua cứ liệu Nôm trong sách «Phật thuyết đại báo phụ mẫu ân trọng kinh». *Ngôn ngữ* 1982(3): 34–41.

Nguyễn, Quang Hồng. 2001. *Khái luận văn tự học chữ Nôm*. Hanoi: NXB Giáo dục.

Nguyễn, Tá Nhí. 1987. Lối đánh dấu cá trong chữ Nôm. *Tạp chí Hán Nôm* 1987(1): 35–38.

Nguyễn, Tài Cẩn. 1979. *Nguồn gốc và quá trình hình thành cách đọc Hán Việt*. Hanoi: NXB Khoa học Xã hội.

Nguyễn, Tài Cẩn. 2008. *Một số vấn đề về ngành Nôm học*. Ms. (Hội nghị Nôm học, April 2008, Temple University).

Nguyễn, Phú Phong, Trí Dõi Trần, Michel Ferlus. 1988. *Lexique Vietnamien-Rục-Français: parler d'une minorité ethnique des montagnes de Quảng Bình, Vietnam avec notes ethnographiques et introduction linguistique*. Paris: Sudestasie.

Nguyễn, Hữu Vinh, Đặng Thế Kiệt, Nguyễn Doãn Vượng, Lê Văn Đặng, Nguyễn Văn Sâm, Nguyễn Ngọc Bích, Trần Uyên Thi. 2009. *Từ Điển Chữ Nôm Trích Dẫn: Dictionary of Nôm characters with excerpts*. Westminster: Viện Việt-Học.

Pan, Wuyun. 2000. *Hànyǔ lìshǐ yīnyùnxué* (汉语历史音韵学). Shanghai: Shanghai jiaoyu chubanshe.

Pulleyblank, Edwin G. 1981. Some notes on Chinese historical phonology. *Bulletin de l'École française d'Extrême-Orient* 69: 277–288.

De Rhodes, Alexandre. 1651. *Dictionarium annamiticum, lusitanum et latinum*. Rome: Sacrae Congregationis de Propaganda Fide. Copy held at the Bayerische Staatsbibliothek, accessed at: http://mdz-nbn-resolving.de/urn:nbn:de:bvb:12-bsb10521857-2.

Sagart, Laurent, William H. Baxter. 2012. Reconstructing the *s- Prefix in Old Chinese. *Language and Linguistics* 13(1): 29–59.

Shimizu, Masaaki. 1996. Kanbun-Chunomubun taiyaku Bussetsu Daihō Bumo Onjūgyō ni miru chunomu ni tsuite (漢文＝字喃文対訳『仏説大報父母恩重経』に見る字喃について). *Ningen-kankyōgaku* 5: 83–104.

Shimizu, Masaaki, Lê Thị Liên, Shiro Momoki. 2005. A Trace of Disyllabicity of Vietnamese in the 14th Century - Chu Nom Characters Contained in the Inscription of Hộ Thành Mountain. *Kōbeshi Gaikokugo Daigaku Gaikokugaku kenkyū* 64: 17–49.

Shimizu, Masaaki. 2015. A reconstruction of Ancient Vietnamese initials using Chữ Nôm materials. *NINJAL research papers* 9: 135–158.

Skal·bzang 'Gyur·med (Gésāng Jūmiǎn), Skal·bzang Dbyangs·can (Gésāng Yāngjīng). 2002. *Zàngyǔ Fāngyán Gàilùn* (藏语方言概论). Beijing: Minzu chubanshe.

Sun, Jackson T.-S. 1998. Nominal morphology in Caodeng rGyalrong. *Bulletin of the Institute of History and Philology, Academia Sinica* 69(1): 103–149.

Sun, Jackson T.-S. 2004. Verb-stem variations in Showu rGyalrong. In: Lin Ying-chin et al. (eds.). *Studies on Sino-Tibetan Languages: Papers in Honor of Professor Hwang-Cherng Gong on His Seventieth Birthday*: 269–296. Taipei: Academia Sinica.

Suwilai, Premsrirat. 1996. Phonological characteristics of So (Thavung), a Vietic language of Thailand. *Mon-Khmer Studies* 26: 161–178.

Suwilai, Premsrirat. 2000. *So (Thavung) Preliminary Dictionary*. Salaya & Melbourne: Institute of Language and Culture for Rural Development, Mahidol University and the University of Melbourne.

Thiều Chửu. 1942. *Hán-Việt Tự Điển*. Hanoi: Đuốc Tuệ xuất bản.

Thompson, Laurence C. 1976. Proto-Viet-Muong Phonology. In: Philip N. Jenner and Laurence C. Thompson and Stanley Starosta (eds.). *Austroasiatic Studies*: 1113–1204. Honolulu: The University Press of Hawaii.

Trần, Trọng Dương. 2010. *Về dấu vết chữ Nôm ky húy trong sách 'Phật thuyết'*. Ms. (Tạp chí Văn hóa Nghệ An 7).

Trần, Trọng Dương. 2011. *Phật thuyết* có phải dịch phẩm Nôm của thế kỉ XII? *Ngôn ngữ* 2011(4): 31–47.

Trần, Trọng Dương. 2012a. *Một số từ gốc Hán có cấu trúc CCVC qua ngữ liệu thơ Nôm trong Quốc âm thi tập*. Presentation made at the Hội thảo kỷ niệm 1 năm ngày mất GS Nguyễn Tài Cẩn.

Trần Trọng Dương. 2012b, Từ nguyên của XE và các điệp thức của nó. In: *Thông báo Hán Nôm học năm 2010–2011*: 557–562. Hanoi: Thế giới.

Trần, Uyên Thi, Hữu Vinh Nguyễn. 2007. *Ai vẽ được, ai xoá được? Dấu vết âm việt cổ: từ song tiết và phụ âm kép*. Presentation made at the Hội Nghị Quốc Tế về Tiếng Việt, Viện Việt-Học, Westminster.

Wang, Li. 1948. Hànyuèyǔ yánjiù (漢越語研究). *Lingnan xuebao* 9(1): 1–96.

Yánmùchū. 2005. Shìlùn Jiāróngyǔ zhōng sɑ de yǔfǎ gōngnéng jí xíngshì (试论嘉戎语中sɑ 的语法功能及形式). *Aba Shifan Gaodeng Zhuanke Xuexiao Xuebao* 22(1): 91–92.

Zhang, Jichuan. 2009. *Zàngyǔ cízú yánjiū: gǔdài zàngzú rúhé fēngfù fāzhǎn tāmende cíhuì* (藏语词族研究: 古代藏族如何丰富发展他们的词汇). Beijing: Shehuikexue Wenxian Chubanshe.

Гун Сюнь. Китайские заимствования с полуторасложной структурой в старовьетнамском языке

В настоящее время практически не остается сомнений в том, что в древнекитайском языке присутствовали начальные сочетания согласных, включавшие в себя пресиллабы; однако прямых указаний на то, какие конкретные формы реконструировать для тех или иных слов, сохранилось очень немного. В данной статье вводится в научный обиход новый источник, ранее остававшийся неисследованным — старовьетнамский язык, достаточно подробно зафиксированный в одном сохранившемся тексте, где ряд слов, которые в современном вьетнамском языке являются односложными, записан в орфографии, предполагающей для них изначальный полуторасложный характер. Для некоторых слов, заимствованных из китайского, только варианты их записи в старовьетнамской орфографии позволяют определить изначальную форму, в которой они были заимствованы. Небольшой список известных на сегодняшний день полуторасложных слов, зафиксированных в старовьетнамском тексте и имеющих китайское происхождение, включает как лексемы, для которых можно надежно реконструировать в китайском сочетания согласных, так и слова, где вероятна префиксация уже на вьетнамской почве.

Ключевые слова: древнекитайский язык, старовьетнамский язык, историческая реконструкция, полуторасложные слова, префиксальная морфология.

Zhang Shuya[†], *Guillaume Jacques*[‡], *Lai Yunfan*[§]

[†] Institut National des langues et des civilisations orientales, Paris; zhang.shuyafr@gmail.com

[‡] Centre National de la Recherche Scientifique, Paris; rgyalrongskad@gmail.com

[§] Max Planck Institute for the Science of Human History, Jena; khroskyabs@gmail.com

A study of cognates between Gyalrong languages and Old Chinese

Gyalrongic languages, a subgroup of the Burmo-Qiangic branch of the Sino-Tibetan family, are spoken in the Western Sichuan Province of China. They are polysynthetic languages, and present rich verbal morphology. Although they are not closely related to Chinese, they are of particular interest for Sino-Tibetan/Trans-Himalayan comparative linguistics with regards to their conservative phonology and morphology. Based on previous studies on Old Chinese phonology, combining with recent fieldwork data, this paper aims to show how Gyalrong languages could shed light on Old Chinese morphology and thus contribute to the Old Chinese reconstruction. It also proposes a list of possible cognates between Old Chinese, Gyalrong languages, indicating also Tibetan cognates when available.

Keywords: Gyalrong languages, Old Chinese language, etymological cognates, comparative morphology, historical reconstruction.

1. Introduction

Although Gyalrongic languages are not closely related to Chinese (Sagart et al. 2019), they are of particular interest for Sino-Tibetan/Trans-Himalayan comparative studies since they are the rare languages in the family exhibiting complex consonant clusters (Lai 2017) and conservative morphologies (Jacques 2016b; Gong 2017). However, they have hitherto been neglected by comparativists. This paper aims at providing easily accessible data on potential cognates between Old Chinese (OC) and Gyalrong, and discusses how these comparisons could contribute to improve OC reconstruction.

In section 2, we show how Gyalrong data could shed light on distinction between the anticausative and passive derivations in OC. Section 3 proposes a hypothesis on the origin of OC *l- from pre-OC *sl-. Section 4 raises some issues concerning medial *-r- in current OC reconstructions.

The last section contains a list of possible cognates shared among the Gyalrong languages, Tibetan and the Old Chinese, classified by OC rhymes[1]. The reason for choosing rhymes as the order of classification is justified by the fact that there is some consensus on the rhyme categories of OC, while there are more divergences concerning the consonants. In addition, correspondences are easier to establish between the rhymes of OC and Gyalrongic than between their onsets. Many of the etymologies in this work have been discussed before, in particular those with Tibetan cognates, for which we cite the reference in Schuessler (2007) as a summary of previous scholarship (in particular Conrady 1896; Simon 1929; Coblin 1986; Peiros and Starostin 1996; Gong 1995), and builds on previous comparative research concerning Gyalrong languages (Jacques 2004, 2005).

---

[1] One of the reviewers has kindly advised to add a long table as data supplement. We agree with this advice, but before making the table accessible as recommended by the reader, it is better to wait until we have data of more languages, in order to provide a more complete database.

The Gyalrong data come from three varieties, Japhug, Brag-bar (Situ) and Cogtse (Situ). For each cognate, we first list the Chinese word, provided with middle Chinese (MC) and OC reconstructions. We systematically cite Baxter and Sagart's (2014) reconstruction, but in cases when the comparisons are incompatible with their model, we propose alternative possibilities[2].

2. Anticausative and passive derivations

Anticausative[3] verbs in Gyalrong languages present initial prenasalizing alternations with regards to the basic transitive verbs (Jacques 2008, 84–87; Zhang 2016, 93–95). This process is no longer productive in modern Gyalrong languages. Jacques (2015c) suggests that this prenasalizing element could be etymologically related to the spontaneous-autobenefactive prefix *nɯ-* in Japhug. While the prenasalizing anticausative has been lexicalized, the spotaneous-autobenefactive has undergone regularization and is still highy productive.

As shown in (1), the anticausative verb Brag-bar (Situ) *kə-mblēt* 'be extinct' (S *ɟavən* 'bee'), presents initial prenasalizing alternation in regard to the transitive verb Brag-bar (Situ) *ka-plēt* 'to extinguish' (A *kəɕtɕə̄k* 'leopard', O *kəɟōk* 'sheep').

(1) a. Brag-bar (Situ) *ka-plēt* 'to extinguish'
         kəɕtɕə̄k   kə   kəɟōk   kəʑē   tə   na-plôt.
         leopard   ERG   sheep   all   DET   PFV-to.extinguish$_{II}$
         'The leopard has eaten all the sheep.'

     b. Brag-bar (Situ) *kə-mblēt* 'be extinct'
         tʂaɕî   ɟavə̄n   rgombâ   kəsə̂m   ˈro-ɕput   rɐnə,
         Bkra.shis bee   box   three   IFR.PFV -to.feed$_I$ but
         u-ɟavə̄n   ɲo   kəʑē   no-ˈkə-mblɐt
         3SG.POSS-bee   PL   all   IFR.PFV-3NS.INTR-be.extinct$_I$
         'Bkra.shis has fed three beehives of bees, but his bees are all died.'

Note that the direction of derivation should be from a transitive/causative to an intransitive/anticausative verb. This process is not productive in modern Gyalrong languages, however Jacques (2008, 86) noticed that in Japhug it is applied to a Tibetan loanword Japhug *χtor* 'to scatter (vt.)' (Tibetan གཏོར *gtor* 'to scatter') ~ Japhug *ʁndʑr* 'be scattered (vi.)', whereas the anticausative counterpart does not exist in Tibetan (Jacques 2015c).

Many languages in the ST/TH family have voicing alternations related to transitivity. In Middle Chinese and attested Sinitic languages the prenasalized element has been lost, leaving only initial voicing alternations. Baxter and Sagart (2014) reconstructed a prenasalizing *N-* prefix to account for this voicing alternations attested in MC, as for instance between 別 *pjet* < B/S *pret* 'to separate' and 別 *bjet* < B/S *N-pret* 'be separated', an example semantically compatible with an interpretation as an anticausative derivation.

However, we also find in OC voicing alternations with meanings that cannot be interpreted as anticausative, such as that between 見 *kenH* < B/S *[k]ˤen-s* 'to see (vt.)' and 現 *ɣenH* < B/S *N-[k]ˤen-s* 'to appear (vi.)'. Anticausative verbs denote spontaneous situations and

---

[2] Since the aim of this study is to illustrate the extent to which Gyalrong languages can contribute to OC reconstruction, we provide some minor amendments to existing reconstruction systems. A full revision of OC historical phonology is beyond the scope of this paper.

[3] This argument demoting mechanism derives a intransitive verb from a transitive verb, by suppressing the A of the basic verb and promoting the original O to S (Dixon and Aikhenvald 2000, 315).

exclude an external cause or agent (Haspelmath 1993), and are thus incompatible with verbs of perception. It is thus difficult to compare the voicing alternation in 現 *ɣenH* < B/S **N-[k]ˤen-s* with the Gyalrong anticausative.

In Gyalrong languages however, we also find a passive prefix (*a-* in Japhug, *o-* in Brag-bar, *ŋa-* in Cogtse) originating from a nasal **ŋa-*. Example (2) shows the triple contrast between a base transitive verb, its anticausative and its passive in Japhug.

(2) Japhug *prɤt* 'to cut (vt.)'
    Japhug *mbrɤt* 'be cut (ANTICAUS, vi.)'
    Japhug *a-prɤt* 'be cut (PASS, vi.)'

While a morphological distinction between passive and anticausative seems to be absent in OC, it is possible that the voicing alternation reconstructed as **N-* by Baxter and Sagart results in fact from the merger of an anticausative **N-* and a passive **ŋ-* derivation (this would not be the only case of merger between etymologically unrelated morphological alternations in OC; the *qusheng* derivation appears to be a similar case, see Jacques 2016a).

The verbs 敗 *bæjH* / 敗 *pæjH* possibly provide evidence for a contrast between passive and anticausative derivations. The reading *bæjH* with a voiced initial has two distinct meanings 'be damaged' (3) and 'be defeated' (4), whereas that with an unvoiced initial *pæjH* only means 'to defeat' (5).

(3) Anticausative 敗 *bæjH* < B/S **N-pˤra[t]-s*

魚	餒	而	肉	敗,	不	食
ŋjo	*nwojX*	*ɲi*	*ɲuwk*	*bæjH*	*pwot*	*ʑik*
fish	decay	CONJ	meat	decay	NEG	to.eat

'He did not eat fish or flesh which has gone bad. (Translation of James Legge).'
《論語·鄉黨》

(4) Passive 敗 *bæjH* < B/S **N-pˤra[t]-s*

梁惠王	以	土地	之	故,	糜
Ljaŋ.ɣwejH.hjwaŋ	*jiX*	*tʰuXdijH*	*tɕi*	*kuH*	*mje*
King.Hui.of.Liang	PREP:because.of	territory	GEN	reason	to.tore

爛	其	民	而	戰	之,	大	敗
lanH	*gi*	*mjin*	*ɲi*	*tɕenH*	*tɕi*	*tajH*	*bæjH*
to.destroy	POSS.3SG	people	CONJ	to.lead...to.war	PRON	great	be.defeat

'The king Hui of Liang, for the matter of territory, tore and destroyed his people, leading them to battle. Sustaining a great defeat. (Translation of James Legge)'
《孟子·盡心下》

(5) Transitive 敗 *pæjH* < B/S **pˤra[t]-s*

冬,	與	越人	水	戰,	大	敗	越人
towŋ	*joX*	*hjwotɲin*	*ɕwij*	*tɕenH*	*tajH*	*pæjH*	*hjwotɲin*
winter	with	Yue.people	water	to.fight	great	to.defeat	Yue.people

'In the winter he had an engagement with that of Yue, on which he inflicted a great defeat. (Translation by James Legge)' 《莊子·逍遙遊》

Since the meaning 'be damaged' is necessarily older than 'be defeated' (a semantic change 'defeat' → 'destroy' seems highly unlikely), this verb seems to provide evidence for the idea that the intransitive 敗 *bæjH* is the base form, and that the transitive 敗 *pæjH* is derived from it by a causative prefix (as was insightfully pointed out by Wang Hongzhi, pc).

However, the reasons for not reconstructing a sigmatic causative to account for these voicing alternations have been discussed at length elsewhere (Jacques 2015b; Sagart and Baxter 2012), and it is possible to reconcile OC and Gyalrong data by supposing that 敗 *pæjH* is cognate of Brag-bar (Situ) *ka-prāt* 'to break', Japhug *prɤt* 'to break', and originally meant 'destroy, damage'. The form 敗 *bæjH* *N-pˤrats in the meaning 'be damaged' would be an anti-causative form of this transitive verb ('become damaged spontaneously, by itself'), itself cognate to Cogtse (Situ) *kɐ-mbrêt* 'to break', Brag-bar (Situ) *kə-mbrāt* 'to break' and Japhug *mbrɤt* 'to break'. The base verb then underwent semantic narrowing to the sense of 'to defeat', from which a passive *ŋ-pˤrats (merging early with *N-pˤrats, and undistinguishable in MC from the anticausative) was derived.

3. On the origin of OC *l̥-

Sagart and Baxter (2012) propose a sound change chain concerning consonant clusters consisting of a presyllable *s(ə)- and a nasal. Pre-OC tight *s*- preinitial consonant clusters result in voiceless nasals in OC, whereas loose clusters became tight clusters in OC, as illustrated in (6):

(6) Origin of voiceless nasal in OC (adapted from Table 1 in Sagart and Baxter 2012)

Pre-OC		OC		MC
*sm-, *sn-, *sŋ-	>	*m̥-,*n̥-,*ŋ̊-	>	*χ, th, χ*
*sə.m-, *sə.n-, *sə.ŋ-	>	*sm-, *sn-, *sŋ-	>	*s*

In two sets (7), *l̥- in OC corresponds to the consonant cluster *ɕl-* in Gyalrong languages, suggesting that one of the origins of OC *l̥- is earlier *sl- (in line with Yakhontov and Starostin 1989, 218).

(7) OC | Gyalrong
失 *ɕit* < B/S *l̥i[t]* 'to lose' | Japhug *ɕluɣ* 'to let sth. fall'
| Cogtse (Situ) ka-ʃlâk 'to fall (from hand)'
脱 *thwat* < B/S*mə-l̥ˤot* 'to take off' | Brag-bar(Situ) *ka-ɕlēt* 'to fall (from hand)'

We thus propose *sli[t] and *slˤot* as pre-OC forms of 失 and 脱.

(8) Origin of OC *l̥

Pre-OC		OC		MC
*sl-	>	*l̥-	>	*th-, ɕ-*
*sə.l-	>	*sl-	>	*s-*

Japhug *ɕluɣ* 'to let sth. fall' is a lexicalized causative of Japhug *luɣ* 'to detach' (the productive causative, expressing a volitional action, is Japhug *suɣluɣ* 'to cause to detach'), and this is a case where the OC preserves a morphological element as indirect trace only.

4. *-r- medial in OC reconstruction

The medial *-r- in OC only partially corresponds to medial *-r-* in languages other than Chinese. In particular, based on comparisons by Gong (1995), Handel (2002) points out that the onsets reconstructed as dental stop or dental affricates+*-r- generally correspond to clusters with preinitial *r*- in Tibetan, and suggests that metathesis from *rC- to *Cr- took place, as summarized in (9):

(9) OC Non-Chinese languages

Grave *Cr- *Cr-

      *r-C (>*C-) *r-C

Acute *r-C (>*Cr-) *r-C

*s- *s- *sr- (before front vowels)

*s- *s- *s- (before non-front vowels)

Comparative data from Gyalrong languages could shed light at least on two aspects.

First, the comparison of Japhug *tɤ-zrɤm* 'root' to 參 B/S *s.rum* 'root' shows that Handel (2002)'s hypothesis that *sr-* merges with *s-* in languages other than Chinese before non-front vowel must be amended (Jacques 2015a).

Second, the grave initial syllables reconstructed with medial *-r-* in OC in present reconstructions correspond in some cases to words with preinitial *r-* in Gyalrong languages (Table 1). We suggest to reconstruct preinitial *r-* in these cases in OC.[4] The difference between preinitial *r-* and medial *-r-* is not detectable on the basis of Chinese-internal evidence alone, though (depending on the relative chronology of sound changes between Chinese and Viet-Muong) it is possible that preinitial *r-* would yield lenition in old loanwords into Vietnamese (see Pulleyblank 1981, 284 for a suggestion in these lines).

Unlike Handel, we do not think that it is necessary to suppose that metathesis took place in OC – for examples of retroflexion of dental stop by preceding liquids, see Burrow (1972) on Indic languages and Kümmel (2007, 231). Rather, the rhotic (and perhaps other types of preinitials, as suggested by Pulleyblank) became a suprasegmental rhotacized voice quality, as proposed by (Miyake 2012).

Table 1. Preinitial *r-* in OC

MC	B/S	Amended OC	Gyalrong cognates
冰 *piŋ* 'ice'	*p.rəŋ*	*rpəm*	Cogtse (Situ) *ta-rpâm* 'ice'
			Brag-bar (Situ) *ta-rpâm* 'ice'
			Japhug *tɤ-jpɣom* 'ice'
眉 *mij* 'eyebrow'	*mr[ə][r]*	*rməj*	Cogtse (Situ) *ta-rɲê* 'hair',
			Brag-bar (Situ) *ta-rɲê* 'hair'
			Japhug *tɤ-rme* 'hair'
顏 *ŋæn* 'face'	*C.ŋˤrar*	*rŋˤan*	Japhug *tu-rŋa* 'face'

There are however a number of unexplained exceptions, such as 熬 *ŋaw* < B/S *ŋˤaw* 'to fry, roast' (Cogtse (Situ) *ka-rŋô* 'to fry', Brag-bar (Situ) *ka-rŋô* 'to fry', Japhug *rŋu* 'to fry') or 名 *mjieŋ* < B/S *C.meŋ* 'name' (Cogtse (Situ) *tə-rmê* 'name', Brag-bar (Situ) *tə-rmiê* 'name', Japhug *tɤ-rmi* 'name') for which no rhotacization can be reconstructed in OC. Note that Tibetan *ming* �མིང་ 'name' also lacks a medial or preinitial *r-*.

---

[4] This hypothesis implies to abandon the reconstruction *r.ŋˤaʔ for 魯 *luX* proposed to account for a *xiesheng* relationship with 魚 *ŋjo* 'fish'.

5. Gyalrong cognates

5.1. Open syllable rhymes

5.1.1. 鱼 yu *a

The rhyme 鱼 *a generally corresponds to the vowel a in Cogtse (Situ), iɛ in Brag-bar (Situ) and a in Japhug. The Brag-bar has undergone the *a >iɛ sound change.

1. 斧 *pjuX* < B/S *p(r)aʔ 'axe', Cogtse (Situ) *ʃa-rpâ* 'axe', Brag-bar (Situ) *ɕɐ-rpiɛ̂* 'axe', Japhug *tɯ-rpa* 'axe'. The first syllable in Situ etyma is the constructed status of the noun Cogtse (Situ) *ʃẽ* 'firewood', Brag-bar (Situ) *ɕɛ̂* 'firewood'. In Japhug, the *tɯ-* prefix is the indefinite possessive. The Gyalrongic data suggest the presence of an *r-* preinitial in Old Chinese (see section 4), but the rime *-ju* in MC is ambiguous as to the presence or absence of a rhotacizing element.

2. 下 *ɣæH* < B/S *m-gˤraʔ-s 'to descend', Cogtse (Situ) *kɐ-ŋgrêk* 'fall down', Brag-bar (Situ) *kə-ŋgriɛ̂* 'fall down', Japhug *ŋgra* 'fall down'. The velar coda in Cogtse is unexplained. In Gyalrongic languages this verb is the anticausative of the transitive verb (see section 2), reflected by Brag-bar (Situ) *ka-kriɛ̂* 'to cause to fall down', Japhug *kra* 'to cause to fall down'. An alternative etymology would be 落 *lak* < B/S *kə.rˤak 'to fall'[5].

3. 蘇 *su* < B/S *s-ŋˤa 'to revive'. OC *sŋ- became s- in MC (Baxter 1992, 225; Sagart 1999, 65; Schuessler 2007, 482), Cogtse (Situ) *kə-mə-sŋâr* 'feel cold', Brag-bar (Situ) *kə-mə-sŋiɛ̂r* 'come to oneself', Japhug *sŋa* 'come to oneself'. The Situ forms have a *mə-* prefix and a *-r* coda which remain unexplained.

4. 吾 *ŋu* < B/S *ŋˤa 'I, my', Tibetan *nga* ᨶ 'I' (Schuessler 2007, 518), Cogtse (Situ) *ŋā* 'I', Brag-bar (Situ) *ŋā* 'I', Japhug *a-ʑo* 'I'. Northern Gyalrong languages have lost the velar nasal initial *ŋ-* of the singular first person pronouns.

5. 咀 *dzjoX* < *dzaʔ 'to chew', Tibetan *za-ba* ᨦᨦ 'to eat' (Schuessler 2007, 323). Cogtse (Situ) *ka-zā* 'to eat', Brag-bar (Situ) *ka-ndziɛ̂* 'to eat', Japhug *ndza* 'to eat'.

6. 五 *ŋuX* < B/S *C.ŋˤaʔ 'five', Tibetan *lnga* ᨥ 'five' (Schuessler 2007, 519). Cogtse (Situ) *kəmŋô* 'five', Brag-bar (Situ) *kəmŋɛ̂j* 'five', Japhug *kɯmŋu* 'five'. The correspondence of the vowel is irregular. In addition, the correspondence between *lŋ-* in Tibetan and *mŋ-* in Gyalrong languages is also unexplained (Jacques 2004, 125).

7. 夜 *jæH* < B/S *[ɢ]Ak-s 'night', Tibetan *zhag* ᨥᨦ 'one night' (Schuessler 2007, 561–62), Brag-bar (Situ) *rə-ɟāk* 'one night', Japhug *tɤ-rʑaʁ* 'night'. The Brag-bar form takes the numeral prefix *rə-* 'one'. While this cognate set is not in doubt, the reconstruction of the onset in OC, Tibetan and Gyalrong is problematic. Japhug *rʑ-* and Situ *ɟ-* could respectively originate from *rj- and *j-, and OC might be better reconstructed with a primary yod initial.

5.1.2. 支 zhi *e

1. 兒 *ɲe* < B/S *ŋe 'child'. Jacques (2004, 223) relates this word to Japhug *nɤ-rɲi* 'baby', although the first syllable in the Japhug etyma is unexplained. This words could also be related to a lexicalized diminutive suffix in Situ, as in Cogtse (Situ) *kho-lɲâ* 'infant' Brag-bar (Situ) *tapə-rɲiɛ̂* 'infant', however, correspondence between *i* in Japhug and *a/iɛ* in Situ seems irregular and needs to be further investigated.

---

[5] Etymology suggested by Laurent Sagart.

2. 髀 *bejX* < B/S **m-pˤeʔ* 'thighbone', Tibetan *dpyi* དཔྱི 'thighbone, hipbone' (Schuessler 2007, 164), Japhug *tɯ-χpɣi* 'thigh'.

<h3 style="text-align:center">5.1.3. 之 zhi *ə</h3>

The rhyme 之 **ə* corresponds to *a* in Cogtse (Situ), **a > iɛ* in Brag-bar (Situ) and *a* in Japhug.

1. 富 *pjuwH* < B/S **pək-s* 'rich'. Sagart (2017) relates this word to Tibetan *phag* ཕག 'pig'. If this etymology is accepted, it is also comparable to Cogtse (Situ) *pāk* 'pig', Brag-bar (Situ) *piāk* 'pig', Japhug *paʁ* 'pig'. Schuessler (2007, 152) relates the 'pig' etymon to 豝 *pae* < B/S **pˤra* instead.

2. 子 *tsiX* < B/S **[ts]əʔ* 'child'. Schuessler (2007, 633) proposes that this word is the ST root for 'offspring, child' and relates it to Tibetan *tsha-po* ཚ་པོ 'grandchild; nephew'. It could be comparable to Brag-bar (Situ) *ta-ˈtsa-pu* 'father's sister's child', Japhug *tɯ-ftsa* 'father's sister's child; sister's child' (for the Gyalrong kinship systems and the designation of this word, see Jacques 2012). The Brag-bar etymon is in an incomplete status of noun-compounding since the word is attested with a penultimate accent instead of the final tonal contrast, it can be resulted by a recent suffixation of diminutive -*pu* (*ta-pū* 'child') on the base noun **ta-(p)tsa* due to changes of its kin terminology. The *p-* preinitial is preserved in the derived social relation collective Brag-bar (Situ) *koɕə-ptsa-pâ* 'paternal cross cousins' daughters'. Correspondence between *p-* preinitial in Situ and *f-* preinitial in Japhug is discussed in Jacques (2004, 269–70). Another possible cognate of 子 is only shared by Situ dialects, Cogtse (Situ) *tə-tsâ* 'boy, son', Brag-bar (Situ) *tə-ziê* 'boy, son'. The Brag-bar etymon has undergone the **ts- > z-* lenition, as in Brag-bar (Situ) **tə-matsa > tə-mazâ* 'mother's sister's child'.

3. 裘 *gjuw* < B/S **[g]ʷə* 'fur garment', Tibetan *gos* གོས 'clothes', Cogtse (Situ) *tə-wê* 'garment', Brag-bar (Situ) *tə-ŋgâ* 'garment', Japhug *ŋga* 'to wear'. The *w-* initial of the Cogtse etymon can be explained by **g- > w-* lenition in this dialect. The Tibetan etymon is suffixed by the -*s* nominalizer.

4. 耳 *ɲiX* < B/S **C.nəʔ* 'ear', Tibetan *rna* རྣ 'ear' (Schuessler 2007, 225–26), Cogtse (Situ) *tə-rnâ* 'ear', Brag-bar (Situ) *tə-rniê* 'ear', Japhug *tɯ-rna* 'ear'.

5. 牛 *ŋjuw* < B/S **[ŋ]ʷə* 'ox', Cogtse (Situ) *nəŋâ* 'cow', Brag-bar (Situ) *nəɲiê* 'cow', Japhug *nɯŋa* 'cow'. The first element in the Gyalrong etyma could be related to the constructed status of the word Cogtse (Situ) *tə-nū* 'udder', Brag-bar (Situ) *tə-nū* 'udder', Japhug *tɯ-nu* 'udder'. Although most Gyalrong languages have lost the **-w-* medial, evidence can be found in Zbu *ŋwéʔ* (Gong 2018, 40).

<h3 style="text-align:center">5.1.4. 侯 hou *o</h3>

The rhyme **o* in OC corresponds to the back vowels in Gyalrong language. It corresponds to *ɯ* or *u* in Japhug, and *u* or *o* in Situ.

1. 坐 *dzwaX* < B/S **[dz]ˤo[j]ʔ* 'to sit', Japhug *a-mdzɯ* 'to sit'.
2. 乳 *ɲuX* < B/S **noʔ* 'milk; nipple', Tibetan *nu-ma* ནུ་མ 'udder' (Schuessler 2007, 446), Cogtse (Situ) *tə-nū* 'udder', Brag-bar (Situ) *tə-nū* 'udder', Japhug *tɯ-nu* 'udder'.
3. 后 *ɣuwX* < B/S **Gˤ(r)oʔ* 'sovereign; queen', Tibetan *mgo* མགོ 'head', Tibetan *'go-pa'* འགོ་བ 'headman' (Schuessler 2007, 279–80), Cogtse (Situ) *ta-kō* 'head', Brag-bar (Situ) *ta-wô* 'head', Japhug *tɯ-ku* 'head'. The initial consonant of the Brag-ba etymon has undergone the **k- > w-* lenition.

4. 後 *ɣuwX* < B/S *[ɢ]ˤ(r)oʔ 'after' , Cogtse (Situ) *mə-ŋkhū* 'be after', Brag-bar (Situ) *mə-ŋkhū* 'be after', Japhug *ma-qhu* 'be after'. The Situ and Japhug etyma have a denominal prefix (Jacques 2004, 407).

5. 腑 *pjuX* < *poʔ 'internal organs', Cogtse (Situ) *tə-po-lɐŋtʃê* 'intestine', Brag-bar (Situ) *tə-vo-loŋtɕâ* 'intestine', Japhug *tɯ-pu* 'intestine'. The initial consonant of the cognate base in Brag-bar has undergone the *p- > v-* lenition.

6. 軀 *khju* < *kʰ(r)o 'body; person', Tibetan *sku* སྐུ 'body, statue' (Schuessler 2007, 435), Cogtse (Situ) *tə-skrū* 'body', Brag-bar (Situ) *tə-skrū* 'body', Japhug *tɯ-skhrɯ* 'body'.

7. 寇 *kʰuwH* < B/S *[k]ʰˤ(r)o-s 'to rob; robber', Tibetan *rku-ba* རྐུབ 'to steal' (Schuessler 2007, 336), Brag-bar (Situ) *kə-mərkâ* 'bandit', Japhug *mɯrkɯ* 'to steal'. The Brag-bar form might be loanword from other Gyalrong languages, since it has an unexpected *ə* vowel.

8. 嘔 *ʔuwX* < B/S *qˤ(r)oʔ 'vomit', Tibetan *skyug* སྐྱུག 'vomit' (Schuessler 2007, 407, 595), Japhug *qioʁ* 'vomit'. The final glottal stop *-ʔ possibly corresponds to Tibetan *-g* and Japhug *-ʁ* (Schuessler 2007, 31–32).

5.1.5. 幽 you *u

The rhyme 幽 *u* correspond to *u* in Cogtse, *u* in Brag-bar and *ɯ* in Japhug.

1. 保 *pawX* < B/S *pˤuʔ 'to take care of, protect'. In Gyalrong languages we found fixed expressions such as Brag-bar (Situ) *u-pû ka-viê* 'to take care of', Japhug *ɯ-pɯ kɤ-pa* 'to take care of', consisting of a possessive prefixing action nominal and a light verb Brag-bar (Situ) *ka-viê* 'to do', Japhug *kɤ-pa* 'to do'. The inalienably possessed noun *u-pû/ɯ-pɯ* in this collocation can be analyzed as an action nominal, as in other non-ambiguous cases Brag-bar (Situ) *tə-ɕmô ka-viê* 'to steal', consisting of the action nominal Brag-bar (Situ) *tə-ɕmô* 'stealing' whose corresponding verbal form is Brag-bar (Situ) *ka-ɕmô* 'to steal'. Such analysis would suggest the existence of the corresponding verbs Brag-bar (Situ) **ka-pû* 'to take care of' Japhug **kɤ-pɯ* 'to take care of' at an earlier stage. An alternative to this light verb construction is a denominal transitive Brag-bar (Situ) *ka-ra-pupū* 'to take care of' (stem I *rapəpā*, stem II *rapəpâ*, stem I' *rapupū*) (see 10). The verb is clearly derived from a nominal base *-pu*, by adding a denominal prefix *ra-* on the reduplicated base. It is possible that in Gyalrong languages the basic verb 'to take care of' has been lost, whereas the derived action nominal has been preserved and became then the base noun. A similar case is the verb *donner* 'to give' in French, which does not come from Lat. *dare* but was recreated from *don* (< *donum* 'gift') (List 2016).

2. 浮 *bjuw* < B/S *m.b(r)u 'float' could be indirectly related to Brag-bar (Situ) *ʑbrū* 'boat' and Japhug *ʑmbru* 'boat', assuming that these nouns are fossilized sigmatic instrumental nominalizations from a verb **mbru* meaning 'float'.

3. 帽 *mawH* < B/S *mˤuk-s 'hat'. Sagart (2017) relates this word to Tibetan *rmog* རྨོག 'hat, helmet'. The word for 'mushroom' shared in all Gyalrong languages could be possible cognate, Cogtse (Situ) *te-jmōk* 'mushroom', Brag-bar (Situ) *ta-jmōk* 'mushroom', Japhug *tɤ-jmɤɣ* 'mushroom' (Breton *tok touseg* 'frog hat' for 'mushroom').

4. 腦 *nawX* < B/S *nˤ[u]ʔ 'brain', Cogtse (Situ) *tə-rnōk* 'brain', Brag-bar (Situ) *ta-rnōk* 'brain', Japhug *tɯ-rnoʁ* 'brain'.

5. 肘 *ʈjuwX* < B/S *t-[k]<r>uʔ 'elbow', Tibetan *gru-mo* གྲུ་མོ 'elbow' (Schuessler 2007, 624), Cogtse (Situ) *tə-krū* 'elbow', Brag-bar (Situ) *tə-krû* 'elbow', Japhug *tɯ-zgru* 'elbow'. The *t-* preinitial in OC could be related to the indefinite possessive prefix *tV-* in Gyalrong languages.

6. 爪 *tʂæwX* < B/S *[ts]ˤ<r>uʔ* 'claw', Cogtse (Situ) *ta-ndzrū* 'nail', Brag-bar (Situ) *ta-ndzrū* 'nail', Japhug *tɯ-ndzru* 'nail'.

7. 九 *kjuwX* < B/S *[k]uʔ* 'nine', Tibetan *dgu* དགུ 'nine' (Schuessler 2007, 320), Cogtse (Situ) *kəngû* 'nine', Brag-bar (Situ) *kəngû* 'nine', Japhug *ngɯt* 'nine'. The final *-t* in the Japhug etyma is an innovation in Northern Gyalrong languages; probably due to analogy from Japhug *kɯrcat* 'eighy' (Jacques 2004, 253).

8. 舅 *gjuwX* < B/S *[g](r)uʔ* 'mother's brother', Tibetan *a-khu* ཨ་ཁུ 'father's brother' (Schuessler 2007, 321), Brag-bar (Situ) *a-kû* 'mother's brother'. The first element in Brag-bar and Tibetan etymon is the vocative prefix. This word is the common Sino-Tibetan root for mother's brother (Benedict 1942), the Tibetan etyma has undergone a semantic shift from mother's brother to father's brother (Nagano 1994). Correspondence between voiced *g-* initial in OC and voiceless *k-* and *kh-* in Gyalrong and Tibetan etyma could be explained in a similar as (Jacques 2017a) proposes for 父 *pjuX* < B/S *[N-p](r)aʔ* and Tibetan *pha* ཕ 'father'. ST kin terms are often prefixed either by vocative or by possessive, therefore it is possible that a nasal element is inserted between the possessive/vocative prefix and the root, as the case in Limbu (Davids and Driem 1985).

9. 韭 *kjuwX* < B/S *s.[k](r)uʔ* 'allium', Cogtse (Situ) *ʃkō* 'onion', Brag-bar (Situ) *ɕkō* 'onion', Japhug *ɕku* 'onion', Tibetan *sgog-pa* སྒོག་པ 'wild onion'.

(10) a. Brag-bar (Situ) *u-pû ka-viɛ̂* 'to take care of'

nə-ta-ka-mbə̄-n		tə	ostō	u-pū	rɐ-viɛ̂-n
1SG PFV-1→2-NMLZ-to.give$_{II}$-2SG	DET	really	3SG.POSS-protection	IMP-to.do$_{I}$-2SG

'Take care of what I gave to you.'

b. Brag-bar (Situ) *ka-ra-pupū* 'to take care of'

nə-ta-ka-mbə̄-n		tə	ostō	rɐ-nə-ra-ˈpəpə-n
1SG PFV-1→2-NMLZ-to.give$_{II}$-2SG	DET	really	IMP-AUTO-DENOM-protection$_{I}$-2SG

'Take care of what I gave to you.'

5.2. *-k* ending rhymes

5.2.1. 鐸 duo *ak

The rhyme 鐸 *ak* in OC corresponds to *-ak* in Situ and *-aʁ* in Japhug. A few examples correspond to open syllables *a/iɛ* and *-a* respectively, a type of correspondence discussed by Sagart (2017).

1. 百 *pæk* < B/S *pˤrak* 'hundred', Tibetan *brgja* བརྒྱ 'hundred', Cogtse (Situ) *pərjâ* 'hundred', Brag-bar (Situ) *pərjē* 'hundred', Japhug *ɣurʑa* 'hundred'. In Brag-bar, *iɛ < *a* is realized as *e* after palatal stops. The initial *p-* of the OC etymon corresponds to a presyllable in Gyalrong languages, showing that OC underwent monosyllabicization in this word. The final stop in OC is discussed in Schuessler (2007, 69–70) and Sagart (2017).

2. 膊 *phak* < *pʰˤak* 'shoulder blade', Tibetan *phrag* ཕྲག 'shoulder' (Schuessler 2007, 170), Cogtse (Situ) *tə-rpāk* 'shoulder', Brag-bar (Situ) *ta-rpiāk* 'shoulder', Japhug *tɯ-rpaʁ* 'shoulder'. Since Tibetan lacks a *rp-* cluster, the comparison between Gyalrong and Tibetan suggest that a metathesis *rp- ➙ *pr-* took place in pre-Tibetan in this etymon.

3. 薄 *bak* < B/S *[b]ˤak* 'thin', Cogtse (Situ) *kə-mbâ* 'to be thin', Brag-bar (Situ) *kə-mbiɛ̂* 'to be thin', Japhug *mba* 'to be thin'.

4. 莫 *mak* < B/S **mˤak* 'there is no X such that X ...', Brag-bar (Situ) *kə-miāk* 'not be', Japhug *maʁ* 'not be'. This a common Sino-Tibetan negative copula, it is also related to the negative prefix *mV-* in Tibeto-Burman languages (Lai 2017, 248). Pulleyblank (2000) (also mentioned by Schuessler 2007, 70) considered the coda *-k* in **mˤak* to be a distributive suffix **-k*, and 莫 is derived from 無 **ma > mju* 'not have'. Another example provided by Pulleyblank (2000) is 或 **[ɢ]ʷˤək > ɣwok* 'someone, something', derived from 有 **[ɢ]ʷəʔ > hjuwX* 'have, exist'. Pulleyblank's hypothesis is not compatible with the comparison presented above.

5. 胳 *kak* < B/S **[C.q]ˤak* 'armpit', Tibetan *lag* ལག 'arm' (Schuessler 2007, 252), Cogtse (Situ) *tɐ-jēk* 'arm', Brag-bar (Situ) *ta-jāk* 'arm', Japhug *tɯ-jaʁ* 'arm'. This comparison would be more compatible with a reconstruction such as **klˤak* in OC.

6. 腋 *jek* < B/S **[ɢ](r)Ak* 'armpit', Tibetan *bzhag.'og* བཞག་འོག 'armpit' Japhug *tɯ-pjaʁpa* 'armpit' (the syllable *-pa* is a noun meaning 'down, bottom part'). This comparisons suggest eihter a primary yod or a lateral in OC rather than a uvular.

5.2.2. 錫 xi **ek*

1. 隻 *tsjek* < B/S **tek* 'single', Tibetan *gcig* གཅིག 'one' (Schuessler 2007, 614), Cogtse (Situ) *kə-tēk* 'one', Brag-bar (Situ) *kə-rīk* 'one', Japhug *tɤɣ* 'one'. The initial consonant *r-* of the Brag-bar etymon is due to the **t- > r-* lenition. A *t-* initial allomorph can be found in Brag-bar (Situ) *zɹa-tēk* 'eleven'.

2. 滴 *tek* < **tek* 'to drop; drop', Tibetan *thigs-pa* ཐིགས་པ 'a drop', Tibetan *'thig-pa* འཐིག་པ 'to drop' (Schuessler 2007, 209), Cogtse (Situ) *nthēk* 'drop CLF', Brag-bar (Situ) *rə-nthāk* 'drop'.

5.2.3. 職 zhi **ək*

A general correspondence between **ək* in OC and *ak* in Gyalrong languages can be found. The vowel *iɛ < *a* is realized as *iaC* before velar codas in Brag-bar.

1. 革 *kɛk* < **kˤrək* 'hide' could be compared with Cogtse (Situ) *ka-klāk* 'peel off', Brag-bar (Situ) *ka-kliāk* 'peel off', under the assumption that the meaning 'hide' derives from 'skin that has been peeled off'.

2. 慝 *thok* < B/S **n̥ˤək* 'evil' Tibetan *nag-po* ནག་པོ 'black' (Schuessler 2007, 493), Cogtse (Situ) *kə-nēk* 'be black', Brag-bar (Situ) *kə-nâk* 'be black', Japhug *ɲaʁ* 'be black'. The meaning 'evil' is also found in the lexicalized nominal form Japhug *ɯ-ɣɲaʁ* 'disastrous consequence', in which the preinitial *ɣ-* is a lenited form of the velar participle prefix.

3. 織 *tɕik* < B/S **tək* 'to weave', Tibetan *btags* བཏགས 'to weave (PST)' (Schuessler 2007, 615), Brag-bar (Situ) *ka-tiāk* 'to weave', Japhug *taʁ* 'to weave'.

4. 色 *srik* < B/S **s.rək* 'color; countenance'. Jacques (2015a) relates this word to Japhug *tɤ-zraʁ* 'shame', despite the semantic gap. Alternative cognates 赤 *tshjek* < B/S **[t-qʰ](r)Ak*, 赫 *xæk* < B/S **qʰˤrak* 'red' are proposed by Schuessler (2007, 451).

5.2.4. **ik*

1. 節 *tset* < B/S **tsˤik* 'joint', Tibetan *tshigs* ཚིགས 'segment' (Schuessler 2007, 312), Brag-bar (Situ) *rə-rtsāk* 'a segment', Japhug *tɯ-rtsɤɣ* 'segment'. The first syllable of the Brag-bar etymon is the numeral. It could also be related to Cogtse (Situ) *ka-ra-ntsīk* 'to cut (into segments', Brag-bar (Situ) *ka-ra-ntsīk* 'to cut (into segments)' (Stem I *ra-ntsāk*, Stem II *ra-ntsîk*, Stem I' *ra-ntsīk*, details of stem alternations in Brag-bar see Zhang 2018).

2. 蝨 *ṣit* < B/S **srik* 'louse'. **-ik* in OC became *-it* in MC (Baxter and Sagart 2014, 236). Schuessler (2007, 461) compares this word to Tibetan *shig* ཤིག་ 'louse'. Japhug *zruɣ* 'louse' is also related. The correspondence between **sr-* in OC, Tibetan *ɕ-* and *zr-* in Japhug is treated in Jacques (2015a).

5.2.5. 屋 wu **ok*

1. 曲 *khjowk* < B/S **kʰ(r)ok* 'to bend' and 局 *gjowk* < B/S **N-kh(r)ok* 'be bent, curved', Japhug *kɤɣ* 'to curve' and its anticausative Japhug *ŋgɤɣ* 'be bent'

2. 啄 *træwk* < B/S **mə-tˤ<r>ok* 'to peck', Cogtse (Situ) *tə-ntôk* 'beak', Brag-bar (Situ) *tə-ntēk* 'beak'. The vowel *e* in the Brag-bar etymon is due to a **o >e* sound change. This noun originates from a verb 'to peck' also indirectly attested as a fossilized participle in the compound Cogtse (Situ) *ʃi-kō kə-ntōk* 'woodpecker', Brag-bar (Situ) *ɕewo-ˈkontek* 'woodpecker', which can be regarded as a lexicalized s/a deverbal noun (Sun and Lin 2007; Jacques 2016c).

3. 殼 *khæwk* < B/S **[kʰ]ˤrok* 'shell', Cogtse (Situ) *tə-rkhō* 'bark', Brag-bar (Situ) *tə-rkhō* 'bark', Japhug *u-rqhu* 'bark', Tibetan *skogs-pa* སྐོགས་པ 'bark'. The absence of *-k* coda in Gyalrong etyma is unexplained.

4. 角 *kæwk* < B/S **C.[k]ˤrok* 'horn'. Schuessler (2007, 309) relates this word to Tibetan *rwa* ར 'horn' (on the rhyme *-wa* in this word, see Jacques 2009). Cognates are found in Cogtse (Situ) *tə-rū* 'horn', Brag-bar (Situ) *tə-rū* 'horn', Japhug *ta-ʁru* 'horn'.

5.2.6. 覺 jue **uk*, **iwk*

1. 毒 *dowk* < B/S **[d]ˤuk* 'poison', Tibetan *dug* དུག་ 'poison' (Schuessler 2007, 216), Cogtse (Situ) *tɜk* 'poison', Brag-bar (Situ) *tə-ndōk* 'poison', Japhug *tɤ-ndɤɣ* 'poison'. Cogtse (Situ) *tɜk* 'poison' is a loanword, otherwise we would expect a voiced initial. Note that in Japhug, the loanword Japhug *tuɣ* 'poison' (Jacques 2004, 166) coexists with the cognate form Japhug *tɤ-ndɤɣ* 'poison'.

2. 六 *ljuwk* < B/S **k.ruk* 'six', Tibetan *drug* དྲུག་ 'six' (Schuessler 2007, 362), Cogtse (Situ) *kə-tʂōk* 'six', Brag-bar (Situ) *kətʂōk* 'six', Japhug *kutʂɤɣ* 'six'. The initial consonant *tʂ-* in Situ and Japhug comes from the Proto-Gyalrong consonant cluster **tr-* (Jacques 2004, 294).

3. 腹 *pjuwk* < B/S **p(r)uk* 'belly', Tibetan *phugs* ཕུགས་ 'innermost part' (Schuessler 2007, 246), Cogtse (Situ) *tə-pōk* 'belly', Brag-bar (Situ) *tə-vōk* 'belly', Tibetan *bug* བུག་ 'hole'. The Brag-bra etymon presents **p- > v-* lenition.

4. 目 *mjuwk* < B/S **C.m(r)[u]k* 'eye', Tibetan *mig* མིག་ 'eye' (Schuessler 2007, 392–93), Cogtse (Situ) *tə-mɲāk* 'eye', Brag-bar (Situ) *tə-mɲāk* 'eye', Japhug *tɯ-mɲaʁ* 'eye'. The consonant cluster *mɲ-* in Gyalrong etyma comes from **mj-* in Proto-Gyalrong (Jacques 2004, 299).

5. 蹙 *tsjuwk* < **tsˤiwk*. The rhyme **iwk* is reconstructed for this word given its *xiesheng* relation 戚 *tshek* < B/S **s.tʰˤiwk*. A comparison with Japhug *sthoʁ* 'to press' would be possible if the sound change **st- → *ts-* is accepted (Bodman 1969).

5.3. *-ŋ* ending rhymes

Gyalrong languages have lost final **-ŋ* in native words, so that cognates with Chinese generally have open syllables corresponding to OC **-ŋ*.

5.3.1. 陽 yang *aŋ

The rhyme 陽 *aŋ of OC corresponds to -o in Japhug and Situ.

1. 紡 *phjaŋX* < *pʰaŋʔ 'twist, spin', Tibetan 'phang-ma འཕང་མ 'spindle' (Schuessler 2007, 232), Cogtse (Situ) *ka-pô* 'to spin', Brag-bar (Situ) *ka-pō* 'to spin', Japhug *pɣo* 'to spin'. The ɣ medial in the Japhug etymon comes from the velarized vowel of Proto-Gyalrong *paˀŋ.

2. 孟 *maengH* < B/S *m̥ˤraŋ-s 'eldest, great' is possibly related to Tibetan *mag-pa* མག་པ 'son-in-law', Cogtse (Situ) *tə-nmak* 'son-in-law', Japhug *tɤ-nmaʁ* 'husband'. Another etymology suggested by L. Sagart relates this word with Cogtse (Situ) *kə-mbrô* 'be tall', Brag-bar (Situ) *kə-mbrô* 'be tall', Japhug *mbro* 'be tall'.

3. 攘 *ɲaŋX* < *naŋʔ 'to oppose; disturb', Cogtse (Situ) *ka-nô* 'to chase', Brag-bar (Situ) *ka-nô* 'to chase', Japhug *no* 'to chase'. 攘 is related to 讓 *ɲaŋH* < B/S *naŋ-s 'to allow', and could be possibly related to Tibetan *g.nang-ba* གནང་བ 'to give, allow'.

4. 想 *sjaŋX* < B/S *[s]aŋʔ 'to think', Cogtse (Situ) *ka-səsō* 'to think', Brag-bar (Situ) *ka-səsō* 'to think', Japhug *suso* 'to think'. The verbs in modern Gyalrong languages are possible reduplicated forms of *saŋ > *so.

5. 剛 *kaŋ* < B/S *kˤaŋ 'strong, hard', Cogtse (Situ) *kə-rkô* 'be hard', Japhug *rko* 'be hard', Tibetan *mkhrang-po* མཁྲང་པོ 'be hard'. The form 勍 *gjæŋ* < (possible reconstructions would include *graŋ, *N-kraŋ or *N-rkaŋ) possibly reflects a variant of the same root with a *r like the Gyalrong and Tibetan cognates.

6. 羊 *jaŋ* < *laŋ B/S *ɢaŋ 'sheep', Tibetan *g.yang-dkar* གཡང་དཀར 'sheep', Cogtse (Situ) *kə-jō* 'sheep', Brag-bar (Situ) *kə-ɟōk* 'sheep', Japhug *qa-ʑo* 'sheep'. The first syllable in Gyalrong date is the prefix designating animals, which could correspond to the *g-* preinitial in the Tibetan etymon (< PT *Gə-jaŋ, Jacques 2013). The *-k* coda in the Brag-bar etymon is of unclear origin. Similar phenomenon has also been reported in Kyom-kyo (Situ), in which some words can have two realisations, either with the final velar stop or not, *kəɟoʔ ~ kəɟok, kəruʔ ~ kəruk* (Prins 2016, 47–48).

7. 癢 *jaŋX* < B/S *Cə.ɢaŋʔ 'to itch', Tibetan *g.ya'-ba* གཡའ་བ 'to itch' (Schuessler 2007, 559), Cogtse (Situ) *ka-rajâk* 'to itch', Brag-bar (Situ) *ka-reɟâk* 'to itch', Japhug *rɤʑa* 'to itch'. *rV-* in the Gyalrong etyma could be the denominal prefix. The *-k* coda in Situ is unexplained. This etymon is better reconstructed with a primary initial yod in OC (Jacques 2013).

5.3.2. 耕 geng *eŋ

The rhyme 耕 *eŋ in OC corresponds to e in Situ and i in Japhug. A group of *i/*e in Brag-bar became ɐj, whose phonetic condition remains to be investigated.

1. 名 *mjieŋ* < B/S *C.meŋ 'name', Tibetan *ming* མིང 'name', (Schuessler 2007, 387). The Gyalrong cognates Cogtse (Situ) *tə-rmê* 'name', Brag-bar (Situ) *tə-rmiê* 'name', Japhug *tɤ-rmi* 'name' have a *r-* preinitial without equivalent in Chinese and Tibetan.

2. 鳴 *mjæŋ* < B/S *m.reŋ 'cry (of birds or animals)', Cogtse (Situ) *kə-mərê* 'be loud', Brag-bar (Situ) *kə-mbrêj* 'be loud', Japhug *mbri* 'be loud'. The consonant cluster *mr-* in Cogtse became two syllables, with the insertion of ə, whereas in Brag-bar and Japhug, *mr- > mbr-* (Jacques 2004, 137).

3. 繩 *ʑiŋ* < B/S *Cə-m.rəŋ 'string, cord', Brag-bar (Situ) *tə-mbrē* 'rope', Japhug *tɯ-mbri* 'rope'.

5.3.3. 蒸 zheng *əŋ

The rhyme 蒸 *əŋ corresponds to *o* in Japhug, after merger with *aŋ.

1. 冰 *piŋ* < B/S *p.rəŋ 'ice', Cogtse (Situ) *ta-rpâm* 'ice', Brag-bar (Situ) *ta-rpâm* 'ice', Japhug *tɤ-jpɣom* 'ice', (Schuessler 2007, 168). The -ɣ- medial of the Japhug etymon comes from the velarized vowel *aᵞ in Proto-Gyalrong. The consonant clusters *rp-* in Situ and *jp-* in Japhug come from the Proto-Gyalrong *lp-. This comparison suggests an alternative reconstruction *rpəm with labial dissimilation (see section 4).

2. 夢 *mjuwŋH* < *muŋs B/S *C.məŋ-s 'dream', Tibetan *rmang-lam* རྨང་ལམ 'dream' (Schuessler 2007, 381), Cogtse (Situ) *ta-rmô* 'dream', Brag-bar (Situ) *ta-rmôk* 'dream', Japhug *tɯ-jmŋo* 'dream'. Like the other velar nasal ending rhymes in OC, the rhyme *uŋ also corresponds to a single vowel in Gyalrong languages. However, the -k coda of the Brag-bar etymon is likely to be secondary for two reasons. First, the cognate forms in other Gyalrong languages all end in an open syllable. Second, the denominal verb *ka-va-rmō* 'to dream' has no coda. The *r-* preinitial in Situ and *j-* preinitial in Japhug comes from the *l- preinitial of Proto-Gyalrong (*lm- > rm- in Situ, jm- in Japhug). In addition, the -ŋ- medial in Japhug is due to the velarized rhyme *lmaᵞŋ in Proto-Gyalrong (Jacques 2004, 44).

3. 乃 *nojX* < B/S *nˤə(ŋ)ʔ 'your', this word is the possessive form of 汝 *nyoX* < *naʔ 'you (SG)' (Schuessler 2007, 446). It is comparable to Cogtse (Situ) *nō* 'you', Brag-bar (Situ) *nə-ɟâ* 'you', Japhug *nɤ-ʑo* 'you'. The second element in the Brag-bar and Japhug etyma is the root of the reflexive pronoun *tɐɟâ* 'oneself' in Brag-bar and *tɯ-ʑo* 'oneself' in Japhug. The cognate pronoun in Brag-bar and Japhug etyma might occur in their constructed status.

4. 蠅 *jiŋ* < B/S *m-rəŋ 'fly', Japhug *ɣʑo* 'bee', Tibetan *sbrang* སྦྲང 'honey' < *smraŋ.

5.3.4. *iŋ

The rhyme *iŋ in OC generally corresponds to a single high vowel in Gyalrong languages, *i* in Japhug and *e* in Situ.

1. 薪 *sin* < B/S *[s]i[n] 'firewood', Tibetan *shing* ཤིང 'tree' (Schuessler 2007, 538–39), Cogtse (Situ) *ʃē* 'wood, firewood' Brag-bar (Situ) *ɕê* 'wood, firewood', Japhug *si* 'wood, firewood'.

2. 新 *sin* < B/S *s.tsʰi[n] 'be new', Cogtse (Situ) *kə-ʃīk* 'be new', Brag-bar (Situ) *kə-ɕək* 'be new', Japhug *ɕɤɣ* 'be new'. This series is also a case of a non-checked rhyme *iŋ in OC corresponding to a checked rhyme in Gyalrong languages.

3. 身 *ɕin* < B/S *ŋ̊i[ŋ] 'body, self', Cogtse (Situ) *tə-ʃnē* 'heart', Brag-bar (Situ) *tə-ɕnê* 'heart', Japhug *tɯ-sni* 'heart', Tibetan *snying* སྙིང 'heart'.

4. 田 *den* < B/S *lˤiŋ 'field', Tibetan *ʑiŋ-kha* ཞིང་ཁ 'farmland' (Schuessler 2007, 496), Japhug *tɯ-ji* 'field'. It is also related to Brag-bar (Situ) *tə-jē* 'farming' and Brag-bar (Situ) *ka-jê* 'to plant', the verb is derived by replacing the nominal prefix *tə-* by the dynamic infinitive prefix *ka-*. Backformation from the noun to the verb is also possible, though less likely.

5.3.5. 東 dong *oŋ

1. 撞 *ɖæwŋH* < B/S *[N-t]ˤ<r>oŋ-s 'strike', Japhug *ʁndu* 'to hit', Tibetan *rdung-ba* རྡུང་བ 'to beat'.

5.4. *-t* ending rhymes

5.4.1. 月 yue **at*

1. 殺 *ʂɛt* < B/S **s<r>at* 'to kill', Tibetan *bsad* བསད་ 'to kill' (Schuessler 2007, 452), Cogtse (Situ) *ka-sāt* 'to kill', Brag-bar (Situ) *ka-siɛt* 'to kill', Japhug *sat* 'to kill'. For the correspondence between **sr-* in OC and *s-* in Gyalrong and Tibetan etyma, see Jacques (2015a).
2. 敗 *pæjH* < B/S **pˤra[t]-s* 'to defeat', Cogtse (Situ) *kɐ-prêt* 'to break', Brag-bar (Situ) *ka-prāt* 'to break', Japhug *prɤt* 'to break'. 敗 also has an anticausative form *bæjH* < **N-pˤra[t]-s/*brats* 'suffer defeat', cognate forms are also found in Gyalrong languages, Cogtse (Situ) *kɐ-mbrêt* 'to break', Brag-bar (Situ) *kɘ-mbrāt* 'to break', Japhug *mbrɤt* 'to break'.

5.4.2. 月 yue **et*

1. 八 *pɛt* < B/S **pˤret* 'eight', Tibetan *brgjad* བརྒྱད་ 'eight' (Schuessler 2007, 152), Cogtse (Situ) *wurjat* 'eight', Brag-bar (Situ) *kɘrcēt* 'eight', Japhug *kɯrcat* 'eight'. The Cogtse form suggests (as in the case of 'hundred') that **pˤret* may have come from an earlier form such as **pV-rˤjat* with primary yod through loss of the vowel in the first syllable and monosyllabicization. Among Gyalrong languages, the Brag-bar and Japhug etyma have the irregular onset *rc-*, for the expected correspondence would be **rj- > rɟ-* and **rj- > rʑ-*, as in Zbu *vɘ-rɟêt* (Gong 2018, 130). The Tibetan etymon has undergone the fortition change **ry > rgy* (Li 1959; Hill 2011).

5.4.3. 月 yue **ot*

1. 脫 *thwat* < B/S **mɘ-l̥ˤot* 'to take off', Brag-bar (Situ) *ka-ɕlēt* 'to fall (from hand)' (stem I *ɕlēt*, stem II *ɕlôt*).
2. 刮 *kwæt* < B/S **[kʷˤ]r[a]t* 'to scrape', Cogtse (Situ) *ka-khrōt* 'to scratch', Brag-bar (Situ) *ka-khrāt* 'to scratch', Japhug *qhrɯt* 'to scratch'.
3. 掘 *gjut; gjwot* < B/S **[g]ot; *[g]ʷat* 'to dig out (earth)', Cogtse (Situ) *ka-sɘkū* 'to bury', Brag-bar (Situ) *ka-sɘkū* 'to bury', Japhug *skɯ* 'to bury', Tibetan *rko* རྐོ་ 'to dig'.

5.4.4. 質 zhi **it*

1. 躓/懫 **tr-lit-s** B/S *[t-l]ri[t]-s* 'to slip' (Schuessler 2007, 619), Japhug *aʁdɤt* 'slip'. Schuessler (2007, 619) also relates this word to Tibetan *'dred-pa* འདྲེད་པ 'to slip'.
2. 滅 *mjiet* < **mit* B/S **[m]et/* 'to destroy', Cogtse (Situ) *kɘ-rmēk* 'to extinguish', Brag-bar (Situ) *kɘ-māk* 'to extinguish', Japhug *mi* 'to extinguish'. The Japhug etymon has lost the final **-k > -ɣ* after the high vowel *i* Jacques (2004, 224). The correspondence between the final *-t* in OC and *-k* in Gyalrong languages is unexplained.
3. 七 *tshit* < B/S **[tsʰ]i[t]* (possibly **s-n̥it*) 'seven'. Schuessler (2007, 419) relates it to 'seven' in Gyalrong languages. Cogtse (Situ) *kɘʃnās* 'seven', Brag-bar (Situ) *kɘɕnēs* 'seven', Japhug *kɯɕnɯz* 'seven'.
4. 失 *ɕit* < B/S **l̥i[t]* 'to lose', Cogtse (Situ) *ka-ʃlɔ̂k* 'to fall (from hand)', Japhug *lɯɣ* 'to fall', Japhug *ɕlɯɣ* 'to let sth. fall without any attention'. The Gyalrong etyma could be evidence showing that the rhyme **-it* originates from **-ik* in this word (Baxter and Sagart 2014, 236). The preinitial *ɕ-* is the lexicalized causative prefix.

5. 痹 *pjijH* < *pits* 'stiff', Japhug *ndʑɯrpɯt* 'stiff (IDEO)'[6]. Syllable break of the Japhug etymon is unclear between *ndʑɯ.rpɯt* and *ndʑɯr.pɯt*.

5.5 *-n/-r* ending rhymes

5.5.1. 元 yuan **an/r*

The rhyme 元 **an* possibly corresponds to *a* in Gyalrong languages, but this cannot be confirmed until more cognates are found.

1. 顔 *ŋæn* < B/S **C.ŋˤrar* 'face' could be compared with Japhug *tɯ-rŋa* 'face', an etymology which would imply an alternative reconstruction **rŋan* in OC. Schuessler (2007, 551) proposes two other etymologies of this word: derivation from 御迓訝 *ngjoH* < **[ŋ](r)a-s* 'to meet' by suffixation of *-n* nominalizer, or cognate of Tibetan *ngar* ᄃᄃ 'front side'.

5.5.2. 元 **en/r*

1. 鮮 *sjen* < B/S **s[a]r* 'fresh', Tibetan *gsar-pa* གསར་པ 'fresh,new' (Schuessler 2007, 528), Brag-bar (Situ) *kə-tsâr* 'fresh', Japhug *sɤr* 'fresh'.

5.5.3. 元 **on/r*

1. 酸 *swan* < B/S **[s]ˤor* 'sour', Cogtse (Situ) *kə-tɕôr* 'be sour', Brag-bar (Situ) *kə-tɕôr* 'be sour', Japhug *tɕur* 'be sour'
2. 晚 *mjonX* < B/S **m[o][r]ʔ* 'late', Cogtse (Situ) *tə-mōr* 'night', Brag-bar (Situ) *tə-mēr* 'night' < **tə-mōr*, Brag-bar (Situ) *rə-môr* 'one night', Japhug *tɯ-ɣmɯr* 'night'. The *ɣ-* preinitial in Japhug etymon comes from wrong segmentation from the numeral prefix 'one', as is the case in Japhug *tɯ-xpa* 'one year' (Jacques 2017b).

5.5.4. 文 wen **ən/r*

1. 蚓 *jinX* < B/S **[ɢ](r)ə[r]ʔ* 'earthworm' (Schuessler 2007, 574), Japhug *qa-ndʑe* 'earthworm', Tibetan *srin* སྲིན 'worm'. Japhug *-ndʑe* comes from Proto **nre-* (Jacques 2004, 295).

5.5.5. 文 wen **un/r*

1. 昏 *xwon* < B/S **m̥ˤu[n]* 'dusk, dark', Japhug *tɯrmɯ* 'afternoon', Tibetan *mun-pa* མུན་པ 'darkness', Tibetan *mun-po* མུན་པོ 'dark'.
2. 熏 *xjun* < B/S **qʰu[n]* 'to smoke', Cogtse (Situ) *tɐ-khâ* 'smoke', Brag-bar (Situ) *ta-khû* 'smoke', Japhug *tɤ-khɯ* 'smoke'.

5.6. *-p* ending rhymes

5.6.1. 緝 ji **əp* **ip*

1. 立 *lip* < B/S **k.rəp* 'to stand', Cogtse (Situ) *ka-rjāp* 'to stand', Brag-bar (Situ) *ka-rʑēp* 'to stand'.

---

[6] Etymology suggested by Gong Xun.

2. 汲 *kip* < *kəp* 'to scoop water', Brag-bar (Situ) *ka-kiɛ̂p* 'to carry water on back', Japhug *kaβ* 'to carry water on back'. Brag-bar (Situ) *sa-kāp* 'well', Japhug *sakaβ* 'well' are lexicalized locative participles, formed by adding the oblique participant nominalizer *sV-* (Sun 2006; Jacques 2016c).

5.7 *-m* ending rhymes

5.7.1. 侵 qin **əm *um*

1. 頜 *ɣomX* < B/S **[ɢ]ˤ[ə]mʔ* 'jaw, chin', Japhug *ta-mɢom* 'pliers'.
2. 枕 *tɕimX* < B/S **[t.k][ə]mʔ* 'pillow', Brag-bar (Situ) *tə-mk�âm* 'pillow', Japhug *tɤ-mkɯm* 'pillow'. The *t-* preinitial in OC could be related to the indefinite possessive prefix in Gyalrong languages.
3. 三 *sam* < **ˤsəm* B/S **s.rum* 'three', Tibetan *gsum* གསུམ 'three' (Schuessler 2007, 449), Cogtse (Situ) *kəsâm* 'three', Brag-bar (Situ) *kəsɔ̂m* 'three', Japhug *χsɯm* 'three'.
4. 參 *ʂim* < **srəm* B/S **s.rum* 'plant root', Cogtse (Situ) *tɐ-srêm* 'root', Brag-bar (Situ) *ta-srâm* 'root', Japhug *tɤ-zrɤm* 'root' (Jacques 2015a).
5. 覃 *dom* < **ˤləm* B/S **N.rˤ[o]m* 'to extend, spread', Brag-bar (Situ) *rə-kcâm* 'length of two arms', Japhug *tɯ-ɟom* 'length of two arms', Tibetan *ndom-pa* འདོམ་པ 'armspan'. The initial consonant *ɟ-* of the Japhug etymon may have come from the Proto **lj-*, and *nd-* in the Tibetan etymon from **nl-* (Jacques 2004, 148). The preinitial *k-* in the Brag-bar etymon remains unexplained.
6. 針 *tɕim* < B/S **t.[k]əm* 'needle', Cogtse (Situ) *ta-kāp* 'needle', Brag-bar (Situ) *ta-wiɛ̂p* 'needle', Japhug *ta-qaβ* 'needle', Tibetan *khab* ཁབ 'needle'. The preinitial **t-* in OC could be related to the indefinite possessive in Gyalrong languages. The *w-* initial in Brag-bar etymon is due the **k- > w-* lenition.
7. 陰 *ʔim* < B/S **q(r)um* 'be cloudy, dark', Brag-bar (Situ) *ta-ncâp* 'dark side of the mountain', Japhug *ɴqiaβ* 'dark side of the mountain'. The consonant cluster with a uvular initial and *-j-* medial in Northern Gyalrong languages corresponds to the palatal initial in Situ, Proto-Gyalrong **uvular-j-* has been palatalized in Situ Jacques (2004, 309).
8. 熊 *hjuwŋ* < B/S **C.[ɢ]w(r)əm* 'bear', Tibetan *dom* དོམ 'bear' (Schuessler 2007, 542), Cogtse (Situ) *tɔwām* 'bear', Brag-bar (Situ) *tɔwām* 'bear'. The Tibetan etymon comes from **dwam*, **wa* could have been monophthongized to *o* (Jacques 2009) as in Japhug **ɕwa > ɕɣa* 'tooth' and Tibetan **swa > so* སོ 'tooth'.

5.8. *-w* ending rhymes

Gyalrong languages lack *-w* coda in native words.

5.8.1. 幽 you **iw*

The rhyme 幽 **iw* corresponds to a non-front vowel in Gyalrong languages, *ə* in Cogtse (Situ), *u* in Brag-bar (Situ) and *ɯ* in Japhug.

1. 梟 *kew* < **ˤkiw* 'owl' (Schuessler 2007, 535), Cogtse (Situ) *pka-khə̄* 'owl', Brag-bar (Situ) *pow-khū* 'owl', Japhug *pɣɤ-khu* 'owl'. The first syllable of the Gyalrong etyma is the word designating 'bird', Cogtse (Situ) *pka-tʃû* 'chicken', Brag-bar (Situ) *pa-tɕû* 'chicken', Japhug *pɣɤ-tɕɯ* 'bird'.

5.8.2. 宵 xiao *aw

1. 熬 ŋaw < B/S *ŋˤaw 'to fry, roast', Tibetan rngo ཞོ 'to fry' (Schuessler 2007, 151), Cogtse (Situ) ka-rŋô 'to fry', Brag-bar (Situ) ka-rŋô 'to fry' (Stem I rŋô, stem II rŋē, stem II' rŋō), Japhug rŋu 'to fry'. The Gyalrong forms could also be loanwords from Tibetan.
2. 号 ɣawH < B/S *[g]ˤaw-s 'command', 號 haw < B/S *[C.g]ˤaw 'call out'.. It is comparable to Cogtse (Situ) ka-ŋa-khō 'to shout, call', Brag-bar (Situ) kə-okhō 'to shout, call', Japhug akhu 'to shout, call'.

5.9. -j ending rhymes

5.9.1. 歌 ge *aj

1. 羆 pje < B/S *praj 'brown-and-white bear', Cogtse (Situ) prī 'Asiatic brown bear', Brag-bar (Situ) prī 'Asiatic brown bear', Japhug pri 'Asiatic brown bear'.

5.9.2. 脂 zhi *ij

1. 屎 ɕijX < B/S *[qʰ]ijʔ 'excrement', Japhug tɯ-qe 'excrement'. An alternative reconstruction *lhijʔ is proposed by Schuessler (2007, 465), it is comparable to Tibetan lci-ba ལྕི་བ 'excrement' < PT *lhyi, and Japhug tɯ-ɣli 'excrement' < Proto-Gyalrong *klij (Jacques 2004, 313).
2. 米 mejX < *mijʔ B/S *(C.)mˤ[e]jʔ 'millet or rice grains', smai-khrī "小米" (Huáng and Sūn 2002, 550) . The second element of this etymon is the widespread word of 'rice' in Situ, Cogtse (Situ) khrī 'rice', Brag-bar (Situ) khrī 'rice', also found in Japhug khrɯ-zwa 'cooked rice'. This word is related to Tibetan khre ཁྲེ 'millet'.
3. 矢 B/S *l̥i[l]ʔ 'arrow', Japhug tɯ-di 'arrow'. The voice initial consonant d- in Japhug might come from *ld- in Proto-Gyalrong (Jacques 2004, 313).
4. 死 sijX < B/S *sijʔ 'to die', Tibetan shi ཤི 'to die' (Schuessler 2007, 478), Cogtse (Situ) ka-ʃî 'to die', Brag-bar (Situ) kə-ɕî 'to die', Japhug si 'to die',
5. 妣 pjijX < B/S *pijʔ-s 'deceased mother', Schuessler (2007, 162) relates this word to Tibetan phyi-mo ཕྱི་མོ 'grandmother', which according to Benedict (1942) is the common Tibeto-Burman root for grandmother. The Gyalrong words Cogtse (Situ) ta-wî 'grandmother', Brag-bar (Situ) ta-wî 'grandmother', Japhug tɤ-wi 'grandmother' are possible cognates.
6. 畀 pjijH < *pij-s B/S *pi[t]-s 'to give', Cogtse (Situ) kɐ-wû 'to give', Brag-bar (Situ) ka-mbī 'to give', Japhug mbi 'to give', Tibetan sbyin སྦྱིན 'to confer'. The initial consonant of the Cogtse etymon presents *b- > w- lenition.
7. 二 ɲijH < B/S *ni[j]-s 'two', Tibetan gnyis གཉིས 'two' (Schuessler 2007, 226–27), Cogtse (Situ) kənês 'two', Brag-bar (Situ) kənâs 'two', Japhug ʁnɯz 'two'.

5.9.3. 微 wei *əj

The rhyme 微 *əj *uj corresponds to a high/mid-high vowel in Gyalrong languages, i/e in Cogtse (Situ), i/e/ɐj in Brag-bar (Situ) and i/e in Japhug.

1. 眉 mij < /*mrəj B/S *mr[ə][r] 'eyebrow' (Schuessler 2007, 377), Cogtse (Situ) ta-rɲê 'hair', Brag-bar (Situ) ta-rɲê 'hair', Japhug tɤ-rme 'hair'. This root is also found in compound nouns Brag-bar (Situ) ta-wɐ-rɲê 'hair', Japhug tɯ-kɤ-rme 'hair', the first syllable is the constructed status of Brag-bar (Situ) ta-wô 'head', Japhug tɯ-ku 'head'.

2. 尾 *mjɨjX* < B/S *[m]əjʔ* 'tail', Cogtse (Situ) *ta-jmî* 'tail', Brag-bar (Situ) *ta-jmî* 'tail', Japhug *tɤ-jme* 'tail'. The *j-* (*jm-*) preinitial in the Gyalrong etyma comes from the **l-* (**lm-*) preinitial before labial initials (Jacques 2004, 271), which has no equivalent in OC.

3. 邇 *nyeX* < B/S *n[ə][r]ʔ* 'near'. Schuessler (2007, 226) relates this word to Tibetan *nye-ba* ཉེ་བ 'near' and Tibetan *snyen-pa* སྙེན་པ 'to come near'. The verb root itself is not attested in Gyalrong languages, but the noun Japhug *tɯ-ɣni* 'friend' is a possible cognate of Tibetan *gnyen* གཉེན 'friend, relative', a noun derived from the verb root by the *g-...-n* circumfix (Jacques 2018).

4. 火 *xwaX* < **m̥əjʔ* (Schuessler 2007, 290–91) B/S *[qʷʰ]ˤəjʔ* 'fire'. The phonetic evolution of this word from OC is irregular, as MC *-wa* normally comes from OC **-aj* or **-oj*. The word 火 rhymes as **-aj* in the *Shijing*, as in 七月流火，九月授衣 (« 豳風 · 七月 »). Schuessler (2007, 290–91) relates this word to Tibetan *me* མེ 'fire' (see also Hill 2013 on an alleged spelling of this word in Tibetan cited by some Sinologists). Cognate is found in Japhug *smi* 'fire'. This word could originally be a compound **sɯ-mi*, in which *sɯ* is the construct status of Japhug *si* 'firewood', and cannot be used as evidence for a cluster **sm-* in Old Chinese.

5.9.4. 微 wei **uj*

1. 虺 *xwojX* < **ˤhməjʔ* B/S *[r̥]u[j]ʔ* 'snake brood; sound of thunder' (Schuessler 2007, 287), Cogtse (Situ) *kha-brē* 'snake', Brag-bar (Situ) *kha-prēj* 'snake', Japhug *qa-pri* 'snake', Tibetan *sbrul* སྦྲུལ 'snake' < **smrul* (Jacques 2004, 137). There is no equivalent of *s-* preinitial of the Tibetan etymon in OC and Gyalrong languages. The Gyalrong etymon for 'snake' could also be possible cognate of 巴 *pæ* < **pˤra* 'snake'.

5.10. Wanderwörter

1. 菽[7] *syuwk* < B/S **s-t(ʰ)uk* 'pulse, beans', Cogtse (Situ) *tɐ-stōk* 'broad bean', Brag-bar (Situ) *ta-stēk* 'broad bean', Japhug *stoʁ* 'broad bean'. The Brag-bar etymon has undergone the **o > ɐ* sound change. Despite the regular phonetic correspondence between OC and Gyalrong words, (Sagart 1999, 185–88) points out that the bean has been domesticated too recently to be a cognate.

2. 馬 *mæX* < B/S **mˤraʔ* 'horse', Cogtse (Situ) *mbrō* 'horse', Brag-bar (Situ) *mbrō* 'horse', Japhug *mbro* 'horse'. Sagart (1999, 196) suggests that 馬 could be an early loanword from TB, after the loss of vowel nasalization of **mraŋ* or **mrã* (also mentioned in Schuessler 2007, 373).

3. 兔 *thuH* < B/S **l̥ˤa-s* 'rabbit, hare', Cogtse (Situ) *ka-lā* 'rabbit', Brag-bar (Situ) *ka-liɛ̄* 'rabbit', Japhug *qa-la* 'rabbit' (Schuessler 2007, 502). The first syllable in Gyalrong etyma is the lexicalized nominal prefix designating animals (Jacques 2008, 52–54).

References

Baxter, William H. 1992. *A Handbook of Old Chinese Phonology*. Berlin: Mouton De Gruyter.

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

Benedict, Paul K. 1942. Tibetan and Chinese Kinship Terms. *Harvard Journal of Asiatic Studies* 6 (3/4): 313–37.

---

[7] These Wanderwörter have related words in many other branches of ST/TH and beyond, and a full examination of the complete dataset is beyond the scope of this work.

Bodman, Nicholas C. 1969. Tibetan *Sdud* 'Folds of a Garment', the Character 卒, and the **st*-Hypothesis. *The Bulletin of the Institute of History and Philology* 39: 327–46.

Burrow, Thomas. 1972. A Reconsideration of Fortunatov's Law. *Bulletin of the School of Oriental and African Studies* 35 (3): 531–45.

Coblin, Weldon South. 1986. *A Sinologist's Handlist of Sino-Tibetan Lexical Comparisons*. Steyler Verlag.

Conrady, August. 1896. *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Ton-accenten*. Leipzig: Otto Harrassowitz.

Davids, Irene Maria Hendrina, George van Driem. 1985. Limbu Kinship Terminology: A Description. *Himalayan Languages* XII (1-2): 115–56.

Dixon, R. M. W., Alexandra Y. Aikhenvald (eds.). 2000. *Changing Valency*. Cambridge University Press.

Gong, Hwang-cherng. 1995. The System of Finals in Proto-Sino-Tibetan. In: William S.-Y. Wang (ed.). *The Ancestry of Chinese*: 41–92. Journal of Chinese Linguistics Monograph Series.

Gong, Xun. 2017. Jiarong yuzu yuyan yu shanggu hanyu de xingtaixue (嘉绒语组语言与上古汉语的形态学). In: Haeree Park, Shaoxuan Cheng (eds.). *Gu wenzi yu hanyu lishi yinyunxue* (古文字与汉语历史音韵学): 134–56. Shanghai: Fudan daxue chubanshe (复旦中华文明研究专刊).

Gong, Xun. 2018. *Le rgyalrong zbu, une langue tibéto-birmane de Chine du Sud-ouest. Une étude descriptive, typologique et comparative.* PhD thesis: Institut National des langues et des civilisations orientales.

Handel, Zev. 2002. Rethinking the Medials of Old Chinese: Where Are the R's? *Cahier de Linguistique Asie Orientale* 31 (1): 3–32.

Haspelmath, Martin. 1993. More on the Typology of Inchoative/Causative Verb Alternations. In: Bernard Comrie, Maria Polinsky (eds.). *Causatives and Transitivity*: 80–120. Amsterdam: John Benjamins.

Hill, Nathan. 2011. An Inventory Of Tibetan Sound Laws. *Journal of the Royal Asiatic Society* 21 (4): 441–57.

Hill, Nathan. 2013. Old Chinese **Sm-* and the Old Tibetan Word for Fire. *Cahier de Linguistique Asie Orientale* 42 (1): 60–71.

Huáng, Liángróng, Hóngkāi Sūn. 2002. *Han Jiarong cidian* (漢嘉戎詞典). Beijing: Minzu chubanshe.

Jacques, Guillaume. 2004. *Phonologie et Morphologie du Japhug (rGyalrong)*. PhD thesis, Paris: Université Paris VII – Denis Diderot.

Jacques, Guillaume. 2005. *Jiarongyu yu shanggu hanyu* (嘉绒语与上古汉语). Hanyu shangguyin guoji yantaohui (汉语上古音国际研讨会): December 2005. Ms., available online at: https://halshs.archives-ouvertes.fr/halshs-00137451/document.

Jacques, Guillaume. 2008. *Jiarongyu yanjiu* (嘉绒语研究). Beijing: Nationalities Press.

Jacques, Guillaume. 2009. Tibetan Wa-Zur and Laufer's Law. *Linguistics of the Tibeto-Burman Area* 32 (1): 141–44.

Jacques, Guillaume. 2012. The Tangut Kinship System in Qiangic Perspective. In: Nathan W. Hill (ed.). *Medieval Tibeto-Burman Languages*: 211–256. Leiden: Brill.

Jacques, Guillaume. 2013. On Pre-Tibetan Semi-Vowels. *Bulletin of the School of Oriental and African Studies* 76 (2): 289–300.

Jacques, Guillaume. 2015a. On the Cluster **Sr-* in Sino-Tibetan. *Journal of Chinese Linguistics* 43 (1): 215–23.

Jacques, Guillaume. 2015b. The Origin of the Causative Prefix in Rgyalrong Languages and Its Implication for Proto-Sino-Tibetan Reconstruction. *Folia Linguistica Historica* 36: 165–98.

Jacques, Guillaume. 2015c. The Spontaneous-Autobenefactive Prefix in Japhug Rgyalrong. *Linguistics of the Tibeto-Burman Area* 38 (2): 271–91.

Jacques, Guillaume. 2016a. How Many **-S* Suffixes in Old Chinese? *Bulletin of Chinese Linguistics* 9 (2): 205–17.

Jacques, Guillaume. 2016b. Le Sino-Tibétain: Polysynthétique Ou Isolant? *Faits de Langues* 47 (1): 61–74.

Jacques, Guillaume. 2016c. Subjects, Objects and Relativization in Japhug. *Journal of Chinese Linguistics* 44 (1): 1–28.

Jacques, Guillaume. 2017a. La Nasale Intrusive En Limbou. *Panchronica* (available online at: https://panchr.hypotheses.org/1654).

Jacques, Guillaume. 2017b. The Morphology of Numerals and Classifiers in Japhug. In: Picus Sizhi Ding, Jamin Pelkey (eds.). *Sociohistorical Linguistics in Southeast Asia*: 135–48. Leiden: Brill.

Jacques, Guillaume. 2018. Fossil Nominalization Prefixes in Tibetan and Chinese. In: *LFK Society Young Scholar Symposium.* Taipei.

Kümmel, Martin J. 2007. *Konsonantenwandel: Bausteine zu einer Typologie des Lautwandels und ihre Konsequenzen für die vergleichende Rekonstruktio*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Lai, Yunfan. 2017. *Grammaire Du Khroskyabs de Wobzi*. PhD thesis, Université Sorbonne Nouvelle – Paris 3.

Li, Fang-Kuei. 1959. Tibetan Glo-Ba-'Dring. In: Søren Egerod, Else Glahn (eds.). *Studia Serica Bernhard Karlgren Dedicata*: 55–59. Copenhagen: Ejnar Munksgaard.

List, Johann-Mattis. 2016. Beyond Cognacy: Historical Relations Between Words and Their Implication for Phylogenetic Reconstruction. *Journal of Language Evolution* 1 (2): 119–36.

Miyake, Marc Hideo. 2012. Complexity from Compression: A Sketch of Pre-Tangut. In: Irina Popova (ed.). *Tanguts in Central Asia: A Collection of Articles Marking the 80th Anniversary of Prof. E. I. Kychanov*: 244–61. Moscow: Vostochnaja Literatura.

Nagano, Sadako. 1994. A Note on the Tibetan Kin Terms Khu and Zhang. *Linguistics of the Tibeto-Burman Area* 17 (2): 103–15.

Peiros, Ilia, Sergei Starostin. 1996. *A Comparative Vocabulary of Five Sino-Tibetan Languages*. Melbourne: University of Melbourne, Department of Linguistics.

Prins, Marielle. 2016. *A Grammar of RGyalrong Jiǎomùzú (Kyom-Kyo) Dialects.* Leiden: Universiteit Leiden.

Pulleyblank, Edwin G. 1981. Some Notes on Chinese Historical Phonology. *Bulletin de L'École Française d'Extrême-Orient* 69: 277–88.

Pulleyblank, Edwin G. 2000. Morphology in Old Chinese. *Journal of Chinese Linguistics* 28 (1): 26–51.

Sagart, Laurent. 1999. *The Roots of Old Chinese*. John Benjamins Publishing Company.

Sagart, Laurent. 2017. A Candidate for a Tibeto-Burman Innovation. *Cahiers de Linguistique Asie Orientale* 46: 101–119.

Sagart, Laurent, William H. Baxter. 2012. Reconstructing the *S- Prefix in Old Chinese. *Language and Linguistics* 13 (1): 29–59.

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill, Johann-Mattis List. 2019. Dated Language Phylogenies Shed Light on the Ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences of the United States of America* 116 (21): 10317–22.

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

Simon, Walter. 1929. Tibetisch-Chinesische Wortgleichungen, ein Versuch. *Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich-Wilhelms-Universität zu Berlin* 32: 157–228.

Starostin, Sergei A. 1989. *Rekonstrukcija Drevnekitajskoj Fonologičeskoj Sistemy [Reconstruction of the Phonological System of Old Chinese]*. Moscow: Nauka.

Sun, Jackson T.-S. 2006. *Caodeng jiarongyu de guanxiju* (草登嘉戎語的關係句). *Language and Linguistics* 7 (4): 905–33.

Sun, Jackson T.-S, Youjing Lin. 2007. Constructional variation in rGyalrong relativization: how to make a choice? In: *Pre-Conference Proceedings of the International Workshop on Relative Clauses*: 205–226. Taipei: Institute of Linguistics, Academia Sinica.

Zhang, Shuya. 2016. *La morphologie et la phonologie du dialecte de Brag-dbar du rgyalrong situ*. Master's thesis: Institut National des langues et des civilisations orientales.

Zhang, Shuya. 2018. Stem Alternations in the Brag-bar Dialect of Situ Rgyalrong. *Linguistics of the Tibeto-Burman Area* 41 (2): 294–330.

Чжан Шуя, Гийом Жак, Лай Юньфань. К вопросу о когнатах между гьярунгскими языками и древнекитайским

Гьярунгские языки, составляющие подгруппу бирмано-цянской ветви сино-тибетской семьи, распространены на западе китайской провинции Сычуань. Они относятся к полисинтетическому типу и обладают богатыми системами глагольной морфологии. Хотя гьярунгские языки не демонстрируют близкого родства с китайским языком, они, тем не менее, представляют большой интерес для сравнительного изучения синотибетских (транс-гималайских) языков из-за архаичной природы их фонологической и морфологической систем. В данной статье, опираясь на предшествующие исследования в области древнекитайской фонологии и сопоставляя их с результатами новейших полевых данных, мы пытаемся показать, каким образом гьярунгские языки могут пролить свет на особенности древнекитайской морфологии и помочь тем самым скорректировать реконструкцию древнекитайского языка. В статье также приводится список возможных когнатов между древнекитайским и гьярунгскими языками, включая тибетские параллели там, где они обнаруживаются.

Ключевые слова: гьярунгские языки, древнекитайский язык, этимологические когнаты, сравнительная морфология, историческая реконструкция.