Georg Orlandi
Osaka University; Italian School of East Asian Studies, Kyoto; thegaygenius@gmail.com

Once again on the history and validity of the Sino-Tibetan bifurcate model

Sino-Tibetan is the hypothesis that postulates a bifurcate genetic relationship between Sinitic and Tibeto-Burman. The history of the subgrouping of its components as well as its overall membership are briefly though exhaustively summarized. Special attention is given to the methodological shortcomings of a recent Bayesian approach that further professes to substantiate this bifurcate Sino-Tibetan model. Future directions and broader affinities of Trans-Himalayan are also discussed.

Keywords: Trans-Himalayan family; Sino-Tibetan family; Bayesian inference; language subgrouping.

1. Introduction

The Trans-Himalayan family (also referred to as Sino-Tibetan or Tibeto-Burman) is a very large linguistic family, comprising around 600 languages (Turin 2007, Owen-Smith and Hill 2014) spoken across a wide geographic area. It is the second populous linguistic family of the world, with nearly 1.4 billion speakers (second only to Indo-European), and one of the most diversified families, encompassing languages with very different morphological systems, from isolating languages such as Chinese and Burmese, to polysynthetic languages such as rGyalrongic or Kiranti. 'Sino-Tibetan' is the hypothetical linguistic family postulating a bifurcate genetic relationship between Sinitic and all other Trans-Himalayan languages, which are thus grouped under a "Tibeto-Burman" rubric.

Today the existence of a Sino-Tibetan/Trans-Himalayan family, whatever be the subgrouping of its components, is widely accepted among linguists (Greenberg 1996: 134; LaPolla 2001: 225). Since the beginning of the 21st century, a number of publications, including scientific articles, books and monographs have been dedicated to the subject, e.g. Ting and Sūn (2000, 2001), van Driem (2001), Beckwith (2002),[1] Gong (2002), Matisoff (2003), Thurgood and LaPolla (2003), Saxena (2004), Sagart *et al.* (2005), Owen-Smith and Hill (2014), Hill (2019), and so on, and the field is becoming more well-covered. Nevertheless, disagreement still persists on some crucial aspects of Sino-Tibetan/Trans-Himalayan, especially for what regards its overall membership, as well as the subgrouping of its components. While some scholars accept a bifurcate Sino-Tibetan family (Bradley 1997; LaPolla 2002; Thurgood and LaPolla 2003; Matisoff 2003), others contend that no serious linguistic evidence which may warrant the separation of Sinitic from the rest of the family has been shown (van Driem 2001, 2005, 2007, 2011; DeLancey 2014). Thus, they have suggested renaming the family with the more agnostic term 'Trans-Himalayan' (van Driem 2014; Owen-Smith and Hill 2014). The present survey article provides an overview of both historical and recent works in Sino-Tibetan/Trans-Himalayan linguistics, with an eye toward a few issues which may be of particular interest to both specialists and non-specialists of this branch of linguistics. But before discussing in more detail these

---

[1] This author, however, contends that Sinitic and Tibeto-Burman are not related. For ST discussions, see in particular Bradley (2002: 73–112) and Beckwith (2002: 113–157).

aspects, a remark seems to be necessary. The existence of Sino-Tibetan/Trans-Himalayan has not yet been proven (Handel 2008: 423), because it cannot, and, by the same token, it cannot be disproved. Just like any other early linguistic relationship, including those which are universally or almost-universally accepted such as Indo-European, Uralic and Semitic, Sino-Tibetan/ Trans-Himalayan, too, necessarily remains confined into the realm of non-demonstrable hypotheses. Some of these hypothetical, now-lost *Ursprachen* are, indeed, well-documented, whereas others are, instead, less-thoroughly explored. The question, nonetheless, is not whether they may or may not be definitely proven (cf. Handel 2008: 423)—because they cannot—but instead how convincing is the evidence that may be brought in support of these specific hypotheses. Although this statement may sound trivial, as it should be well-known to any serious worker in the field of historical linguistics, there are specialists who, occasionally, seem to be unaware of this fact.

2. The historical classification of Sino-Tibetan

The history of the Sino-Tibetan hypothesis has been illustrated in a number of publications by George van Driem (1997 et seq.), which are perhaps the most informative sources on the argument. This section, however, does not constitute a recapitulation of van Driem's detailed history of the Western linguistic development of the hypothesis of a Sino-Tibetan language family. While a section of this length is necessarily incomplete, the present paragraph demonstrates that the history of Sino-Tibetan is more controversial, debated and complex than usually assumed. According to van Driem, workers in Sino-Tibetan divide in two types: those who work in the received tradition of John Leyden's Indo-Chinese paradigm, which he believes to be what most specialists now refer to as Sino-Tibetan, and those who operate within the framework of Julius Klaproth's Tibeto-Burman. Van Driem reckons Sinitic as one of the many branches of Tibeto-Burman and therefore he still refers to the family as a whole as Tibeto-Burman (or Trans-Himalayan) instead of Sino-Tibetan. Van Driem's arguments in favour of a Tibeto-Burman/Trans-Himalayan family chiefly rely on three points: (i) Chinese and Limbu, a Kiranti language spoken in Nepal, exhibit certain vestiges of early morphological connections (van Driem 1997: 463–471). He also adduced lexical evidence which further purports to prove the existence of this early, now-lost "Sino-Bodic" connection (van Driem 1997: 471–484); (ii) from a historical point of view, the alleged bifurcation of Sino-Tibetan was nothing else than a mere scholastic artefact. Chinese appeared to be sensibly different from other Tibeto-Burman languages only because early reconstructions of Chinese, especially those championed by the great Swedish Sinologist Bernhard Karlgren (1889–1978) were methodologically flawed and thus unable to reveal many of its morphophonemic characteristics (see below for a more detailed discussion); (iii) extra-linguistic evidence, such as archaeology and the history of migrations, seems to support his subgrouping (van Driem 2001, 2007; cf. Handel 2008: 429). Other scholars have dismissed van Driem's subgrouping of Sino-Tibetan as an example of 'neosubgroupitis' (cf. Matisoff 2000); still others have rejected the arguments in favour of both Sino-Bodic and Trans-Himalayan because they find them "less than convincing, due to problems of fact and argumentation" (LaPolla 2016). Although certain problems are recognized,[2] it is not necessarily mistaken to agree with van Driem on two fundamental

---

[2] The present writer chiefly disagrees on the following points: (i) no matter what is our opinion of Heinrich Julius von Klaproth (1783–1835), Klaproth's work is merely of literary interest, and does not indicate an extraordinary foresight. Klaproth worked mainly in the received strand of nineteenth century linguistics, with its particular

points: (i) the bifurcate Sino-Tibetan model was never anything more than a mere impressionistic label; (ii) had more refined reconstructions of the sound system of Old Chinese been available to Benedict and other pioneers, they probably would not have outlined a bifurcate Sino-Tibetan phylum.

We start from the second point. For more than forty years, Karlgren's reconstructions have been almost the sole reference available in Europe, as well as in East Asia and the United States. However, Karlgren's works were more historical than comparative. In fact, like any other before (and perhaps even after) him, Karlgren made use of the "comparative method," a method which he openly distrusted (Karlgren 1915–1926: 45), only when the identification of the phonetic value of a certain sound class was not otherwise knowable. Furthermore, being a severe opponent of phonemics, he apparently lost himself into the triviality of reconstructing as many phonetic contrasts as the number of sound classes of medieval Chinese rime tables required. This led him to reconstruct a system which was studded with contrasts that were hardly describable with the acoustic-articulatory attributes known from other living languages, including other Sino-Tibetan/Trans-Himalayan tongues. But, more relevant to our case is the fact that Karlgren (1931) embarked in an open polemic with Walter Simon (1893–1981), when he tried to compare Chinese with Written Tibetan in the attempt at finding the solution to certain problems pertaining the reconstruction of final consonants in Old Chinese. Karlgren did not deny that Chinese was a branch of Sino-Tibetan (which he called Indo-Chinese). Rather, he thought that the radical phono-semantic changes that had occurred in Chinese lexicon, as well as the huge time-depth that existed between Sinitic and its sister languages, rendered infeasible one-to-one comparisons between Chinese and Tibetan in Indo-European fashion, let alone the identification of regular sound laws. Hence, Karlgren urged to build a stronger knowledge for Chinese lexicon. This was the idea which lied behind Karlgren's emphasis on word families (*cízú* 詞族). Unfortunately, the complexity of Karlgren's approach to word families was such that later scholars, such as Robert Shafer (1967: 9), completely misunderstood Karlgren's real intentions, as it is clearly illustrated by Shafer's critique to Karlgren's alleged advocacy for a "relaxation" of *Lautgesetze*.

When Paul Benedict moved to Berkeley in 1938 to join Kroeber's Sino-Tibetan Philology project, Karlgren (1915–1926) was the only complete system of reconstructions available to him, and his reconstructed Old Chinese was visibly different from Tibeto-Burman. In fact, only several years after Karlgren's death, the structure of Old Chinese morphemes and roots began to be better understood, especially after Sagart (1993: 238, 255–257) criticized the absence of a framework to handle affixation processes in Baxter's system (1992). We cannot know whether Benedict and other pioneers would have accepted a bifurcate Sino-Tibetan family, had more refined reconstructions of Old Chinese been available to them, but there are good reasons to believe that they probably would not have outlined such a bifurcate scheme, if Karlgren's system, as many other systems reconstructed after him, had contained an *s- causative and denominative prefix (Mei 1985: 334–343, 1989; Baxter and Sagart 1998: 53; Schuessler 2007: 52–54) which could also form iteratives and intensives (Schuessler 2007: 53), an *m- prefix for the direction/causativity dichotomy (Takashima 1996; J. Sun 1993), a *-t suffix for transitivization (Benedict 1972: 98–102; Michailovsky 1985; van Driem 1988), and a nominalizing *-n suffix (LaPolla 1994; Jīn 1998; Sagart 1999 [but abandoned in later publications]; Schuessler 2007).

---

emphasis for the typological classification of languages. Although Klaproth's work may be regarded as better than the work of other scholars of his time, his classifications too were based on criteria that can no longer satisfy contemporary scientific requirements; (ii) the "fallen leaves" model of the family appears to imply the existence of a *Stammbaum* of which, however, we have no trace (cf. LaPolla 2016: 291).

The other point discussed by van Driem is that the establishment of a bifurcate Sino-Tibetan phylum has never been verified in accordance with serious linguistic methods, using accepted protocols of evaluation of results. Although we may agree on the fact that Sino-Tibetan linguistics is relatively young (cf. Matisoff 1991: 469), and that only Benedict and few other scholars may be credited for pioneering Sino-Tibetan linguistics, we must remark, however, that certain embryonic forms of Sino-Tibetan or Trans-Himalayan subgrouping can be traced to earlier times. As argued in this section, van Driem claims that workers in Sino-Tibetan linguistics divide in two groups: those who followed the Indo-Chinese tradition started by John Leyden (1775–1811), and those who instead still adhere to Julius von Klaproth's Tibeto-Burman model, which was allegedly "accepted by British scholars such as Forbes, Houghton and Cust" (van Driem 2007: 215). However, this dichotomy is not necessarily correct. Surely, the name 'Tibeto-Burman' was soon popularized by British scholars, such as Capt. Charles James Forbes (?–1879) and Robert Needham Cust (1821–1909), but Chinese was never explicitly included within this phylum. For instance, Robert Cust (1881), in an article presented at the Fourth International Congress of Orientalists in Florence (1878), grouped together some non-Aryan languages of 'East India.' He distinguished six language families: Dravidian, Kolarian, Tibeto-Burman, Khasian (now also part of Austroasiatic), Tai and Mon-Anam (both being Austroasiatic). He never mentioned Chinese, therefore we do not know whether in his opinion Chinese was part of that Tibeto-Burman family, or whether it was reckoned as a linguistic isolate. For what regards Forbes's 'Tibeto-Burman,' as explained by he himself, the term was nothing else but the convenient designation of a very wide family of languages which corresponded more or less to Max Müller's Gangetic and Lohitic 'classes' (Forbes, 1878:210). He mainly revised the previous Tibeto-Burman scheme outlined by the American missionary Francis Mason (1799–1874), who included, in this family, the dialects spoken by eleven different Burmese tribes, such as "Burmese, Arracanese, Mugs, Kanyans, Toungooers, Tavoyers, Yos, Yebains, Pyus, Kados and Danus" (Forbes 1878: 211). Forbes left only six of Mason's original eleventh tribes under his Burman sub-phylum, to which he added also Naga, Singpho and Karen. As we can see, Julius von Klaproth could not exercise a decisive influence over these two British scholars. Rather, it seems that Max Müller (1823–1900) and his Turanian family exerted a much wider influence.[3] Müller is credited with being the first to have outlined the existence of a Tibeto-Burman phylum (Grierson and Konow 1909: 13; Campbell and Poser 2008: 113). In fact, Müller's *Bothiya* class of languages includes many Trans-Himalayan idioms. Most of these languages had been extensively studied by Brian Houghton Hodgson (1801–1894), who also championed the Turanian hypothesis, and had pointed out similarities in the use of personal pronouns between languages of the Circassian area and rGyalrongic. Müller divided Bothiya into two groups, Lohitic and Sub-Himalayan (or Gangetic). They included languages such as Burmese, Tibetan, Garo, Naga, Bodo, Kachin, Kuki-Chin, Singpho, Mising and many other dialects of North Assam (Müller 1854: 97–126).

What most scholars have overlooked, however, is the fact that a Trans-Himalayan phylum had been described long ago by a member of the Indian Civil Service, Joseph Edward Lyall Brandreth (1823–1907). Writing in the *Journal of Royal Asiatic Society* (1878), Brandreth outlined the richest and most exhaustive classification of Trans-Himalayan languages. Brandreth consulted the many communications that Robert Cust had made to the Philological Society, as well as other works and reports written by Sir George Campbell (1824–1892) and Sir William Wilson Hunter (1840–1900). Brandreth's classification of Trans-Himalayan languages was also grounded on Müller's Turanian, and his very aim was to correct some deficiencies

---

[3] For more information about 'Turanian', see Orlandi (2020).

that the Turanian class contained (Brandreth 1878: 2). Brandreth divided the non-Indo-European (non-Aryan in his words) languages of India into six groups, *viz.* Dravidian, Kolarian, Tibeto-Burman, Khasi, Tai, and Mon-Anam. For what regards Brandreth's classification of Tibeto-Burman languages, he further divided this huge family into nineteen different classes on the basis of mutual intelligibility and other grammatical features. That this was a Trans-Himalayan classification, rather than a Tibeto-Burman one, is highlighted by the fact that although the family was referred to as 'Tibeto-Burman,' Brandreth collocated Tibetan, together with Sherpa and other dialects spoken across Sikkim and Nepal, under the second class, whereas Burmese stood as an isolate in his sixteenth class. Although Brandreth recognized certain phonetic parallelisms between the two languages, he kept them separated because "Burmese has not the determinative syllables or letters of the Tibetan, for which the tones are the substitutes," and also because "[t]he Burmese verb has no person endings, but it has a plural suffix *kra*" (Brandreth 1878: 22). Verbal aspects, plus other features related to word order were reckoned by Brandreth as diagnostic for the sub-classification of Tibeto-Burman (i.e. Trans-Himalayan) languages.

As we can see, there was no pure dichotomy between Indo-Chinese and Tibeto-Burman (perhaps the only dichotomy was between monophyletic vs polyphyletic origin of Asian languages). These terminologies were often imprecise and covered different types of linguistic families. Furthermore, claiming that a dichotomy was likewise existing between followers of Leyden's Indo-Chinese and Klaproth's Tibeto-Burman would require us to ignore both Brandreth's Trans-Himalayan (clearly a different type of hypothesis) and Müller's Bothiya. It is also unlikely that the racist views of Indo-Chinese workers were mostly concerned with typological classification and had little to do with genetic subgrouping, as claimed by van Driem (2004, 2007, 2014). It is true that, e.g., monosyllabism was often associated with backwardness, yet many influential scholars such as Jean-Pierre Abel-Rémusat (1788–1832) and Charles-Joseph de Harlez de Deulin (1832–1899) were unconvinced of this erroneous causality.[4] German scholars had rejected the monosyllabic view at least since Theophilus (Gottlieb) Siegfried Bayer (1703: 106), and, most importantly, Georg von der Gabelentz (1840–1893), in his article presented at the Fourth International Congress of Orientalists (1878), ventured into a *genealogical* classification of Indo-Chinese languages, i.e. not based on typological criteria. Gabelentz started his investigation from morphology. After having identified a resemblance in the morphology of Indo-Chinese languages, he sought to establish the *Lautverschiebungsgesetze* (phonetic displacement) between these tongues. In the attempt to obtain as reliable data on lesser known Indo-Chinese languages as possible, Gabelentz mainly worked on second-hand materials, especially on those studies which had been published some years earlier in the *Journal of the Indian Archipelago*. An intellectual step forward is remarked by the fact that he did not compare lexical items in their modern forms. For instance, when he compared the word for 'ear' in Chinese, Karen and Tibetan, he did not take the character 耳 in its Mandarin clothing but in its medieval Chinese reading *ni (Gabelentz 1881: 289, 292). Moreover, although Gabelentz observed a phonetic correspondence between Thai liquid consonant *l* and Cantonese *h* (Mandarin *f*), he also cautioned about some alleged cases of cognancy between Chinese, Thai and Hmong. Gabelentz also remarked that monosyllabism proved nothing, and was not a diagnostic feature for typological, let alone genealogical classification. To sum up, Gabelentz's *Indochinesisch* encompassed Chinese, Siamese (Thai), Shan, Kiranti, Lao, Miao (Hmong), Tibetan and Burmese. Some years later, Conrady (1896, 1915), like Trombetti (1905) and other proponents of the Indo-Chinese theory, grouped Tibeto-Burman from one side

---

[4] See, respectively, Abel-Rémusat (1825: 47, 60) and Harlez (1878: 67–88)

(whose subgroup also included a *nepalische und assamesische Gruppe*) and Sino-Siamese from the other. He studied extensively the grammar of these languages, especially of Chinese, and also recognized some important connections between certain types of tones and voiced initials.

However, van Driem (2007: 221–226, 2011) is probably correct when he concludes that the Indo-Chinese legacy still defines current Sino-Tibetan phylogeny, not because Sino-Tibetanists still accept the typological classification and/or the racial prejudices which were the run of the mill until the first half of nineteenth century, but because the Sino-Tibetan hypothesis still defines a Tibeto-Burman taxon which encompasses all non-Sinitic languages after the removal of Daic (wrongly reckoned as being a branch of Sino-Tibetan by Whitney (1867: 336), Kuhn (1889), Trombetti (1905), Finck (1909), and others). Van Driem's other critics argue that those who subscribe to the standard bifurcate Sino-Tibetan phylogeny certainly do not do so because of influence from nineteenth century typologists, but mainly because of Li Fang-kuei's analysis of the languages of China and also later comparative works by Benedict (1972, 1976) and Matisoff (1973, 1991, 2003) (LaPolla 2016: 288). However, Li Fang-kuei's Sino-Tibetan (1937) is a "macro Sino-Tibetan" taxon, encompassing also Kra-Dai (*ZhuàngDòngyǔ* 壯侗語) and Hmong-Mien (*MiáoYáoyǔ* 苗瑤語), in Gabelentz and Conrady fashion.[5] Lamenting that Western scholars had scarce familiarity with 'tones,' which he regarded as important as vowels and consonants, Li sought to explain the striking similarities between certain Kra-Dai, Hmong-Mien and Sinitic languages, by assuming that they were inherited from a common ancestor, and not by assuming that they reflect areal linguistics or other types of linguistic contacts.[6]

Furthermore, Robert Shafer (1955: 98–99, 1965: 445–446), whom Benedict (1975) saw as the first Sino-Tibetanist, also described the Sino-Tibetan family as encompassing Daic, albeit he rejected Konow's opinion (1909: 1) according to which Sino-Tibetan was divided into an eastern "Chinese-Siamese" sub-family and a western "Tibeto-Burman" one (Shafer 1955: 94).[7]

Another early work on Sino-Tibetan was written between 1964 and 1965 by Charles Frederick Voegelin (1906–1986) and Florence Marie Voegelin (1927–1989). This lengthy work, which was divided into five fascicles published in *Anthropological Linguistics*, also presents a Sino-Tibetan phylum which encompasses Sinitic, Hmong-Mien, Kra-Dai (called Kam-Tai and including a south-western group [Tai-Shan languages], a central group [Nung-Tho languages], and a northern group [Kam-Sui languages]), Naga-Kuki-Chin, Karen, Lolo-Burmese, Tibetan and 'Gyarung-Mishmi.' The Sino-Tibetan family, as presented by Voegelin and Voegelin (1964–1965), is visibly very similar to the *Indochinesisch* family described by Gabelentz and Conrady.

Benedict (1972) is of particular interest. Although he proposed a bifurcate Sino-Tibetan model, he was inclined to question the adaptability of traditional tree models for Tibeto-Burman. In fact, he regarded Kachin (i.e. Jingpho) as being central to the family, both geographically and typologically (Handel 2008: 427). Hence, the sub-grouping of Tibeto-Burman languages is not represented in a family tree, but radiates out from Kachin. To put it in Matisoff's words (1978: 2), it represents "an interlocking network of fuzzy-edged clots of languages, emit-

---

[5] This is the view which most Chinese scholars (e.g. Xíng Gōngwǎn 1991, Sūn Hóngkāi 1991, Pān Wùyún 2003, Zhèngzhāng Shàngfāng 2003, Ting and Sūn 2006) still hold today. The approach toward the issue occasionally remains politically motivated, and national and political preferences also figure, though not consistently. On the other side, scholars from Peking University implicitly contend that Chinese and Tibeto-Burman are not related (e.g. Guō Xīliáng 2002, 2003), mainly because of the time gap which occurs between the writing systems of the two main branches of Sino-Tibetan (Chinese and Tibeto-Burman). It is not very clear, however, whether these scholars do not differentiate between language and writing (also a possibility), or whether they claim that the time gap in the emergence of written documentation renders infeasible the correct application of the comparative method.

[6] This claim has been comprehensively disproved. See DeLancey (1990).

[7] Daic was removed from Sino-Tibetan by Benedict. See Benedict (1942).

ting waves of mutual influence from their various nuclear ganglia." Matisoff (1978) also agreed on the fact that evidence from Tibeto-Burman languages does not support the integrity of a *Stammbaum*.[8] What is most striking, however, is the fact that at the end of his laborious scientific work on Sino-Tibetan (1972), Benedict seemed to favour the idea that the connections between Sinitic and Tibeto-Karen might be of a substratum and superstratum kind (Künstler 2019: 19):

> It might be argued that the ST elements constitute only a superstratum in Chinese, and that the substratum is of distinct origin. In historical terms, the Chou people might be regarded as bearers of a ST language, which became fused with, or perhaps immersed in, a non-ST language spoken by the Shang people.[9]

Although at a first glance one may get the impression that Benedict supported the integrity of a bifurcate Sino-Tibetan family, as the title of his book suggests, at the end of his epoch-making study, Benedict (1972: 195–196) exposes four major arguments against the alleged kinship between Sinitic and Tibeto-Karen: (i) lack, from the Sinitic side, of the fairly elaborate Tibeto-Karen morphological system; (ii) presence of a restricted number of common roots; (iii) differences in the phonological system of the two branches which hardly become reconcilable at some points; (iv) non-correlation between the tonal systems of the two stocks.

When Benedict worked on his project in 1942, Karlgren's system of Old Chinese reconstructions was the only system available to him. Although Karlgren opposed to the idea that Sinitic lacked morphology, he tried to show that proto-Chinese possessed a perspicuous system of inflection in Indo-European fashion. This type of flectional morphology, of course, was very different from the derivational morphology exhibited by many Trans-Himalayan dialects, and Benedict could not relate it to any language of his Tibeto-Karen stock. Because those languages grouped together under the Sino-Tibetan phylum lacked exactly those morphological anomalies which facilitated the studies of the pioneer Indo-Europeanists,[10] workers in Sino-Tibetan had to plunge into comparative philology, trying to identify cognate roots mainly on the basis of overall phonetic similarities plus semantic congruity. This was exactly what Karlgren sought to avoid when he proposed his approach toward word-families, and it is probably this fact that led many linguists (perhaps including Benedict) to think that the genetic relationship between Sinitic and Tibeto-Burman relied only on a number of common lexical roots, from which only tentative phonological generalizations could be deduced (Künstler, 2019: 18–25).

Benedict (1976), nonetheless, devoted another influential work on Sino-Tibetan after his *Conspectus* was edited and published in 1972 by Matisoff. The main findings of this *aggiornamento* are that Sino-Tibetan is a well-established family, that Daic and Hmong-Mien should remain confined outside the Sino-Tibetan family, and that lexical analysis supported the taxonomic arrangement of the *Conspectus*, which setted Sinitic apart from Tibeto-Burman, albeit the position of Karen remained indeterminate.[11] It is important to remark, however, that

---

[8] Various Sinologists such as Edwin Pulleyblank (1991: 442) have also argued that the *Stammbaum* model is inappropriate for Sinitic languages. Many dialectologists, especially from China (e.g. Zhang et al. 2018), agree on this point. Nevertheless, others have pointed out that many arguments which seek to reject the traditional tree model are grounded on a series of misconceptions towards the nature and the scopes of the *Stammbaum* model (Orlandi 2019a: 132–134). For what regards how the tree model should be utilized in historical linguistics, see Jacques and List (2019). See also Kalyan and François (2019) for a competing hypothesis.

[9] Benedict (1972: 197).

[10] Indo-European shows a wide range of "aberrant features," such as thematic and athematic conjugations not reducible to the form of active plus affixes. Sino-Tibetan, for typological causes, cannot show this wide range of morphological markers into a string, nor can show a lucid system of apophony (so called *Ablaut*).

[11] However, few scholars have exposed or at least recognized the limitations of the Tibeto-Burman comparanda contained in Benedict's *Conspectus*. In fact, Benedict's reconstruction of Tibeto-Burman roots resembles

Benedict's view on the classification of Sino-Tibetan language was haphazard and at times self-contradictory. As he favoured the idea that proto-Sino-Tibetan was an SOV language, he was tempted to consider the SVO word-order in Sinitic and Karen a sufficient reason for treating Sino-Karen as the outgroup of the family.

In the meanwhile, contributions to the field of historical Chinese phonology had been done, most notably by Li Fang-kuei (1971) and Nicholas Bodman (1973). Benedict (1976: 172) was struck by "the numerous special lexical links between Chinese and Tibetan" which were cited by Bodman (1973), who suggested that Chinese stands in a special relationship to Tibetan. Notwithstanding Bodman's efforts to clarify the position of Sinitic within the Sino-Tibetan family, Benedict's previous Sino-Tibetan scheme practically remained unaltered. However, we must remember that the morphology of Old Chinese was still poorly explored, let alone understood, and that Benedict's argument for a bifurcation of the family was based on a tally of core vocabulary cognates (Swadesh 100-word list), which is not the most recommended practice in historical linguistics. Later discoveries of Old Chinese morphological mechanisms, in fact, have dramatically sharpened the burden between Sinitic and Tibeto-Burman languages.[12]

To conclude this lengthy paragraph, (i) although the alleged intellectual dichotomy between Leyden's Indo-Chinese and Klaproth's Tibeto-Burman is problematic and does not faithfully portrays the vibrant academic endeavours of nineteenth century linguistics, it is probably accurate to assert that the Sino-Tibetan phylum as understood by Chinese scholars (including Li Fang-kuei, *pace* LaPolla) and pioneers such as Shafer and Voegelin and Voegelin is a repackage of the late Indo-Chinese orthodoxy; (ii) the various taxonomic proposals presented in this paragraph, to which one could also include the two challenging models of Starostin (Sino-Kiranti) in 1994 and Bodman (Sino-Himalayan) in 1980 which sought to defy the traditional Sino-Tibetan scheme, are a clear indicator of the debated status of the *Stammbaum* of the family; (iii) the bifurcation of Sinitic vs non-Sinitic varieties of Trans-Himalayan was probably due to the inadequateness of reconstructions of Old Chinese; (iv) bifurcate Sino-Tibetan is just one among many other models that have been proposed from nineteenth century onward.

3. Common Innovations of non-Sinitic Languages

While it is not possible to assess all the attempts to justify the bifurcated Sino-Tibetan model, some of the proposed common innovations of the non-Sinitic languages may equally deserve a brief discussion.

Schuessler (2007: 30–31) already observed that some Chinese words seem to correspond to "Tibeto-Burman" final *-ʔ, whereas other words seem to correspond to TB final *-k. Baxter

more a teleological "reconstructive exercises," where proto-forms are often postulated through an imaginative process of syllable canon, than a reconstruction arrived by means of the orthodox comparative method. Beckwith (2002: xiii–xiv) is one of the few scholars to have exposed this problem: "[T]he only premodern Tibeto-Burman languages cited (unsystematically) in the book are what Benedict calls 'Written Tibetan' and 'Written Burmese'. These terms are in fact equatable with whatever is contained in, respectively, Jàschke's Tibetan-English dictionary and Judson's two Burmese dictionaries, both of which, though excellent, contain many modernisms alongside numerous archaisms."

[12] Regarding Benedict and Matisoff's reconstructions, there has been a heated debate in *Cahier de Linguistique Asie Orientale* on word-families and reconstructions between Fellner and Hill (2019ab) and Handel (2019), Thurgood (2019) and Schuessler (2019). While the present writer does not agree with every objection raised by Fellner and Hill (2019), it is felt that their advocacy for a stricter adherence to the traditional comparative method is essentially correct.

(1992: 182) speculated about the existence of a *-kʔ final cluster. Baxter and Sagart (2014) adduced good evidence in favour of the existence of a contrast between *-k and *-ʔ codas in Old Chinese. Sagart (2017) suggested that for the correspondence between Old Chinese *-k and Tibeto-Burman (i.e., non-Sinitic) *-k it would be natural to postulate a proto-Sino-Tibetan coda *-k, whereas for the correspondence between Old Chinese *-ʔ and Tibeto-Burman *-k, a compelling hypothesis would be to reconstruct proto-Sino-Tibetan *-q. Under this scenario, which Sagart referred to as scenario A, the merger of proto-Sino-Tibetan *-k and *-q codas would constitute a shared innovation uniquely observed among non-Sinitic languages. The other scenario (scenario B), however, would imply to treat the final glottal stop as a secondary development, perhaps through dialect stratification. This explanation would definitely rule out the alleged merger, provided that some doublets with an alternation between *-k and *-ʔ are attested in Old Chinese. Evidence for the existence of this type of doublets was adduced by Baxter (1992: 205), as Sagart himself (2017: 106) acknowledged in his study.[13] Nevertheless, Sagart writes that

> Explanation B is simpler and we could prefer it on that ground. However evidence external to Sino-Tibetan argues that explanation A is correct. An argument that the ST and Austronesian language families are genetically related was made in Sagart (2005), supported by sound correspondences, morphological parallels and arguments from cultural history, in particular relating to agriculture.

Unfortunately, Sino-Tibeto-Austronesian (STAN) remains unproven (cf. Blust 1995; Li 1995), and I have already shown in my rebuttal of STAN (2018: 6–7) that Old Chinese *-ʔ actually corresponds to multiple proto-Austronesian (PAN) final consonants without explained conditions, and in some cases it is even unmatched. Furthermore, it is true that certain affixes may fossilize in one language while remaining productive in another, however, under normal circumstances, external evidence should not take the place of internal evidence, all the more since Austronesian and Trans-Himalayan are not demonstrably related. In other words, I would rather be inclined to agree with Sagart that scenario B is simpler (Occam's razor) and, as such, preferable to scenario A. Note also that it has been demonstrated that uvulars are cross-linguistically disfavoured in a phonemic system without glottals also co-occurring (Sylak-Glassman 2014: 39).

For what regards other non-Sinitic innovations, it is often mentioned the merger of *a and *ə, proposed by Handel (2008) and discussed by Hill (2019: 29–31), in languages other than Chinese. It is true that the merger likely occurred in both Tibetan and Burmese, but it still needs to be demonstrated that it also occurred in all the non-Sinitic tongues of the Trans-Himalayan family. Also, sound changes involving vowels are quite frequent, and without further corroborative evidence from other domains (e.g., morphology, syntax) it would be premature to assume a bifurcation between Sinitic and non-Sinitic languages. Also, it would be helpful to explain why phonemic evidence should be more important than other types of evidence. For instance, DeLancey (2010, 2015, 2017) has written some articles on the study of verb agreement and verbal person marking, which is a near-neglected system in Trans-Himalayan. He spoke convincingly in favour of the existence of a central branch within Trans-Himalayan on the basis of morphological considerations. He showed that there is strong morphological

---

[13] Old Chinese reconstructions are in Baxter's system (1992): 採 *sri(k)ʔ > 'gather' 嗇 *srjik 'reap'; 采 *sri(k)ʔ 'colourful' 色 *srjik 'colour'. However, an anonymous reviewer has rightly pointed out to me that, in this specific case, Baxter (1992) postulated an odd reconstruction which violates its own reconstruction system, since OC *sr- does not yield MC *ṣ-, as it normally should. The semantic correlation between such pairs does not appear to be sufficiently strong.

evidence connecting Jinghpaw with both Northern Naga and Kuki-Chin (as in the classification of Voegelin and Voegelin 1965). Furthermore, according to DeLancey (2017), verbal person marking systems are lacking in Tibetic, Tani, Lolo-Burmese (and also Sinitic), but are widely diffused among other clades of the Trans-Himalayan family, most conspicuously in rGyalrongic, Qiang, Naga, Kuki-Chin, Kiranti and West Himalayan.

In a more recent study, Jacques (2019) has proposed that, contrary to common assumptions, the labial causative prefixes found in various Trans-Himalayan languages of Northeastern India are not innovations, but related to labial causative prefixes found in rGyalrongic and, possibly, also in other branches of Trans-Himalayan, including perhaps Sinitic. If this is correct, it may have further implications for what concerns the subgrouping of Trans-Himalayan languages.

Morphology plays a central role in historical linguistics, especially when attention is on homologous and not analogous features.[14] However, it is also true that structural features may be less informative than other types of evidence, such as lexicon, for reconstructing linguistic phylogenies, all the more since a huge typological gaps exists between morphologically rich rGyalrongic and Kiranti languages on the one hand, and Lolo-Burmese and Chinese on the other hand (Arcodia and Basciano 2020), as in the case of Trans-Himalayan. Since morphology can also be dramatically lost, some authors feel that comparing morphology would be especially tricky, as it will only be informative for the morphologically rich clades of the family. Indeed, the affixes which are not attested in Chinese may either never have existed, or have been lost without leaving traces (cf. DeLancey 2014; Gong 2017; Zhang, Lai and Jacques 2019).

From the phonological side, potential cases of common innovations between Lolo-Burmese and rGyalrongic (and even Chinese) have been adduced in recent publications, such as Jacques and Michaud (2011).[15] In their work on proto-Naish (Na, Naxi and Laze), Jacques and Michaud (2011: 491) have demonstrated that a merger between *n̥- and *l̥- (> *l̥-) is likely to have occurred in that proto-language. Something similar was described in a well-known source such as Baxter (1992: 194) is thought to have happened in Old Chinese as well, as clearly pointed out also by the authors of that study (2011: 492): e.g., Old Chinese *hn- and *hl- > medieval Chinese *tʰ- or *ɕ-, depending on the type of syllable.

Even if lexicon is given a special status, the subgrouping of Trans-Himalayan is not likely to bifurcate between Sinitic and non-Sinitic. Zhang, Jacques and Lai (2019) have arranged a substantial list of lexical cognates between Old Chinese and rGyalrongic, including some body parts. While it is not thought that everything is correct (e.g., rime *-eŋ, p. 84, should be dropped), some of the lexical comparanda adduced in this study, indeed, are fairly robust. Other promising cases of lexical innovations which point toward a closer relationship between Burmic and rGyalrongic are, instead, listed in Sagart et al. (2019b). Although not all the cases adduced by Sagart et al. (2019b: 26) have convinced the present writer, we may list at least four of the most convincing ones, viz. Japhug *tɯ-rtsʰɤz* 'lung' and Old Burmese *ʔachut* 'lung'; Japhug *ɯ-ʁɤri* 'the front side' and Old Burmese *rheʔ* 'the front side'; Japhug *zdɯm* 'cloud' and Old Burmese *tim* 'cloud'; Japhug *nɯqambɯmbjom*, Wobzi Khroskyabs *jmbjɥ̂m* 'to fly' and Old Burmese *pjaṃ* 'to fly'. While the Japhug form is much longer than the Burmese one, the au-

---

[14] See, for instance, the role that morphology plays in the validation of Kiranti (Jacques 2017, but Gerber and Grollman (2018) for a contrasting view) or in the definition of rGyalrongic (J. Sun 2000).

[15] The following examples, however, do not necessarily reflect the opinion of the present writer. For I do not think that the pieces of evidence mentioned below are robust enough to imply a closer relationship between the languages in question, and, in any case, the authors of such proposals never claimed so. Nevertheless, these examples may indicate where further work might prove rewarding.

thors have argued that it is due to the presence of a reduplication as well as of additional prefixes. Another good piece of evidence (which, incidentally, contradicts also a cognate found in the STEDT) in favour of a Tibeto-Derung connection, adduced in the supplementary material in Sagart (2019b), is the parallel between Derung *tuɯ³¹ wan⁵³* and Japhug *tɤjpa*. In particular, they argue that the comparison with Tibetan, Chepang and Kaman adduced in the STEDT (#471) is probably mistaken, because the Tibetan forms *wal* (Amdo) and *kha⁵⁵ wa⁵³* (Derge) included do not have final *-l (Sagart et al. 2019b: 25).

Also, although phylogenetic studies do not support the integrity of Sino-Bodic, potential cases of morphological and lexical connections between Chinese and Kiranti languages such as Limbu and Dumi were presented by van Driem (1997). More specifically, van Driem (1997: 466) has shown that the morphological alternations in the stem finals of the type */-k ~ -ŋ/, */-t ~ -n/ and */-p ~ -m/ exhibited by medieval Chinese verbs, preserved in the form of polyphonic readings of Chinese characters, "are precisely the type of regular morphophonological alternations manifested by the various classes of verb stem in Kiranti languages such as Limbu and Dumi". Lexical isoglosses are also included in van Driem's detailed study on Sino-Bodic, albeit they are less palatable for a number of motives.[16] However, two interesting parallels (cf. van Driem 1997), namely Old Chinese *hjuʔ 手 'hand' *vs* Limbu *huk* 'arm, hand', and Chinese *waŋ 王 'king' *vs* Limbu *haŋ* 'king', may also be discussed here. The Old Chinese initial of the former is uncertain (Schuessler, 2009:178). Some authors such as Baxter (1992, 1995) and Schuessler (2009) reconstruct an initial *h- phoneme, whereas Unger (1995: 133–135) reconstructed an Old Chinese nasal initial *n- mainly on the basis of textual glosses and graphic alternations. Unger (1995) points out that the *Shìmíng* 釋名 (200 CE), a gloss dictionary of Hàn dynasty (202 BC–220 AD), glosses this word with the character *xū* 須, which possibly had a cluster initial composed of a sibilant prefix plus a nasal phoneme, e.g. *sn-.[17] In addition, the *Guǎngyùn* 廣韻 uses the graph *chǒu* 杻 for *Qièyùn* 切韻 *niǔ* 杻 (also written as 狃), as indicated by Wáng Lì (1982: 231). Unger, thus, suggested that the word for 'hand' may be related to that for 'finger', viz. *niǔ* 狃 (or 丑). Alternatively, if the Old Chinese initial really was *n-, Schuessler (2007: 469) proposes that the word *shǒu* 手 'hand' perhaps may be derived from *shōu* 收 *nhiu 'to gather up, to collect, harvest.' In this case the rising tone would be the endoactive morpheme: 'hand' < lit. 'the thing which is doing the act of collecting.' If this is not correct, and the original initial consonant was *h-, then Limbu *huk* may eventually be a working hypothesis. Note, however, that in Limbu *h-* has several origins (Jacques 2017: 193–194), and that Kiranti languages have very different words for 'arm, hand'. For what regards the word for 'king,' the etymology is uncertain and a number of explanations have been proposed. Difficulties in the phonology of the sister languages (e.g. Written Burmese *dbaŋ* 'power' *vs* Old Chinese *waŋ 王) seem to exclude the Trans-Himalayan origin of this word. Alternatively, Schuessler (2007: 507–508) has indicated that it may be associated with an Austroasiatic homophone: Old Khmer *vaṅ ~ vāṅ(n)* 'royal palace' (also Tai *waŋ* 'palace'), which in turn is cognate to *luəŋ* 'king' (Tai *luaŋ* 'royal'). Note that Shaughnessy (1991: 197) has pointed out an example in bronze inscriptions where *waŋ 王 refers to a royal place and not to the king. Other

---

[16] While certain cognate sets present regular correspondences with proto-Kiranti, other cognate sets agree more closely with the initial of the modern Limbu form (van Driem 1997: 480). To give but one example, in van Driem (1997), Old Chinese *prôŋ 邦 'country' is compared with Limbu *pa:ŋphe:* 'village.' The final syllable *-phe:* is unmatched. In addition, the word for 'country' is possibly the same etymon as *poŋ 封 (Wáng Lì 1982: 388). Schuessler (2009: 169) has also shown that this lexeme may form an area etymon together with Lushai *puŋH* and Written Burmese *phuŋ-po*, further complicating van Driem's purported cognancy.

[17] The reader should be alerted, however, on the fact that sources based on late-Han era phonology cannot be taken at face value. Indeed, it is well-known that major changes had already occurred, especially in Type B syllables.

less parsimonious theories have indicated a poorly plausible connections between *waŋ and *gwaŋ 狂, because of certain theories which link archaic Chinese kingship to shamanism (Keightley et al. 1995: 132). In this case, the Sino-Bodic comparison may enter this controversy as a valid alternative to all other explanations. Corroborative evidence in favour of Sino-Bodic has been adduced in Hill (2011: 715). In the explicit attempt to reconcile Matisoff and Gong's presentations of the origins of Written Tibetan *o*, he pointed out that both Sinitic and Tibetan merged proto-Trans-Himalayan *o and *ow as *o* (although he concludes that Written Tibetan has three additional origins for *o*, viz. *wa, *wə, *aw).[18]

To sum up, the subgrouping of Trans-Himalayan languages still remains unclear and debated. However, with the fine tuning provided by recent scholarship, it is becoming more evident that the bifurcation between Sinitic and non-Sinitic languages is scarcely supported by linguistic facts.

4. Recent Bayesian Models in Trans-Himalayan Linguistics and the Validity of Bifurcate Sino-Tibetan

As argued in sections 1 and 2, although the kinship of Trans-Himalayan languages is not disputed,[19] the subgrouping of its components still remains debated. Recently, two important studies that employ computational approaches have been dedicated to the 'origin and spread' of Sino-Tibetan, one by Zhang et al. (2019) in *Nature*, and one by Sagart et al. (2019) in *PNAS*. In the present section special attention is given mostly (but not exclusively) to the former. In fact, although Sagart et al. (2019) call their tree Sino-Tibetan (perhaps as a face-saving practice), their study apparently comes up with a Trans-Himalayan tree, where the first split is between Sino-Sal (or Sino-Brahmaputran) and the rest of the family.

Zhang et al. (2019) applied "a particular model of probability testing known as Bayesian phylogenetic modelling" (LaPolla 2019: 1). While Bayesian modelling, a character-based method, also frequently makes use of the Swadesh 100-word list and partly resembles distant-based methods in that they are a measure of lexical agreement, this approach is different from lexicostatistics, because it does not claim that lexicon is replaced at a constant rate. Yet, it is necessary to distinguish between tools and methods. Bayesian phylogenetic models are a mathematical tool, and Bayes' rule is a "vote counting" procedure (Jones and Love 2011; Anderson 2011: 189, see below for more details). Now, there are two procedures which one may follow in order to evaluate a proposed method: one is testing its practicability (perhaps re-reproducing the experiment); the other is assessing the validity of the basic assumptions which the method requires. The second procedure is followed here.

The first objection is addressed to Zhang et al.'s (2019: 112) basic assumption, according to which, as a consequence of the fact that "language carries cultural information," "the evolution of language provides insight into prehistoric human *culture* (emphasis added)." It may be true that in a homogeneous cultural area, especially if it is characterized by the lack of historical discontinuity (is this the case?), the relationship between cultural and linguistic similarities may, as Sapir (1921: 234) said, "move along parallel lines." However, this writer is afraid that the authors of this study were driven rightly into the *cul-de-sac* which Sapir (1921: 234) warned about, in his famous though misconceived passage:

---

[18] But see Matisoff (2015) for a rebuttal.

[19] Among the methodologies that have been employed to "prove" the existence of Sino-Tibetan/Trans-Himalayan, we may list core vocabulary cognate counts (Benedict 1976), lexical analysis (Baxter 1995), morpho-syntactic relations (van Driem 1997), and shared innovations (Burling 1983; Thurgood 2003).

> It goes without saying that the mere content of language is intimately related to culture. A society that has no knowledge of theosophy need have no name for it; aborigines that had never seen or heard of a horse were compelled to invent or borrow a word for the animal when they made his acquaintance. In the sense that the vocabulary of a language more or less faithfully reflects the culture whose purposes it serves it is perfectly true that the history of language and the history of culture move along parallel lines. But this superficial and extraneous kind of parallelism is of no real interest to the linguist except in so far as the growth or borrowing of new words incidentally throws light on the formal trends of the language. The linguistic student should never make the mistake of identifying a language with its dictionary.

Second, the entire enterprise is flawed by the *contradictio in adjecto* inherent in their approach. As Zhang et al. (2019: 115) themselves explain, in order to understand the prehistory of human populations as well as their *culture*, they collated 949 lexical root-meanings across 109 Sino-Tibetan languages from the STEDT database, "[a]ccording to the items in the Swadesh 100-word list," which, unfortunately, seeks to develop a 'culturally independent' test list of universally used meanings.[20]

Third, it seems that Zhang et al.'s (2019: 114) understanding of some historical linguistic concepts such as 'comparative method,' 'glottochronology' and 'language classification' is imprecise:

> In historical linguistics, the comparative method—using abundant contemporary materials and historical documents—is an approach that has widely been used to establish language relationships. Glottochronology, which is an extension of the comparative method, uses lexical data to estimate absolute divergence times. However, glottochronology has considerable limitations (such as assuming a constant rate of language evolution) and does not account for different evolutionary rates of languages owing to contact, environmental change or varied rates of substitution among different categories of words.

First, the comparative method is *not* used to establish linguistic kinship (Hoenigswald 1960: 119; Newman 1970: 39; Thomason and Kaufman 1988: 201–202). The comparative method is a multi-task enterprise. On the one hand, it aims at producing as output a phonemic inventory, which is supposed to be a more or less close approximation of the phonological system of a language *X*, by taking as input phonemically transcribed lexical items from related languages (Hoenigswald 1960: 64); on the other, it is the method by which the hypothesis of an

---

[20] To put in the words of Bergsland and Vogt (1962: 151, also quoted in Teeter 1963: 645), "It is difficult to see ... how the various social, cultural, and linguistically patterned factors of lexical change could be determined from short word lists designed ... to eliminate those very factors." Another internal contradiction can also be observed in the lexicostatistical approach developed by Morris Swadesh. For instance, given that lexicostatistics per se relies on a list of culturally independent lexical items, it is of crucial importance that this list be universally valid. Hence, practitioners and supporters of lexicostatistics have been very active in revising and gradually shortening this list, at the expense of respect for the principles of statistical sampling, which by definition would imply that more trust should be placed in *longer* lists of lexical items, not shorter ones. The reality is that these aims are opposed in practice. As Teeter (1963: 642) has rightly pointed out, "the closer we get to universal validity, the fewer the items we have." Consider, for example, the Yakhontov list, which represents a mere 35-word subset of the Swadesh list. Although I do not contest the choice of specific lexical items, it is undeniable that this list may not consist of enough items to handle statistically. It is not a case that, as Holman et al. (2008) have shown, in identifying the kinship between the various Sinitic languages, the Yakhontov-35 list produces less accurate results than the original Swadesh-100 list. As an anonymous reviewer has pointed out to the present writer, however, it should be clarified that the various lexicostatistical tests of varying length may be more or less useful depending on the chronological distance that is measured by the tests. The fact that the Yakhontov-35 list is less useful for measuring dialectal differentiation within the framework of Sinitic is due to the fact that the so-called "stable items" items on them would show few—if any—variation, in the same way longer lists would produce scarce results for measuring the distance between the most distant branches of Trans-Himalayan, since the less-stable items would likely end up being replaced.

already presumed genetic relationship is demonstrated. While its ancillary procedure (i.e. 'internal reconstruction') in the broader sense might refer to any reconstructive technique that is not comparative,[21] the comparative method is the "procedure whereby morphs of two or more sister (i.e. genetically related) languages are matched in order to reconstruct the ancestor language" (Hoenigswald 1960: 119). In their supplementary material, they claim that "[t]here are challenges for the application of the historical linguistic comparative methods to systematically investigating the ST languages," because "the time depth of the ST language is far beyond all the historical accounts and linguistic documentation." This, however, holds true for many other linguistic families, some of which do not even possess a written documentation. This fact, however, is not known to have impeded the application of the comparative method (Campbell 1994, Orlandi 2019b).

Second, why glottochronology is supposed to be an extension of the comparative method we do not learn. Glottochronology, a method which Broca developed out of Dumont d'Urville's "lexicostatistics," is a method for estimating the time length since two or more languages diverged from a common ancestor tongue. The employment of the comparative method, on the other hand, involves (i) the establishment of systematic phonological correspondences in words of same or related meaning; (ii) the reconstruction of phonemic systems; (iii) the establishment of grammatical (especially morphological) correspondences; (iv) the reconstruction of grammatical systems (Thomason and Kaufman 1988: 202). Since glottochronology provides no way of going beyond lexical analysis, it is hard to believe that it is an extension of the comparative method.

Furthermore, one may also point out certain grave mistakes, such as putting the location of Tibetan in Lhasa rather than in the Yarlung/Chongye valley (Hill 2019: 3; Takeuchi 2012: 4; Stein 1972). However, what is most striking is that they did not publish their underlying data, but just relied on one scholar's cognate judgments who strongly believed the theory Zhang et al. (2019) claim to confirm. All we are presented with is a table with ST reconstructions (mostly PTB) marked for presence/absence of reflexes in daughter languages in their supplementary materials. The list does not include a number of well-known languages such as Meitei, Newar, or Lepcha. Furthermore, it would be useful to point out that not including Sinitic languages besides modern Mandarin and so called "Old Chinese" (not properly a language itself, but a set of hypothetical and reconstructed sound classes) could potentially represent a methodological shortcoming. For instance, Sagart et al. (2019a) also include other Sinitic languages, such as Guǎngzhōuhuà, Cháozhōuhuà, Xìngnínghuà, Lónggǎnghuà, etc. Sinitic is a much more diversified sub-family than they assume, but just how short is the list of languages included within their Sinitic group is particularly stunning.

In addition, because not every reader may have familiarity with Bayesian phylogenetic modelling, it is necessary to briefly introduce this method. Bayesian inference is a method of statistical inference where Bayes' theorem [$P(H|E) = P(H) \times P(E|H)/P(E)$] is used to update the probability for a hypothesis as more evidence becomes available.[22] Bayes' theorem describes

---

[21] In a stricter sense, by 'internal reconstruction' is intended the comparison of allomorphs in order to arrive at a reconstructed *invariant* form of a morpheme. The reconstruction entails sound change and may also say something about allophonic relationships, though not necessarily.

[22] *H* is the hypothesis under investigation and *E* is the new evidence. The symbol "|" means "given", so that $P(H|E)$ represent the probability that the hypothesis is true given the new evidence, i.e. 'posterior probability', whereas $P(H)$ is the prior probability (i.e. the probability that the hypothesis is true based on prior evidence). $P(E|H)$ represents the so called 'likelihood function'. $P(E)$ is the dominator and represents the probability of the new evidence.

the probability of an event based on prior knowledge of related conditions, and in so doing it seeks to specify the optimal way to combine old evidence with new information. Bayesian phylogeny in linguistics is a character-based method which typically involves three procedures: encoding, representation and interpretation (Heggarty 2006: 184; Campbell 2013: 474). With the exception of few applications (e.g, Eska and Ringe 2004), in most Bayesian phylogenetic approaches, lexical items are the only characters coded.

More specifically, Bayesian phylogenetic applications involve the following procedures: (i) compiling a list of basic vocabulary items (normally the Swadesh 100-word list); (ii) finding their equivalents in the languages investigated; (iii) searching for or relying on cognates already identified by linguists; (iv) convert them into binary or multi-state representations; (v) using a computer software to construct a phylogenetic tree (Campbell 2013: 476–477). Of course the software produces a myriad of trees, which need to be further assessed. Bayesian methods have experienced a tremendous progress in recent years, following the fine-tuning of estimation and probability theories (Chater et al. 2006), which prompted many scholars to adopt Bayesian approaches in several areas of science. Because Bayesian inference is thought be able to explain all the data without invoking arbitrary explanations, and to allow uncertainties to be estimated in a natural way, Bayesian approaches are not only at the forefront of phylogenetic development, but are also enjoying a position of prominence in historical linguistics, with authors who have gone as far as claiming that they constitute the future of historical linguistics (cf. Verkerk 2017).[23]

Proponents of Bayesian phylogenies have been active in reassuring that these quantitative approaches should not be regarded as substitute for linguistic innovations, and that statistical models will not replace the role of historical linguists. Many historical linguists (cf. Holm 2007, Campbell 2013: 482–490) argue that, although Bayesian models are certainly more sophisticated than distance-based models such as glottochronology and lexicostatistics, they are equally subject to the limitations resulting from their near-exclusive reliance on lexical items, which nonetheless are widely regarded as being incapable of revealing a sufficient range of linguistic change.[24] Greenhill et al. (2017) have demonstrated that, however, structural features are generally less informative and more prone to change than basic lexicon for reconstructing linguistic phylogenies.

As argued above, Bayesian phylogenetic modelling, like any regression model, attempts to infer a tree-generating model that is consistent with the priors in the given data. This does not imply that Bayesian inference can validate a given linguistic genealogy, nor does it mean that it can measure the probability that the studied hypothesis is true. While it is not denied that they might be useful to historical linguistics, these methods, more often than not, can at best yield results that are very close to those hypothesized by historical linguists. In fact, it is not very clear how they might advance our knowledge of the linguistic families they are applied to. For successful attempts in utilizing Bayesian models simply led to a recapitulation of existing theories, whereas other less successful attempts (e.g., Foster and Toth 2003; Serva and Petroni 2008) led to results that are too absurd to be taken seriously (such as, e.g., placing Al-

---

[23] Other specialists, instead, disagree. See Eska and Ringe (2004), Donohue, Denham and Oppenheimer (2012), Pereltsvaig and Lewis (2015).

[24] Lexicon had a central role in the first applications of "lexicostatistics" by Dumont d'Urville (1790–1842) not because lexicon was reckoned as the most stable part of a language, but because vocabulary was all the evidence available to him. Also, the alleged scarce "borrowability" among the items of the famous Swadesh 100-word list has been criticized (Haspelmath and Tadmor 2009). However, supporters of Bayesian inference (e.g., Greenhill and Gray 2012; Scarborough 2019) in linguistics claim that 'borrowings' are not a major issue in Bayesian approaches.

banian together with Romance languages as one branch of Indo-European, cf. Serva and Petroni 2008). For this reason, many historical linguists are left with the impression that these Bayesian models are "reinventing the wheel" (Campbell 2013: 485). For instance, Gray et al. (2000, 2009) argue that their classification of Austronesian languages matches the comparative method well (but curiously enough, one of the most eminent experts of Austronesian linguistics, Robert Blust, who incidentally has supplied the cognancy judgments which their Bayesian experiment relied upon, has not abandoned the conventional scheme of Austronesian languages in favour of the "new" Bayesian phylogeny), but it offers few, if any, original insights which we did not already know. Even within the framework of Sino-Tibetan, the classification of languages whose placement within the family is controversial or not sufficiently understood is not clarified.[25] The phylogenetic study by Zhang et al. (2019) would have been much more useful, had they clarified, e.g., the subgrouping of languages such as Lepcha and Newar that are difficult to classify. As such, despite its prestige and growing diffusion, this Bayesian study has yet to deliver on its promise to advance the empirical study of ST/Trans-Himalayan historical linguistics.

With Zhang et al.'s (2019) understanding of the 'comparative method' as a tool which merely explores the lexicon of two or more languages, and their identification of language with vocabulary, it is not surprising that they chose to put emphasis only on lexical data. However, it is well known since the times of Lorenzo Hervás y Panduro (1735–1809) and James Burnett, Lord Monboddo (1714–1799) that genetic relationship must be proven by grammatical evidence (as Meillet and Sapir believed), especially paradigmatic morphology (i.e. the combination of morphological markers into a string). To put in Nichols's (1996: 64) words:

> The diagnostic evidence is grammatical, and it combines structural paradigmaticity (usually multiple paradigmaticity) and syntagmaticity with concrete morphological forms. The Indo-Europeanists' intuitive feel for what was diagnostic evidence of relatedness corresponds to a computable threshold of probability of occurrence, and the main purpose of this chapter has been to give a simple rule of thumb for judging evidence of relatedness. A grouping can be regarded as established by the comparative method if and only if it rests on individual-identifying evidence.

Nevertheless, certain aspects outlined in Zhang et al. (2019) may, indeed, deserve more attention. For instance, the close relationship outlined in Zhang et al. (2019) between Tani and Digarish or Naga and Kuki Chin (but already acknowledged by DeLancey 2012 et seq.) is probably correct. On the other hand, Sagart et al. (2019a) placed Tibetan together with Derung, and Sinitic with Sal.[26] Potential cases of common lexical innovations are offered in the supplementary material (Sagart et al. 2019b: 24–25), and this might be a good way to encourage further works to follow this direction.

---

[25] Actually it was not even discussed.

[26] DeLancey (2015) reckons the Sal languages as being a sub-branch of the wider Central Tibeto-Burman group. The Sal sub-family is generally represented with three branches, consisting of Bodo-Garo, Konyak and Kachinic or Jingpho-Luish languages (Burling 2003: 175; Thurgood 2003: 11). Divergent opinions are also encountered. Shafer originally put Konyak and Bodo-Garo under his 'Baric' rubric. Bradley (1997: 20) also combines them into a single sub-branch. Matisoff (2013: 20, 2012) divides Sal languages into two main branches, Bodo-Garo and Jingpho-Konyak. This division was foreshadowed by Benedict (1972:7), who noted that Bodo-Garo, Konyak and Kachin languages shared, together with the extinct Chairel language (a Luish language), distinctive roots for 'sun' and 'fire.' Van Driem (2001: 397–398, 403) originally included Bodo-Garo, Konyak, Dhimalish and Kachin-Luic in his Brahmaputran branch. However, in his most recent publications, he presents this family as encompassing only Bodo-Garo and Konyak. See also Coupe (2012).

Nevertheless, it should also be remembered that the consensus trees provided in both studies are not "just so" stories, but only partial representations of the Trans-Himalayan *Stammbaum*. For instance, Zhang et al. (2019) obtained a posterior probability of 0.6 for the non-Sinitic group, which means that this group is only "weakly supported".[27] In Sagart et al.'s (2019) consensus tree, the Sino-Sal branch has a low posterior probability (0.33, cf. Sagart et al. 2019b), albeit it remains the highest candidate. This means that either Sinitic is the outgroup or that it sides with Sal, though they are not necessarily the outgroup.

4. Future directions: the reconstruction of the parent language and the broader affinities of Trans-Himalayan

As argued in sections 1 and 3, the *Stammbaum* of Trans-Himalayan is still debated. Regarding language subgrouping, the identification of shared innovations is the most recommended practice of historical linguistics. However, as LaPolla (2001: 245) rightly pointed out, very few of the proposed subgrouping schemes for ST give reasons for their outlines. A number of recent publications such as Honda (2014), Hyslop (2014) and Čašule (2014), all in the volume edited by Thomas Owen-Smith and Nathan W. Hill, nonetheless tried to sharpen the demarcation line between innovations, inheritances and borrowings. What emerges from these new endeavours, however, is that our knowledge of Trans-Himalayan historical phonology is still limited (Owen-Smith and Hill 2014: 6). In order to distinguish innovations from borrowings, the reconstruction of proto-Trans-Himalayan (or of its branches) seems to be indispensable (cf. Hyslop 2014). Previous attempts at reconstructing the ancestor language of Trans-Himalayan (though called Sino-Tibetan) include Coblin (1986), Gong (1995) and Peiros and Starostin (1996). Reconstructions of proto-Tibeto-Burman are found in Benedict (1972) and Matisoff (2003). Coblin's work (1986), one of the first attempts at incorporating more refined OC reconstructions, unfortunately suffers from some shortcomings. First, his handlist shows a "non-historical" treatment of linguistic material (Künstler 2019: 63). In fact, we find the proto-words for 'iron' or 'metal' (Coblin 1986: 98–99), although we have no evidence that the material cultures which are often associated to Tibeto-Burman populations effectively possessed iron and/or metal in general. If they had no knowledge of iron, it is hardly credible that they had a word for it. Second, this putative handlist of Sino-Tibetan words presents methodological peculiarities. For far too many of his reconstructed linguistic units, typically consisting of long strings of reconstructed phonemes, are tolerated to merge with zero in several attested Sino-Tibetan languages without explained conditions.

Gong (1995) has worked extensively on the most well-documented languages of the family, *viz.* Written Tibetan, Written Burmese, and Old Chinese (mainly as reconstructed by Li Fang-kuei, albeit with some minor revisions). The present article shall not treat in detail Gong's reconstructions, in accordance with its aim of discussing principle and not practice. It suffices to point out that, perhaps, reconstructing a wide family, comprising around 600 languages, just from the surface of three of them will produce, in all probability, a proto-language plagued by the Berkson's fallacy, no matter how rigorous is the reconstruction attempted. In fact, if we reconstruct the ancestor language of a very wide and still insufficiently explored family merely on the basis of Written Tibetan, Written Burmese and Old Chinese, we inevitably introduce a bias towards what is regular and frequent in the morphemic system of these three languages, which would further tempt us to exclude all deviant features a priori. Future

---

[27] I wish to thank an anonymous reviewer for his/her thoughtful comment on this aspect.

reconstructions of the proto-language will have to take into account also other languages in addition to Tibetic, Burmic and Sinitic.

Uncertainties and difficulties with the reconstruction of proto-Trans-Himalayan are testified by a significant divergence in the typology of the proto-languages. While some scholars, most notably LaPolla (1989, 1992, 1994), have argued that verbal agreement systems are independent innovations in various "Tibeto-Burman" branches of the family, others such as van Driem (1993), DeLancey (1989, 2010) and Jacques (2012) hold the view that the proto-language was a highly synthetic tongue whose subgroups later experienced varying degrees of phonological erosion (cf. Owen-Smith and Hill 2014). While none of these typological proposals has weathered the years free of difficulties, this writer's predilection is for the reconstruction of a synthetic proto-language. DeLancey (2014) argues that a process of creolization, following their emergence as *linguae francae*, was the cause of the simpler typology of the analytic languages of Trans-Himalayan. This makes perfectly sense also in view of the process of urbanization and state centralization which is frequently associated with languages such as Chinese. In addition, McWhorter (1998: 798, 2001) lists three typological features which distinguish creoles from other languages, *viz*. (i) little or no inflectional affixation; (ii) little or no use of tone to lexically contrast monosyllables or encode syntax; (iii) semantically regular derivational affixation. However, others (cf. Ansaldo and Matthews 2001: 315–317) have rightly pointed out that this is of little help in East and South Asian context. Other features of creole languages have been listed by Aronoff et al. (2005: 302–312, 326). Some of these include (i) little inflection, or very little derivation, largely consisting of affixation with no simultaneous morphology; (ii) absence of polymorphemic classifier constructions. Most importantly, they have pointed out that the sociolinguistic conditions of usage in a given community may vary, and these variances may give rise to a difference in the grammatical properties of the creole (Aronoff et al. 2005: 304). For instance, certain sociolinguistic factors may favour a creole continuum with mesolectal varieties that are grammatically more similar to the dominant, lexifier language (Holm 1988). Although these ideas are far from being presentable at the moment, this is something future works will inevitably have to take into account.

We can now turn to the broader affiliation of Trans-Himalayan. Since a discussion about all these proposed distant classifications of Sino-Tibetan goes far beyond the scope of the present article, only some lesser-known attempts at linking Trans-Himalayan with other families are surveyed here. These proposals are not discussed in the specific. Overall, Sino-Tibetan has been linked with Na-Dene (Sapir 1920; Bengtson 1994) and Athabaskan (Shafer 1952, 1957, 1969; Swadesh 1952, 1965), Yeniseian and Caucasian (Starostin 1982, 1984, 1989, 1991; Starostin and Ruhlen 1994; Nikolaev 1991), Austroasiatic and Hmong-Mien (Starosta 2005), Indo-European (Chang 1988; Pulleyblank 1983, 1995, 1996), Austronesian (Sagart 1994, 2005, 2011, 2016, 2017), Hmong-Mien, Kradai, Austroasiatic and Austronesian (Pān 2003; Zhèngzhāng 2003). This last one, known as Sino-Austric (*Huá'Ào* 華澳) or Yangtzean, is a megalomaniac and very speculative distant relationship. Although it has been only poorly and vaguely described, one hesitates to cite specific examples from this putative though unclear list of genetic factors which are supposed to purport the integrity of this megalophylum only because one is not willing to give further currency to these pseudo-linguistic speculations. However, one comes up with the impression that mere geographical proximity, together with certain political adumbrations, are at the basis of this megalophylum. Sino-Caucasian has been surveyed elsewhere (van Driem, 2005b), therefore is not discussed here. Sagart's Sino-Tibeto-Austronesian (STAN) is perhaps the most promising among these proposals of distant genetic relationship. It is grounded on both linguistic and archaeological (especially for what regards the domestication of two types of cereals) arguments. Linguistically, it was rejected by Blust

(1995, 2014) with two devastating critiques. More recently, the theory has been re-examined by Orlandi (2018), who appears to be more open-minded to it, although he also concludes that the evidence in favour of STAN is still insufficient for proving this distant relationship, given the non-obvious relationship of the languages involved. Sino-Dene is much less popular, and has not gained much support besides that of its original proponents. Although the theory goes back to Sapir, connections between Trans-Himalayan and the languages of the Americans were already noted by the English engineer and philologist Henry Harcourt Hyde Clarke (1815–1895).[28] He also argued that Lepcha, a Trans-Himalayan language spoken by Lepcha people in Sikkim and parts of West Bengal, Nepal and Bhutan, was genetically related to Chadic (Houssa in Clarke's terminology). Of course, the history of Sino-Dene is unrelated to these absurd speculations. In brief, in 1920, Sapir noted that Na-Dene was fundamentally different from the other languages spoken in North America. He wrote a series of letters to Alfred Kroeber where he enthusiastically spoke of a connection between Na-Dene and "Indo-Chinese." In 1925, a supporting article summarizing his thoughts, albeit not written by him (cf. Bengtson 1994: 208), entitled "The Similarities of Chinese and Indian Languages," was published in *Science Supplements*. Sapir mainly compared Na-Dene classifiers with a similar verb prefix *s-* found in Tibetan. The Sino-Dene hypothesis never gained foothold in the United States, besides Sapir's circle, though it was later revitalized by Shafer (1952, 1957, 1969) and Swadesh (1952, 1965). Nevertheless, their Sino-Tibetan-Athabaskan connections were assailed by other scholars (e.g. Matisoff, note 13 of Benedict's *Conspectus* 1972: 3; Campbell 1988: 593). Today very few scholars have expressed their predilection for Sino-Dene. A notable supporter is Bengtson, who also holds the view that Sino-Dene and STAN could be reconciled at sufficient time depths (pc, 2018).

5. Remarks on the Trans-Himalayan *Urheimat*

The present article intentionally overlooked the ongoing debate on the purported Trans-Himalayan *Urheimat*. Some remarks, however, are also recommendable. Some scholars favour autochthony and argue for a Trans-Himalayan/Sino-Tibetan *Urheimat* in present-day northern China; others favour language diversification and argue for an *Urheimat* in northern India and Nepal.[29] Trans-Himalayan, however, is not the sole linguistic family where we do have two conflicting theories about the placing of a homeland. Even the hypothetical homeland of a well-explored family such as Indo-European is still controversial. One theory purports that

---

[28] See Clarke (1881: 118). However, Clarke misinterpreted the kinship of his Kolarian family, which was supposed to include Sino-Tibetan languages (Kuki-Chin-Naga, Bodo), Dravidic (Konda) and Austroasiatic (Sora) languages. Clarke believed that other Himalayish languages, generally put under the Indo-Chinese phylum, were instead related to Ugric tongues. In Clarke's view, Ugric languages encompassed Samoyedic, Hungarian (Magyar), Khanty (Ostyak), Mordvinic, Mari (Cheremis), Udmurt (Votyak) and Lapp. Clarke saw lexical evidence for a genealogical relationship between Ugric and the Himalayan languages of East Nepal, such as Bantawa (Rungchengbung), Bahing and other languages which today are classified as Kiranti. In addition, there were other languages which show a genetic relationship with Ugric, albeit they were reckoned as being more distantly related: Takpa (Tshangla) and the Qiangic language Manyak (Muya). Kiranti languages of west Nepal were also considered genetically related, though not as close as the Kiranti tongues of east Nepal. Sunwar, Gurung, Magar, Newar and Tamang (Moormi). Clarke further included Vayu, Abhor-Minyong (Abor group) and Orunodoi (Sivasagar, previously Sibsagar), but left out the so-called 'prehistoric languages of north-east Nepal,' such as Bodo, Kachari (both being Bodo-Garo languages) and Dhimal (Clarke 1878: 47).

[29] See, for instance, Blench and Post (2014).

Indo-European originated somewhere within the Steppes near the Caspian Sea around 5,000~7,000 years ago. While there is no direct archaeological evidence for their dispersal north to Scandinavia, south to the Mediterranean, west to Europe and east to India, many linguists accept this provisional *Urheimat* because, based on what is known about rates of language changes, Indo-European could not have existed prior to 7,000~8,000 years ago. Nevertheless, the British archaeologist and palaeolinguist Colin Renfrew (1990) has pointed out that evidence from climatic changes and technological developments would suggest that Indo-Europeans originated as early as 9,000 years ago in Anatolia (cf. Bouckaert et al. 2012).

The typological classification of languages is generally inaccurate and can tell us few about the genetic relationship of languages.[30] The identification of the *Urheimat*, nonetheless, is a typological-cultural enterprise, as we are supposed to identify the typology of the speakers of a reconstructed language with the typology of early cultures, as reconstructed by means of archaeological data (Kitson 1997: 184). In fact, the association of a hypothetical group of proto-speakers with their possible homeland is an intuitive assessment of the time covered by the internal changes reconstructible for the *Ursprache* and by those socio-archaeological cultures which they are associated with (Kitson 1997: 183–184). Many of the homeland theories related to Trans-Himalayan are strictly connected to the assumption according to which "language families arise through demographic processes driven by favourable changes in food procurement" (Sagart et al. 2019a: 10320; cf. Diamond and Bellwood 2003). This 'language dispersal' model is not new, as it has been emphasized by Renfrew (1973 et seq.) and Bellwood (1991 et seq.), and was pioneered a long time ago by von Heine-Geldern (1914). Renfrew associated the spread of Indo-European languages in Europe with the spread of agriculture in the Old Continent, which he interpreted by the 'wave of advance' model of Ammermann and Cavalli-Sforza (1973, 1979, 1984; cf. Campbell and Poser 2008: 337). A wealth of publications have demonstrated that language and agriculture may follow separate vectors, and at least for what regards many other linguistic families, the model is simply untenable (Kaufman 1976; Vansina 1995; Gardner 2000; cf. Campbell and Poser 2008: 344–346). It would be also a good practice to illustrate why the language/agriculture spread model is better than the generally accepted 'centre of gravity' model, based on the principle of minimum moves and maximum diversifi-

---

[30] Some authors, however, claim otherwise. See, for instance, Greenhill et al. (2010), Wichmann and Saunders (2007), Dediu (2011), Wichmann (2013). Some of these studies are indeed very insightful and informative about how to use typological features. However, they do not invalidate the statements that typological classification cannot tell us much about the genealogical classification of languages: Hopi is very different typologically from Aztec, Chamic from Rukai, and Manchu from any other Tungusic language. Within the framework of Trans-Himalayan, we have SVO languages such as Chinese and Kiranti, and SOV languages, highly synthetic languages such as rGyalrongic and isolating languages such as Burmese. Yet Chamic and Rukai are Austronesian, and Manchu is a Tungusic language. Japanese and Korean are typologically close to the point that, in most cases, one can literally translate morpheme by morpheme from one language to the other, and yet they are unrelated. It is clear that typology is not a reliable criterion for language classification, albeit one may effectively look at typology in a different way. However, even some of these typological studies had to conclude that the degree of confidence toward the acceptance of hypotheses concerning genetic relationships generated by algorithms on the basis of typological data is still unknown to us. Indeed, the present writer suspects that this task is at odds with the genealogical classification of languages. Postulating a genetic relationship is a very different type of ability from the fast processing toward a specific end that the new quantitative approaches seek to pursue. It is well-known that the genetic relationship among languages is heuristic in principle, and relies upon some key human abilities such as intuition and common sense. It is very difficult to devise an algorithm for these key abilities, since all that an algorithm can do is to behave in a logical way (which is quite different from 'reasoning'). In fact, there is a neat difference between the type of fluid intelligence demanded by historical linguistics and the type of 'crystallized' intelligence that comes from prior learning and/or past experience which is proper to quantitative approaches, including Bayesian inference.

cation, which is widely used for assigning linguistic homelands to accepted linguistic families (Campbell and Poser 2008), or why it is preferred over other approaches which go beyond the traditional comparative method, such as the 'punctuated equilibrium' model (Dixon 2002), or the 'spread zone' approach (Nichols 1992, 1997).

Genetic studies, sometimes, seem to support these cultural and typological identifications. However, language diversity cannot be explained by genetic factors, albeit the relationship between the two is often stimulating. For example, it is well-known that, as a result of what evolutionary biologists call the 'founder effect,' genetic diversity is higher in Africa and diminishes with increasing distance from it. Curiously, languages in Africa exhibit the highest number of phonemes, including a peculiar type of plosive consonants produced with a velaric ingressive (velaric suction in the Ladefoged and Williamson feature system) airstream, and the farther away phonemic inventories are from Africa, the simpler they become, touching their lowest levels in the languages of the Pacific. Yet this may be only a coincidence, as several exceptions are found in a number of languages of the north-west coast of the Americas. While it is not denied the usefulness of linguistic palaeontology, a term coined by Adolphe Pictet (1859–1863) in a romantic attempt to reconstruct the world of the proto-Indo-Europeans, it should be remembered that the primary scope of language reconstruction is to test and validate an assumed genetic relationship between languages, and not to give indications about the culture or the homeland of their alleged speakers.

6. Conclusions

The detailed history of the Western linguistic development of the hypothesis of the Sino-Tibetan family presented in this article has demonstrated that the hypothesis of a Sino-Tibetan phylum bifurcated into Sinitic and non-Sinitic (i.e. Tibeto-Burman) has never been really substantiated, and as such many scholars may not necessarily be mistaken in claiming that it reflects an obsolete view (cf. McColl et al. 2018: 362). The more agnostic term Trans-Himalayan is the most neutral and accurate name for the family (Simon and Hill 2015: 381). Given that the recent Bayesian study by Sagart et al. (2019) has produced a Trans-Himalayan tree and that the one by Zhang et al. (2019), which purports to substantiate the bifurcate Sino-Tibetan model, contains many methodological peculiarities, it is concluded that, for the present at least, the bifurcate Sino-Tibetan model is weakened rather than enhanced by these statistical approaches. Hugely important is the fact that, even though one is persuaded of similar Bayesian approaches, the two most recent works of this kind have come up with two divergent scenarios, one of which produced a Trans-Himalayan tree.

References

Abel-Rémusat, Jean-Pierre. 1825. *Mélanges Asiatiques*. Paris: Librairie Orientale de Dondey-Dupré Pères et Fils.

Ammerman, Albert J., Luigi Luca Cavalli-Sforza, Colin Renfrew. 1973. The explanation of culture change: models in prehistory. In: Colin Renfrew (ed.). *A population model for the diffusion of early farming in Europe; Models in Prehistory*: 343–357. Pittsburgh: University of Pittsburgh Press.

Ammerman, Albert J., Luigi Luca Cavalli-Sforza. 1979. The wave of advance model for the spread of agriculture in Europe. In: Colin Renfrew, Kenneth L. Cooke (eds.). *Transformations, mathematical approaches to culture change*: 275–293. New York, NY: Academic Press.

Ammerman, Albert J., Luigi Luca Cavalli-Sforza. 1984. *The Neolithic transition and the genetics of populations in Europe*. Princeton, NJ: Princeton University Press.

Anonymous. 1925. The similarity of Chinese and Indian languages. *Sci. Suppl.* 62(1607): xii.

Ansaldo, Umberto, Stephen J. Matthews. 2001. Typical creoles and simple languages: The case of Sinitic. *Linguistic typology* 5(2): 311–325.

Arcodia, Giorgio Francesco, Bianca Basciano. 2020. Morphology in Sino-Tibetan languages. In: Aronoff, Mark (ed.). *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. doi 10.1093/acrefore/ 9780199384655.013.530.

Aronoff, Mark, Irit Meir, Wendy Sandler. 2005. The paradox of sign language morphology. *Language* 81(2): 301–344.

Atkinson, Quentin. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337(6097): 957–960.

Atkinson, Quentin D., Russell D. Gray. 2005. Curious parallels and curious connections – phylogenetic thinking in biology and historical linguistics. *Systematic biology* 54(4): 513–526.

Atkinson, Quentin D., Geoff Nicholls, David Welch, Russell D. Gray. 2005. From words to dates: water into wine, mathemagic or phylogenetic inference? *Transactions of the Philological Society* 103(2): 193–219.

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

Baxter, William H. 1995. 'A stronger affinity... than could have been produced by accident': A probabilistic comparison of Old Chinese and Tibeto-Burman. *Journal of Chinese Linguistics, monograph series* 8: 1–39.

Baxter, William H., Laurent Sagart. 1998. Word formation in old Chinese. In: Jerome L. Packard (ed.) *New approaches to Chinese word formation: morphology, phonology and the lexicon in modern and ancient Chinese*: 35–76. New York: Walter de Gruyter.

Bayer, Theophilus Siegfried. 1730. *Museum Sinicum.* St. Petersburg.

Beckwith, Christopher (ed.). 2002. *Medieval Tibeto-Burman languages*. Leiden, Netherlands: Brill.

Bellwood, Peter. 1991. The Austronesian dispersal and the origin of languages. *Scientific American* 265(1): 88–93.

Bellwood, Peter. 1994. An archaeologist's view of language macrofamily relationships. *Oceanic linguistics* 13: 391–406.

Bellwood, Peter. 2000. The time depth of major language families: An archaeologist's perspective. In: Colin Renfrew, April McMahon, Robert Lawrence Trask (eds.) *Time depth in historical linguistics*. Cambridge, UK: McDonald Institute for Archaeological Research.

Bellwood, Peter. 2001. Early agriculturalist population diasporas? Farming, languages, and genes. *Annual review of anthropology* 30(1): 181–207.

Benedict, Paul K. 1942. Thai, Kadai, and Indonesian: a new alignment in Southeastern Asia. *American anthropologist* 44(4): 576–601.

Benedict, Paul K. 1975. *Austro-Thai language and culture, with a glossary of roots*. New Have: Human Relations Area Files Press.

Benedict, Paul K. 1976. Sino-Tibetan: another look. *Journal of the American Oriental Society* 96(2): 167–197.

Benedict, Paul K., James A. Matisoff. 1972. *Sino-Tibetan: a conspectus*. Berkeley, CA: University of California Press.

Bengtson, John D. 1994. Edward Sapir and the "Sino-Dene" hypothesis. *Anthropological Science* 102(3): 207–230.

Bergsland, Knut, Hans Vogt. 1962. On the validity of glottochronology. *Current Anthropology* 3(2): 115–153.

Blench, Robert, Mark Post. 2014. Rethinking Sino-Tibetan phylogeny from the perspective of Northeast Indian languages. In: Thomas Owen-Smith, Nathan W. Hill (eds.). *Trans-Himalayan linguistics: Historical and descriptive linguistics of the Himalayan area*: 71–104. Berlin: De Gruyter Mouton.

Blust, Robert. 1995. An austronesianist look at Sino-Austronesian. *Journal of Chinese Linguistics, monograph series* 8: 283–298.

Blust, Robert. 2014. Some recent proposals concerning the classification of the Austronesian languages. *Oceanic linguistics* 53(2): 300–391.

Bodman, Nicholas C. 1973. Some Chinese Reflexes of Sino-Tibetan S-Clusters. *Journal of Chinese linguistics* 1(3): 383–396.

Bodman, Nicholas C. 1980. Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship. *Cornell linguistic contribution Ithaca 3*: 34–199.

Bouckaert, Remco, Philippe Lemey, Michael Dunn, Simon J. Greenhill, Alexander V. Alekseyenko, Alexei J. Drummond, Russell D. Gray, Marc A. Suchard, Quentin D. Atkinson. 2012. Mapping the origins and expansion of the Indo-European language family. *Science* 337(6097): 957–960.

Bradley, David. 1997. Tibeto-Burman languages and classification. In: David Bradley (ed.). *Tibeto-Burman languages of the Himalayas (Papers in Southeast Asian Linguistics, 14)*: 1–72. Canberra, Australia: Department of Linguistics, The Australian National University.

Brandreth, Edward Lyall. 1878. On the non-Aryan languages of India. *Journal of the Royal Asiatic Society* 10(1): 1–32.

Burling, Robbins. 1983. The Sal languages. *Linguistics of the Tibeto-Burman area* 7(2): 1–32.

Burling, Robbins. 2003. The Tibeto-Burman languages of Northeastern India. In: Graham Thurgood, Randy J. La-Polla (eds.). *The Sino-Tibetan languages*: 167–192. London / New York: Walter de Gruyter.

Campbell, Lyle. 1988. Review article of Language in the Americas, by J. H. Greenberg. *Language* 64: 591–615.

Campbell, Lyle. 1994. Linguistic reconstruction and unwritten languages. In: R. E. Asher, J. M. Y. Simpson (eds.). *Encyclopedia of language and linguistics*: 3475–3480. Ann Abor: Pergamon Press.

Campbell, Lyle, William J. Poser. 2008. *Language classification: History and method*. Cambridge University Press.

Čašule, Ilija. 2014. Burushaski kinship terminology of Indo-European origin. In: Thomas Owen-Smith, Nathan W. Hill (eds.). *Trans-Himalayan linguistics*: 181–220. Berlin: Walter de Gruyter.

Chater, Nick, Joshua B. Tenenbaum, Alan Yuille. 2006a. Probabilistic models of cognition: Conceptual foundations. *Trends in cognitive sciences* 10(7): 287–291.

Chater, Nick, Christopher D. Manning. 2006b. Probabilistic models of language processing and acquisition. *Trends in cognitive sciences* 10(7): 335–344.

Clarke, Hyde. 1878. Himalayan origin and connection of the Magyar and Ugrian. *The Journal of the Anthropological Institute of Great Britain and Ireland* 7: 44–65.

Clarke, Hyde. 1881. Osservazioni sulla lingua onorifica ovvero cerimoniale della tribù lepcha o ròng di Dorjeling, nella regione Himalaia. In: *Atti del IV Congresso Internazionale degli Orientalisti*: 117–122. Firenze, Italia: Coi Tipi dei Successori Le Monnier.

Coblin, Weldon South. 1986. *A sinologist's handlist of Sino-Tibetan lexical comparisons*. Nettetal: Steyler Verlag.

Conrady, August. 1896. *Eine indochinesische Causativ-Denominativ-Bildung und ihr Zusammenhang mit den Tonaccenten, ein Beitrag zur vergleichenden Grammatik der indochinesischen Sprachen, insonderheit des Tibetischen, Barmanischen, Siamesischen und Chinesischen, von Dr. August Conrady*. Wien: Otto Harrassowitz.

Conrady, August. 1915. Der altchinesische Fragesatz und der steigende Ton. *Mitteilungen des Seminars für orientalische Sprachen* 1(18): 261–297.

Coupe, Alexander R. 2012. Overcounting numeral systems and their relevance to sub-grouping in the Tibeto-Burman languages of Nagaland. *Language and linguistics* 13(1): 193–220.

Cust, Robert. 1881. Sulle lingue non ariane delle indie orientali. *Atti del IV Congresso Internazionale degli Orientalisti Tenutosi in Firenze 1878*: 123–128. Firenze, Italia: Coi Tipi dei Successori Le Monnier.

Dài, Qìngxià. 1997. Guānyú HànZàngyǔ fēnlèi de sīkǎo [Thoughts on subgrouping in Sino-Tibetan]. *Yǔyán jiàoxué yǔ yánjiū* 4: 4–10.

Dài, Qìngxià, Huáng Bùfán, Fù Àilán, Rénzēng Wàngmǔ, Liú Júhuáng (eds.). 1991. *Zàngmiǎnyǔ shíwǔ zhǒng* [Fifteen Tibeto-Burman languages]. Beijing: Běijīng yànshān chūbǎn shè.

Dediu, Dan. 2011. A Bayesian phylogenetic approach to estimating the stability of linguistic features and the genetic biasing of tone. *Proceedings of the Royal Society B: Biological Sciences* 278(1704): 474–479.

DeLancey, Scott. 1987. The Sino-Tibetan languages. In: Bernard Comrie (ed.). *The World's major languages*: 799–810. New York: Oxford University Press.

DeLancey, Scott. 1989. Verb agreement in Proto-Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 52(2): 315–333.

DeLancey, Scott. 1991. Sino-Tibetan languages. In: William Bright (ed.). *The International Encyclopedia of Linguistics*: 445–449. New York: Oxford University Press.

DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9(1): 1–39.

DeLancey, Scott. 2014. Creolization in the divergence of the Tibeto-Burman languages. In: Thomas Owen-Smith, Nathan Hill (eds.). *Trans-Himalayan linguistics*: 41–70. Berlin / Boston: Walter de Gruyter.

DeLancey, Scott. 2015. Morphological evidence for a central branch of Trans-Himalayan (Sino-Tibetan). *Cahiers de linguistique Asie orientale* 44(2): 122–149.

DeLancey, Scott. 2017. Hierarchical and accusative alignment of verbal person marking in Trans-Himalayan. *Journal of South Asian languages and linguistics* 4(1): 85–105.

Diamond, Jared M., Peter Bellwood. 2003. Farmers and their languages: the first expansions. *Science* 300(5619): 597–603.

Dixon, Robert M. W. 2002. *Australian languages: Their nature and development (Vol. 1)*. Cambridge University Press.

Donohue, Mark, Tim Denham, Stephen Oppenheimer. 2012. New methodologies for historical linguistics?: Calibrating a lexicon-based methodology for diffusion vs. subgrouping. *Diachronica* 29: 505–522.

van Driem, George. 1988. Reflexes of the Tibeto-Burman *-t directive suffix in Dumi Rai. In: David Bradley, E. J. Henderson, Martine Mazaudon, Richard Keith Sprigg (eds.). *Prosodic Analysis and Asian Linguistics: To Honour R. K. Sprigg*: 157–167. Canberra: Pacific Linguistics.

van Driem, George. 1993. The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* 56(2): 292–334.

van Driem, George. 1997. Sino-Bodic. *Bulletin of the School of Oriental and African Studies* 60(3): 455–488.

van Driem, George. 2001. *Languages of the Himalayas: An ethnolinguistic handbook of the greater Himalayan region (Vol. 2)*. Leiden: Brill.

van Driem, George. 2003. Review of Thurgood and LaPolla 2003. *Bulletin of the School of Oriental and African Studies* 66(2): 282–284.

van Driem, George. 2005a. Tibeto-Burman vs Indo-Chinese: implications for population geneticists, archaeologists and prehistorians. In: Roger Blench, Laurent Sagart, Alicia Sanchez-Mazas (eds.). *The Peopling of East Asia*: 109–134. London: Routledge.

van Driem, George. 2005b. Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory. *Contemporary issues in Nepalese linguistics*: 285–338.

van Driem, George. 2007. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1(2): 211–270.

van Driem, George. 2011. The Trans-Himalayan phylum and its implications for population. *Communication on contemporary anthropology* 5(1): 135–142.

van Driem, George. 2014. Trans-Himalayan. In: Thomas Owen-Smith, Nathan W. Hill (eds.). *Trans-Himalayan linguistics*: 11–40. Berlin / Boston: Walter de Gruyter.

Eska, Joseph F., Don Ringe. 2004. Recent work in computational linguistic phylogeny. *Language* 80: 569–582.

Fellner, Hannes A., Nathan W. Hill. 2019a. Word families, allofams, and the comparative method. *Cahiers de linguistique Asie orientale* 48(2): 91–124.

Fellner, Hannes A., Nathan W. Hill. 2019b. The differing status of reconstruction in Trans-Himalayan and Indo-European. *Cahiers de linguistique Asie orientale* 48(2): 159–172.

Finck, Franz Nikolaus. 1909. *Die Haupttypen des Sprachbaus*. London: Teubner.

Forbes, Charles James F. S. 1878. On Tibeto-Burman languages. *Journal of the Royal Asiatic Society* 10(2): 210–227.

Gabelentz, Georg von der. 1881. Sur la possibilité de prouver l'existence d'une affinité généalogique entre les langues dites Indochinoises. *Atti del IV Congresso Internazionale degli Orientalisti Tenutosi in Firenze 1878*: 283–293. Firenze, Italia: Coi Tipi dei Successori Le Monnier.

Gardner, Peter M. 2000. *Bicultural versatility as a frontier adaptation among Paliyan foragers of South India (Vol. 2)*. Lewiston, MN: Edwin Mellen Press.

Gerber, Pascal, Selin Grollmann. 2018. What is Kiranti? A critical account. *Bulletin of Chinese Linguistics* 11(1/2): 99–152.

Gong, Hwang-cherng. 1995. The system of finals in Proto-Sino-Tibetan. *Journal of Chinese Linguistics, monograph series* 8: 41–92.

Gong, Xun. 2017. Jiāróng yǔzǔ yǔyán yǔ shànggǔ Hànyǔ de xíngtàixué [The rGyalrong languages and the morphology of Old Chinese]. In: Haeree Park, Shaoxuan Cheng (eds.). *Gǔwénzì yǔ hànyǔ lìshǐ yīnyùn xué* [Ancient scripts and historical Chinese phonology]: 134–156. Shanghai: Fùdàn zhōnghuá wénmíng yánjiū zhuānkān.

Greenberg, Joseph H. 1996. In defense of Amerind. *International journal of American linguistics* 62(2): 131–164.

Greenhill, Simon J., Quentin D. Atkinson, Andrew Meade, and Russell D. Gray 2010. The shape and tempo of language evolution. *Proceedings of the Royal Society B: Biological Sciences* 277(1693): 2443–2450.

Greenhill, Simon J., Russell D Gray. 2012. Basic vocabulary and Bayesian phylolinguistics: Issues of understanding and representation. *Diachronica* 29(4): 523–537.

Greenhill, Simon J., Chieh-Hsi Wu, Xia Hua, Michael Dunn, Stephen C. Levinson, Russell D. Gray. 2017. Evolutionary dynamics of language systems. *Proceedings of the National Academy of Sciences* 114(42): E8822–E8829.

Grierson, George A. 1909. *Linguistic survey of India: Vol. III, Part I, Tibeto-Burman family: Tibetan dialects, the Himalayan dialects and the North Assam group*. Calcutta: Superintendent of Government Printing, India.

Guō, Xīliáng. 2002. Lìshǐ yīnyùn xué yánjiū zhōng de jǐ gè wèntí——bó Méi Zǔlín zài Xiānggǎng yǔyán xuéhuì niánhuì shàng de jiǎnghuà [Some problems related to historical Chinese phonology: Refuting Mei Tsu-lin's arguments at the Hong Kong Language Society annual meeting]. *Gǔ Hànyǔ yánjiū* 3: 2–9.

Guō, Xīliáng. 2003. Yīnyùn wèntí dá Méi Zǔlín [Problems in historical phonology, A response to Mei Tsu-lin]. *Gǔ Hànyǔ yánjiū* 3: 2–17.

Handel, Zev. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2(3): 422–441.

Handel, Zev. 2019. A brief response to Fellner and Hill's "Word families, allofams, and the comparative method". *Cahiers de linguistique Asie orientale* 48(2): 125–141.

Harlez, Charles. 1878. Les langues monosyllabiques. *Actes du dixième congrès international des orientalistes*: 67–88. Leiden: Brill.

Heggarty, Paul. 2006. Interdisciplinary indiscipline? Can phylogenetic methods meaningfully be applied to language data – and to dating language. In: Peter Forster, Colin Renfrew (eds.). *Phylogenetic methods and the prehistory of languages*: 183–194. Cambridge: McDonald Institute for Archaeological Research.

Heine-Geldern, Robert. 1914. *Die Bergstämme des nördlichen und nordöstlichen Birma*. Wien: Elisabeth Stiglmayr.

Hill, Nathan W. 2011. Multiple origins of Tibetan o. *Language and linguistics* 3(12): 707–721.

Hill, Nathan W. 2019. *The historical phonology of Tibetan, Burmese, and Chinese*. Cambridge University Press.

Hoenigswald, Henry. 1960. *Language change and linguistic reconstruction*. Chicago University Press.

Holm, John A. 1988. *Pidgins and creoles: theory and structure. Vol. 1*. Cambridge University Press.

Holm, Hans J. 2007. The new arboretum of Indo-European "trees", Can new algorithms reveal the phylogeny and even prehistory of Indo-European? *Journal of quantitative linguistics* 14: 167–214.

Holman, Eric W., Søren Wichmann, Cecil H. Brown, Viveka Velupillai, André Müller, Dik Bakker. 2008. Explorations in automated language classification. *Folia Linguistica* 42(3/4): 331–354.

Honda, Ideo. 2014. Internal diversity in the Tamangic lexicon. In: Thomas Owen-Smith, Nathan W. Hill (eds.). *Trans-Himalayan linguistics*: 131–154. Berlin/Boston: Walter de Gruyter.

Hyslop, Gwendolyn. 2014. A preliminary reconstruction of East Bodish. In: Thomas Owen-Smith, Nathan W. Hill (eds.) *Trans-Himalayan linguistics*: 155–180, Berlin / Boston: Walter de Gruyter.

Jacques, Guillaume. 2012. Agreement morphology: the case of Rgyalrongic and Kiranti. *Language and Linguistics* 13: 83–116.

Jacques, Guillaume. 2017. A reconstruction of Proto-Kiranti verb roots. *Folia Linguistica Historica* 38: 177–215.

Jacques, Guillaume, Alexis Michaud. 2011. Approaching the historical phonology of three highly eroded Sino-Tibetan languages: Naxi, Na and Laze. *Diachronica* 28(4): 468–498.

Jacques, Guillaume, Johann-Mattis List. 2019. Save the trees. Why we need tree models in linguistic reconstruction (and when we should apply them). *Journal of Historical Linguistics* 9(1): 128–167.

Jīn Lǐxīn. 1998. HànZàngyǔ de míngcí hòuzhuì -n [The Sino-Tibetan noun suffix -n]. *Mínzú Yǔwén* 1: 43–48.

Karlgren, Bernhard. 1915–1926. *Études sur la phonologie Chinoise* [Studies on Chinese phonology]. Archives d'études orientales. Vols I-IV. Leiden / Stockholm: E. J. Brill.

Karlgren, Bernhard. 1931. Tibetan and Chinese. *T'oung Pao* 28: 25–70.

Karlgren, Bernhard. 1933. Word families in Chinese. *Bulletin of the Museum of Far Eastern Antiquities* 5: 5–120.

Kaufman, Terrence. 1976. Archaeological and linguistic correlations in Mayaland and associated areas of Meso-America. *World archaeology* 8(1): 101–118.

Kalyan, Siva, Alexandre François. 2019. When the waves meet the trees. *Journal of Historical Linguistics* 9: 167–176.

Kitson, Peter R. 1997. Reconstruction, typology, and the "original homeland" of the Indo-Europeans. In: Jasek Fisiak (ed.). *Linguistic reconstruction and typology* 96: 183–239. Berlin: Walter de Gruyter.

Klaproth, Heinrich Julius. 1823. *Asia polyglotta*. Berlin: A. Schubart.

Kuhn, Ernst. 1889. *Beiträge zur Sprachenkunde Hinterindiens*. München: Verlag der K. Academie. In Commission bei G. Franz.

Kuhn, Thomas S. 1962. *The structure of scientific revolutions*. Chicago University Press.

Künstler, Mieczysław Jerzy. 2019. *The Sinitic languages: A contribution to sinological linguistics*. London: Routledge.

LaPolla, Randy J. 1989. Verb agreement, head-marking vs. dependent-marking, and the 'deconstruction' of Tibeto-Burman morpho-syntax. In: *Annual meeting of the Berkeley Linguistics Society, Vol. 15*: 356–366. Berkeley: University of California.

LaPolla, Randy J. 1992. On the dating and nature of verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 55(2): 298–315.

LaPolla, Randy J. 1994. Variable finals in Proto-Sino-Tibetan. *Bulletin of the Institute of History and Philology Academia Sinica* 65: 131–173.

LaPolla, Randy J. 2001. The role of migration and language contact in the development of the Sino-Tibetan language family. In: Alexandra Y. Aikhenwald, Robert M. W. Dixon (eds.). *Areal diffusion and genetic inheritance: Case studies in language change*: 225–254. Oxford University Press.

LaPolla, Randy. 2013. Subgrouping in Tibeto-Burman. In: Balthasar Bickel, Lenore Grenoble, David Peterson, Alan Timberlake (eds.). *Language typology and historical contingency*: 463–474. Amsterdam: John Benjamins.

LaPolla, Randy J. 2016. Once again on methodology and argumentation in linguistics. *Linguistics of the Tibeto-Burman area* 39(2): 282–297.

LaPolla, Randy J. 2019. The origin and spread of Sino-Tibetan languages. *Nature, News and Views* 1–2. doi.org/10.1038/d41586-019-01214-6.

Leyden, John. 1808. *On the languages and literature of the Indo-Chinese nations*. London: T. Hubbard at the Hindoostanee Press.

Li, Fang-Kuei. 1936–1937. Languages and dialects. *The Chinese year book* (1936): 121–128. Reprinted in the *Journal of Chinese Linguistics* 1(1): 1–13.

Li, Fang-kuei. 1971. Studies on archaic Chinese phonology. *Tsinghua Journal of Chinese Studies* 9(1/2): 1–61.

List, Johann-Mattis, Mary Walworth, Simon J. Greenhill, Tiago Tresoldi, Robert Forkel. 2018. Sequence Comparison in Computational Historical Linguistics. *Journal of Language Evolution* 12: 130–144.

Matisoff, James A. 1973. *The grammar of Lahu*. Berkeley: University of California.

Matisoff, James A. 1978. *Variational semantics in Tibeto-Burman: The "organic" approach to linguistic comparison*. Berkeley: University of California.

Matisoff, James A. 1991. Sino-Tibetan linguistics: present state and future prospects. *Annual review of anthropology* 20(1): 469–504.

Matisoff, James A. 2002. On 'Sino-Bodic' and other symptoms of neosubgroupitis. *Bulletin of the School of Oriental and African Studies* 63(3): 356–369.

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press.

Matisoff, James A. 2013. Re-examining the genetic position of Jingpho: Putting flesh on the bones of the Jingpho/Luish relationship. *Linguistics of the Tibeto-Burman area* 36(2): 15–95.

Matisoff, James A. 2015. On the demise of the Proto-Tibeto-Burman mid vowels. *Bulletin of the National Museum of Ethnology* 39(3): 375–395.

McColl, Hugh, Fernando Racimo, Lasse Vinner, Fabrice Demeter, Takashi Gakuhari, J. Víctor Moreno-Mayar, George Van Driem. 2018. The prehistoric peopling of Southeast Asia. *Science* 361(6397): 88–92.

McWhorter, John H. 1998. Identifying the creole prototype: Vindicating a typological class. *Language* 74(4): 788–818.

Mei, Tsu-lin. 1970. Tones and prosody in Middle Chinese and the origin of the rising tone. *Harvard Journal of Asiatic Studies* 30: 86–110.

Mei, Tsu-lin. 1989. The causative and denominative functions of the *s-prefix in Old Chinese. *Proceedings of the Second International Conference on Sinology*: 31–51. Taipei: Academia Sinica.

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.

Michailovsky, Boyd. 1985. Tibeto-Burman dental suffixes: evidence from Limbu (Nepal). In: Graham Thurgood, James A. Matisoff, David Bradley (eds.) *Linguistics of the Sino-Tibetan area: the state of the art; papers presented to Paul K. Benedict for his 71st birthday*: 363–375. Canberra: Dept. of Linguistics, Research School of Pacific Studies, The Australian National University.

Müller, Friedrich Max. 1854. *Letter to Chevalier Bunsen on the classification of the Turanian languages*. Oxford.

Müller, Friedrich Max. 1866. *Lectures on the science of language: Delivered at the Royal Institution of Great Britain in April, May, and June 1861*. London: Longmans Green.

Newman, Paul. 1970. Historical sound laws in Hausa and in Dera (Kanakuru). *Journal of West African languages* 1(1): 39–51.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Nichols, Johanna. 1996. The comparative method. In: Mark Durie, Malcolm Ross (eds.). *The comparative method reviewed: Regularity and irregularity in language change*: 39–71. Oxford University Press.

Nichols, Johanna. 1997a. The epicenter of the Indo-European linguistic spread. In: Roger Blench, Roger & Matthew Spriggs (eds.). *Archaeology and language I. Theoretical and methodological orientations*: 122–148. London: Routledge.

Nichols, Johanna. 1997b. Modeling ancient population structures and movement in linguistics. *Annual review of anthropology* 26(1): 359–384.

Nikolaev, Sergei L. 1991. Sino-Caucasian languages in America. In: Vitaly Shevoroshkin (ed.). *Dene-Sino-Caucasian languages: Materials from the first international interdisciplinary symposium on language and prehistory, Ann Arbor, 8.–12. November 1988*: 67–129. Bochum: Brockmeyer.

Orlandi, Georg. 2018. Evaluating the Sino-Tibeto-Austronesian hypothesis. *Journal of Language Relationship* 16(1): 1–18.

Orlandi, Georg. 2019a. On the classification of the Ng Yap dialects: some thoughts on the subgrouping of Sinitic languages. *Journal of Language Relationship* 17(1): 128–152.

Orlandi, Georg. 2019b. Yǔyán chóngjiàn bùzhǐ yīlài shūmiànyǔ (biǎo yīn wénzì) [Language reconstruction does not rely on written languages (phonographic scripts)]. *Yǔyán kēxué* 18(3): 225–235.

Orlandi, Georg. 2020. The state of the art of the genetic relationship of Japonic: the Turanian and Altaic Hypotheses. *International Journal of Eurasian Linguistics* 2(1): 29–69.

Owen-Smith, Thomas, Nathan W. Hill (eds.). 2013. *Trans-Himalayan linguistics: Historical and descriptive linguistics of the Himalayan area* (Vol. 266). Boston / Berlin: Walter de Gruyter.

Pān, Wùyún. 2000. *Hànyǔ lìshǐ yīnyùnxué* [Chinese historical phonology]. Shanghai: Shànghǎi jiàoyù chūbǎnshè.

Peiros, Ilia, Sergei A. Starostin. 1996. *A comparative vocabulary of five Sino-Tibetan languages*. 6 vols. Parkville: Univ. of Melbourne, Dept. of Linguistics and Applied Linguistics.

Pictet, Adolphe. 1859–1863. *Les origines indo-européennes ou, Les Aryas primitifs: essai de paleontologie linguistique / par Adolphe Pictet*. Paris: Joel Cherbuliez.

Pulleyblank, Edwin G. 1983. The Chinese and their neighbors in prehistoric and early historic times. In: David N. Keightley (ed.). *The origins of Chinese civilization*: 411–466. Berkeley: University of California Press.

Pulleyblank, Edwin G. 1991. Chinese dialect studies. *Journal of Chinese Linguistics, monograph series* 3: 429–453.

Pulleyblank, Edwin G. 1995. The historical and prehistorical relationships of Chinese. *Journal of Chinese Linguistics, monograph series* 8: 145–194.

Pulleyblank, Edwin G. 1996. Early contacts between Indo-Europeans and Chinese. *International Review of Chinese Linguistics* 1(1): 1–25.

Renfrew, Colin. 1973. Problems in the general correlation of archaeological and linguistic strata in prehistoric Greece: the model of autochthonous origin. In: R. A. Crossland, Ann Birchall (eds.). *Bronze Age migrations in the Aegean: archaeological and linguistic problems in Greek prehistory*: 263–276. London: Duckworth.

Renfrew, Colin. 1990 [1987]. *Archaeology and language: the puzzle of Indo-European origins*. Cambridge University Press.

Renfrew, Colin. 1991. Before Babel: speculations on the origins of linguistic diversity. *Cambridge Archaeological Journal* 1(1): 3–23.

Renfrew, Colin. 1994. World linguistic diversity. *Scientific American* 270(1): 116–123.

Renfrew, Colin. 2003. World linguistic diversity and farming dispersals. In: Roger Blench, Matthew Spriggs (eds.). *Archaeology and language I: theoretical and methodological orientations*: 104–112. London: Routledge.

Renfrew, Colin. 2000. At the edge of knowability: towards a prehistory of languages. *Cambridge Archaeological Journal* 10(1): 7–34.

Renfrew, Colin. 2002. The emerging synthesis: the archaeogenetics of farming/language dispersals and other spread zones. In: Peter Bellwood, Colin Renfrew (eds.). *Examining the farming/language dispersal hypothesis*: 3–16. Cambridge: McDonald Institute for Archaeological Research.

Ryder, Robin. 2009. *Phylogenetic models of language diversification*. Ph.D. dissertation. Oxford University. Online at: https://tel.archives-ouvertes.fr/tel-00661866.

Sagart, Laurent. 1993. New views on Old Chinese phonology. *Diachronica* 10(2): 237–260.

Sagart, Laurent. 1994. Proto-Austronesian and Old Chinese evidence for Sino-Austronesian. *Oceanic linguistics* 33(2): 271–308.

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

Sagart, Laurent. 2005. Sino-Tibetan-Austronesian: An updated and improved argument. In: Laurent Sagart, Roger Blench, Alicia Sanchez-Mazas (eds.). *The peopling of East Asia: Putting together archaeology, linguistics and genetics*: 161–176. London: Routledge.

Sagart, Laurent. 2011. The homeland of Sino-Tibetan-Austronesian: where and when. *Communication on contemporary anthropology* 5: 143–147.

Sagart, Laurent. 2016. The wider connections of Austronesian. *Diachronica* 33(2): 255–281.

Sagart, Laurent. 2017. A candidate for a Tibeto-Burman innovation. *Cahiers de linguistique Asie orientale* 46(1): 101–119.

Sagart, Laurent, Roger Blench, Alicia Sanchez-Mazas (eds.). 2005. *The peopling of East Asia: Putting together archaeology, linguistics and genetics*. London: Routledge.

Sagart, Laurent, Tze-Fu Hsu, Yuan-Ching Tsai, Yue-Ie C. Hsing. 2017. Austronesian and Chinese words for the millets. *Language dynamics and change* 7(2): 187–209.

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill, Johann-Mattis List. 2019a. Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences* 116(21): 10317–10322.

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill, Johann-Mattis List. 2019b. Supplementary information for the Paper 'Dated language phylogenies shed light on the ancestry of Sino-Tibetan languages'. *Proceedings of the National Academy of Sciences* 116(21): 1–38.

Sapir, Edward. 1920. *Comparative Sino-Tibetan and Na-Dene dictionary*. Ms. American Philosophical Society 20a 3.

Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace.

Saxena, Anju (ed.). 2004. *Himalayan languages: past and present*. Berlin / New York: Walter de Gruyter.

Schuessler, Axel. 2007. *ABC etymological dictionary of Old Chinese*. Honolulu: University of Hawaii Press.

Schuessler, Axel. 2019. Comments on Fellner and Hill: "Word families, allofams, and the comparative method". *Cahiers de linguistique Asie orientale* 48(2): 142–153.

Shafer, Robert. 1952. Athapaskan and Sino-Tibetan. *International journal of American linguistics* 18(1): 12–19.

Shafer, Robert. 1955. Classification of the Sino-Tibetan languages. *Word* 11(1): 94–111.

Shafer, Robert. 1957. Note on Athapaskan and Sino-Tibetan. *International Journal of American Linguistics* 23(2): 116–117.

Shafer, Robert. 1965. The Eurasial linguistic superfamily. *Anthropos* 1(6): 445–468.

Shafer, Robert. 1966–1967. *Introduction to Sino-Tibetan. Vol. 1*. Wiesbaden: Otto Harrassowitz.

Shafer, Robert. 1969. A few more Athapaskan and Sino-Tibetan comparisons. *International Journal of American Linguistics* 35(1): 67.

Simon, Camille, Nathan Wayne Hill. 2015. Tibetan. In: Nicola Grandi, Lívia Körtvélyessy (eds.). *Edinburgh handbook of evaluative morphology*: 381–388. Edinburgh University Press.

Simon, Walter. 1929. Tibetisch-chinesische Wortgleichungen, ein Versuch. *Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich-Wilhelms-Universität zu Berlin* XXXII 1: 157–228.

Starosta, Stanley. 2005. Proto-East Asian and the origin and dispersal of the languages of East and Southeast Asia and the Pacific. In: Laurent Sagart, Roger Blench, Alicia Sanchez-Mazas (eds.). *The peopling of East Asia*: 210–226. London: Routledge.

Starostin, Sergei A. 1982. Praenisejskaja rekonstrukcija i vneshnije sv'azi enisejskix jazykov [Proto-Yeniseian reconstruction and the external relations of the Yenisei languages]. *Ketskij sbornik*: 144–237. Leningrad: Nauka.

Starostin, Sergej A. 1984. Gipoteza o geneticheskix sv'az'ax sinotibetskix jazykov s enisejskimi i severnokavkazskimi jazykami [Hypothesis on the genetic relationship of Sino-Tibetan languages with the Yeniseian and North Caucasian languages]. In: *Lingvisticheskaja rekonstrukcija i drevnejshaja istorija Vostoka* [Linguistic reconstruction and the ancient history of the East]: 19–38. Moskva: Nauka.

Starostin, Sergei A. 1989. Nostratic and Sino-Caucasian. In: Vitaly Shevoroshkin (ed.). *Explorations in language macrofamilies*: 42–66. Bochum: Brockmeyer.

Starostin, Sergei A., Merritt Ruhlen. 1994. Proto-Yeniseian Reconstructions, with Extra-Yeniseian Comparisons. In: Merritt Ruhlen (ed.). *On the origin of languages: Studies in linguistic taxonomy*: 70–92. Stanford, CA: Stanford University Press.

Sūn Hóngkāi. 2001. Shì lùn Zhōngguó jìngnèi Zàngmiǎn yǔ de pǔxì fēnlèi [A classification of Tibeto-Burman languages in China]. In: Tatsuo Nishida (ed.). *Languages and history in East Asia: A festschrift for Tatsuo Nishida on the occasion of his 60th Birthday*, *Vol. I*: 61–73. Kyoto: Shokado.

Sūn Hóngkai, Chénglóng Huáng, Máocǎo Zhōu. 2001. *Róuruòyǔ yánjiu* [A study of Rouruo]. Beijing, China: Zhongyang Mínzú Dàxué Chubanshè.

Sun, Jackson Tien-Shin. 1993. The linguistic position of Tani (Mirish) in Tibeto-Burman: A lexical assessment. *Linguistics of the Tibeto-Burman Area* 16(2): 143–188.

Sun, Jackson Tien-Shin. 2000. Parallelisms in the verb morphology of Sidaba rGyalrong and Lavrung in rGyalrongic. *Language and linguistics* 1(1): 161–190.

Swadesh, Morris. 1952. Review article of 'Athapaskan and Sino-Tibetan', by R. Shafer. *International journal of American linguistics* 18: 178–181.

Swadesh, Morris. 1965. Lingvisticheskie sv'azi Ameriki i Evrazii [Linguistic ties of America and Eurasia]. *Etimologija 1964* [Etymology 1964]: 271–322. Moskva: Nauka.

Takashima, Kenichi. 1996. Language and palaeography. In: Gary F. Arbuckle (ed.). *Studies in early Chinese civilization: Religion, society, language and palaeography. Vol. 1: Text. Vol. 2: Tables and notes*. Kansai Gaidai University Publication.

Takeuchi, Tsuguhito. 2012. Formation and transformation of Old Tibetan. In: Tsuguhito Takeuchi, Norihiko Hayashi (eds.). *Historical development of the Tibetan languages. Proceedings of the workshop B of the 17th Himalayan languages symposium, Kobe, 6th–9th September*: 3–17. Kobe: Research Institute of Foreign Studies.

Teeter, Karl V. 1963. Lexicostatistics and Genetic Relationship. *Language* 39(4): 638–648.

Thomason, Sarah Grey, Terrence Kaufman. 1988. *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Thurgood, Graham. 2019. A review of Fellner and Hill's "Word families, allofams, and the comparative method". *Cahiers de linguistique Asie orientale* 48(2): 154–158.

Thurgood, Graham, Randy J. LaPolla (eds.). 2003. *The Sino-Tibetan languages (Vol. 3)*. London: Routledge.

Ting Pang-Hsin, Hóngkai Sun. 2000. *Hàn-Zàngyǔ yánjiu de lìshi huígù* [Retrospective history of Sino-Tibetan studies]. (*Hàn-Zàngyǔ tóngyuáncí yánjiu* [Cognate words in Sino-Tibetan languages], 1). Nanning: Guǎngxī Mínzú Chubanshè.

Ting Pang-Hsin, Hóngkai Sun. 2001. *Hàn-Zàng, Miáo-Yáo tóngyuáncí zhuantí yánjiu* [Special topics in Sino-Tibetan and Hmong-Mien comparison]. (*Hàn-Zàngyǔ tóngyuáncí yánjiu* [Cognate words in Sino-Tibetan languages], 2). Nanning: Guǎngxī Mínzú Chubanshè.

Trombetti, Alfredo. 1905. *L'unità d'origine del linguaggio*. Bologna: Libreria Treves di Luigi Beltrami.

Turin, Mark. 2007. *Linguistic diversity and the preservation of endangered languages: A case study from Nepal*. Kathmandu: International Centre for Integrated Mountain Development (ICIMOD).

Vansina, Jan. 1995. New linguistic evidence and 'the Bantu expansion'. *The Journal of African history* 36(2): 173–195.

Verkerk, Annemarie. 2017. Phylogenies. *Language dynamics and change* 7(1): 127–140.

Voegelin, Charles Frederick, Florence Marie Voegelin. 1964-1965. *Languages of the World: Sino-Tibetan fascicle one-five*. Ms. Archives of languages of the world, Anthropology department, Indiana University.

Wichmann, Søren, Arpiar Saunders. 2007. How to Use Typological Databases in Historical Linguistic Research. *Diachronica* 24(2): 373–404.

Wichmann, Søren. 2013. Diachronic Stability and Typology. In: Claire Bowern, Bethwyn Evans (eds.). *The Routledge Handbook of Historical Linguistics*: 212–224. London: Routledge.

Xíng, Gōngwǎn. 1991a. Guānyú Hànyǔ Nándǎoyǔ de fāshēng xué guānxì wèntí — L· Shājiā'ěr "Hànyǔ Nándǎoyǔ tóngyuán lùn" shùpíng bǔzhèng [On the genetic relationship of Chinese and Austronesian——Review and corrections of L. Sagart's 'Sino-Austronesian theory']. *Mínzú yǔwén* 3: 1–14.

Xíng, Gōngwǎn. 1991b. Hànyǔ Nándǎoyǔ shēngmǔ de duìyìng — L· Shā jiā'ěr "hànyǔ nándǎo yǔ tóng yuán lùn" shùpíng bǔ zhèng [Sound correspondences in the initials between Chinese and Austronesian — Review and corrections of L. Sagart's 'Sino-Austronesian theory']. *Mínzú yǔwén* 4: 23–35.

Xíng, Gōngwǎn. 1991c. Hànyǔ nándǎo yǔ shēngmǔ jí yùnwěi fǔyīn de duìyìng — L· Shā jiā'ěr "hànyǔ nándǎo yǔ tóng yuán lùn" shùpíng bǔzhèng [Sound correspondences in the initials and final consonants between Chinese and Austronesian—Review and corrections of L. Sagart's 'Sino-Austronesian theory']. *Mínzú yǔwén* 5: 13–25.

Zhang, Menghan, Wu-Yun Pan, Shi Yan, Li Jin. 2018. Phonemic evidence reveals interwoven evolution of Chinese dialects. *arXiv preprint arXiv:1802.05820*.

Zhang, Menghan, Shi Yan, Wuyun Pan, Li Jin. 2019. Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature* 569(7754): 112–115 doi.org/10.1038/s41586-019-1153-z.

Zhang, Shuya, Guillaume Jacques, Yunfan Lai. 2019. A study of cognates between Gyalrong and Old Chinese. *Journal of Language Relationship* 17(1): 73–92.

Zhèngzhāng Shàngfāng. 2003. *Shànggǔyīnxì* [Old Chinese Phonology]. Shanghai: Shànghǎi jiàoyù chūbǎnshè.

Георг Орланди. Еще раз к вопросу об истории и степени обоснованности бинарной модели классификации сино-тибетских языков

В рамках сино-тибетской гипотезы, как правило, постулируется бинарное разделение всех языков семьи на китайскую (синитическую) и тибето-бирманскую ветви. В данной статье кратко, но исчерпывающе излагается общая история гипотезы, за которой сле-

дует критическое обсуждение методологических недостатков недавно предложенной байесовской классификации, якобы подтверждающей достоверность такого бинарного разделения. В завершение статьи обсуждаются возможные пути дальнейшего развития сравнительного изучения трансгималайских языков.

Ключевые слова: трансгималайские языки; сино-тибетские языки; байесовские классификации; филогенетическая классификация языков.