# 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 archaeo-logical 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 im-portance of the Sinitic subgroup, focusing on what they consider the *core* of the language fam-ily, *Tibeto-Burman* (Matisoff 2015). Other scholars emphasize the limits of our current knowl-edge with respect to the detailed affiliation of the numerous subgroups, demanding an unbi-ased 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 inten-sive 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
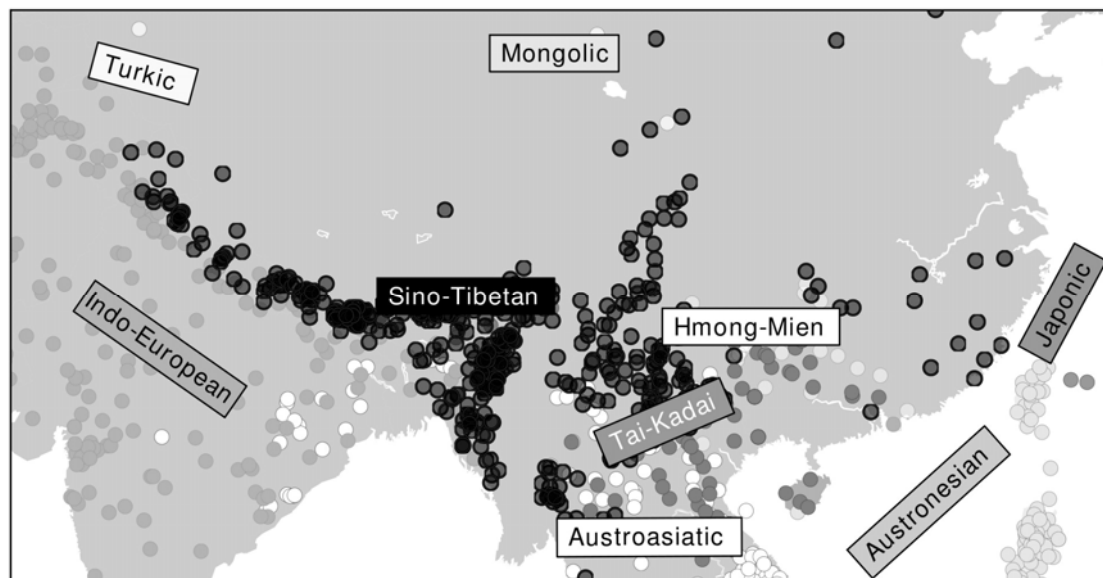
Sino-Tibetan and neighboring language families

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 Hawaiʻi 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.