

9 To which language family does Chinese belong, or what's in a name?

George van Driem

There are at least five competing theories about the linguistic prehistory of Chinese. Two of them, Tibeto-Burman and Sino-Tibetan, originated in the beginning of the 19th century. Sino-Caucasian and Sino-Austronesian are products of the second half of the 20th century, and East Asian is an intriguing model presented in 2001. These terms designate distinct models of language relationship with divergent implications for the peopling of East Asia. What are the substantive differences between the models? How do the paradigms differently inform the direction of linguistic investigation and differently shape the formulation of research topics? What empirical evidence can compel us to decide between the theories? Which of the theories is the default hypothesis, and why? How can terminology be used in a judicious manner to avoid unwittingly presupposing the veracity of improbable or, at best, unsupported propositions?

1. The default hypothesis: Tibeto-Burman

The first rigorous polyphyletic exposition of Asian linguistic stocks was presented in Paris by the German scholar Julius Heinrich von Klaproth in 1823.¹ His *Asia Polyglotta* was more comprehensive, extended beyond the confines of the Russian Empire and included major languages of East Asia, Southeast Asia and Polar America. Based on a systematic comparison of lexical roots, Klaproth identified and distinguished twenty-three Asian linguistic stocks, which he knew did not represent an exhaustive inventory. Yet he argued for a smaller number of phyla because he recognized the genetic affinity between certain of these stocks and the distinct nature of others. One of the major linguistic phyla identified by Klaproth was the language family which comprised Burmese, Tibetan and Chinese and all languages which could be demonstrated to be genetically related to these three.

Klaproth explicitly excluded languages known today to be members of the Daic or Kra-Dai family, e.g. Thai, or members of the Austroasiatic family, e.g. Vietnamese and Mon (1823a: 363–5). Yet Klaproth did not devise labels for each of the many distinct language phyla which he identified in Asia. From 1852 onwards, John Logan became one of the first to use the term ‘Tibeto-Burman’ in print for the language family identified by Klaproth, and to which Logan added Karen and other related languages.

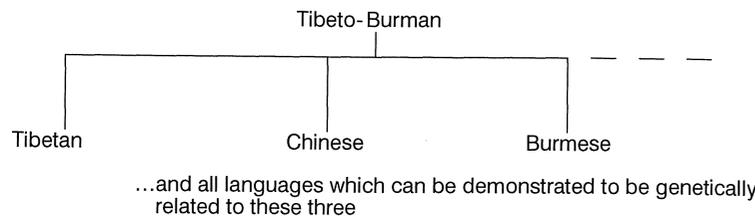


Figure 9.1 One of the language families identified by Julius Heinrich von Klaproth in his polyphyletic view of Asian linguistic stocks (1823a, 1823b). He explicitly excluded languages today known to be Daic, e.g. Thai, or Austroasiatic, e.g. Mon, Vietnamese.

Yet Logan, like many other scholars of his day in the British Isles, was an adherent of the Turanian theory dreamt up by Friedrich Max Müller in Oxford. So he treated Tibeto-Burman as an ingredient in this hypothetical Turanian family, which supposedly encompassed all languages of the world other than the Indo-European and Afroasiatic languages. Logan later also coined the label 'Chino-Tibetan' for a subset of ancient Tibeto-Burman tribes between East and Central Asia (1856: 16).² Subsequently, Charles Forbes observed:

The term 'Tibeto-Burman' has latterly crept into use as the convenient designation of a very large family of languages which appear more or less to approximate to each other. (1878: 210)

Scholars such as Bernard Houghton, who worked on languages in Burma, followed Klaproth in recognizing Chinese to be a member of this Tibeto-Burman family. Houghton observed that in Tibeto-Burman far-reaching phonological change had altered the appearance of many shared roots, particularly in the 'tonic languages' which had 'suffered much from phonetic decay'. False cognates that look alike ought not to be confused with genuine shared Tibeto-Burman roots:

If many such exist in Burmese, where phonetic decay is comparatively moderate, how much more must it be the case in extreme cases like Chinese (even the re-construction of the old sounds in this language barely brings it to the same stage as modern Burmese) and Sgaw-Karen, in which latter every final consonant, even nasals, has been elided. (Houghton 1896: 28)

Robert Cust likewise followed Klaproth in treating 'Tibeto-Burman', including Karen, as a family distinct from the 'Tai' and the 'Mon-Anam' families (1878).

Epistemologically, Klaproth's model makes the fewest assumptions and thus continues to represent the most agnostic theory about the genetic relationship of Chinese. The Tibeto-Burman theory asserts that Tibetan, Burmese and Chinese are genetically related. Furthermore, the theory assumes that there is a family of languages that can be demonstrated to be genetically related to these three languages, and that, at this reconstructible level of relationship, Tibeto-Burman

excludes both the Daic or Kra-Dai languages and the Austroasiatic languages. No new nomenclature is proposed. Tibeto-Burman is used in its original sense to denote the family tree recognized by Julius von Klaproth and accepted by scholars such as Forbes, Houghton and Cust. The Tibeto-Burman theory makes no explicit assertions about the internal subgrouping of the family. So, what is the evidence for the Tibeto-Burman theory?

A vast body of data and comparative work has come to fill the literature on Tibeto-Burman ever since Nicolaes Witsen published the first Tibetan word list and first specimens of Tibetan script in the West in 1692. Most of this literature is cited in the bibliography of my handbook (van Driem 2001), and a number of outstanding contributions have appeared since, e.g. Burling (2004), Coupe (2003), Genetti (2003), Haller (2004), Hari and Chhegu (2004), Hildebrandt (2003), Jacques (2004), Lahaussais (2002), Opgenort (2004, 2005), Plaisier (2005), Strahm and Maibaum (2005), Turin (2005), Wāng (2004), Watters (2002, 2004). All early and recent descriptions of Tibeto-Burman languages support the Tibeto-Burman theory. Comparative historical studies, reconstructions of Proto-Tibeto-Burman and of Tibeto-Burman subgroups such as Old Chinese, all bear out Klaproth's original model, even when some of the scholars who have marshalled this evidence entertained different, less agnostic theories of language relationship, e.g. Shafer (1963, 1966, 1967, 1968, 1974), Benedict (1972, 1976), Matisoff (2003).³

As the most agnostic and best supported theory about the genetic affinity of Chinese, the Tibeto-Burman theory constitutes the default hypothesis. No additional evidence need be adduced to bolster the case of Tibeto-Burman. Rather, the burden of proof lies on proponents of theories that make a greater number of assertions about the genetic relationship of Chinese. We shall now turn to four of these other theories and assess the weight of evidence in their favour.

2. Tibeto-Burman proper vs. pinioned 'Tibeto-Burman'

Both monophyletic models, Indo-Chinese or Turanian, lumped most Asiatic languages into a single grand stock and obscured the genetic position of Chinese. Adherents of either Indo-Chinese or Turanian remained confused about Chinese and undertook to treat Sinitic as something outside of Tibeto-Burman. Müller's Turanian was mentioned above. Indo-Chinese was the invention of the Scottish travelling scholar John Leyden (1806, 1808), whose hypothetical language family encompassed all faraway tongues of Eurasia and Oceania. The anomalous treatment meted out to Chinese within both monophyletic conceptions was due to various causes.

Race and language used to be confused by many laymen and even by some linguists. Much was made of the fact that the Chinese appeared to be racially different from the Burmese, for example; though linguists such as Klaproth and Müller stressed the absolute distinction in principle between race and language, many remained deaf to their explanations.⁴

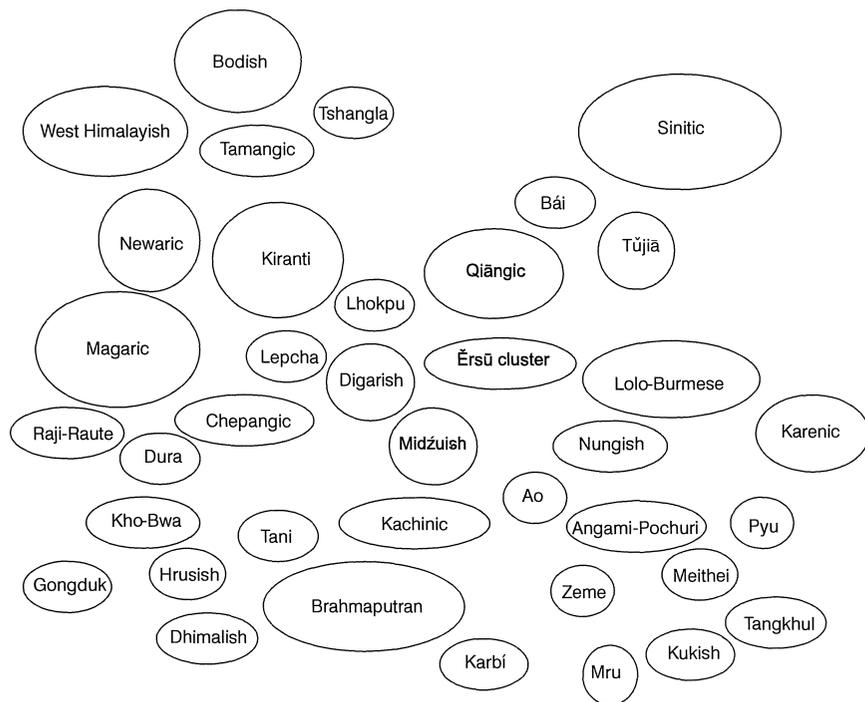


Figure 9.2 Tibeto-Burman subgroups identified since Julius von Klaproth. Brahmaputran may include Kachinic and Dhimlish. Various other subgrouping proposals are discussed in my handbook (van Driem 2001).⁵

A second source of confusion was language typology. In 1782, Rüdiger proposed that structural differences between languages were the result of differences in the stage of development attained by various language communities. Language types therefore reflected a hierarchy of thought. The morphological simplicity of Chinese puzzled typologists who wondered how a people speaking a language at the bottom of the ladder in terms of structural complexity could have produced a great civilization.

In 1854, Arthur de Gobineau attempted to resolve this quandary by speculating that Chinese, whilst a primitive tongue, had been successful because the language was male. Half of the world's languages, he reasoned, were male, and half were female. Male languages are naturally endowed with greater precision than female languages, which are replete with vague notions and emotive terms. Other linguists like Ernest Renan resolved the apparent contradiction in their minds by ascribing a 'sècheresse d'esprit et de cœur' and all sorts of other nasty attributes to the Chinese. Wilhelm von Humboldt and August Friedrich Pott were amongst the linguists who challenged racist notions propagated by the language typologists.

Scholars in Germany working in the tradition of Klaproth had sound intuitions about historical phonology and lucid insights into its implications for historical

grammar. Carl Richard Lepsius insisted that Chinese tones were phonological and could not be equated with either musical tones or intonation. In comparing Tibetan and Southern Chinese dialects with Mandarin, Lepsius recognized that 'die Chinesischen Tonaccente' had arisen from the loss of syllable finals and the loss of distinctions between older syllable initials. Therefore, Lepsius argued both against the diachronic implication of the ladder of language evolution invented by the typologists and against the independent genetic status accorded to Chinese by the monophyleticists. In terms of their historical phonology, Chinese dialects did not represent 'embryonische unentwickelte Ursprachen'. Rather, Chinese dialects were much evolved languages whose apparent 'Einsilbigkeit' was the result of sound changes which had obscured their genetic proximity to their closest cousins.

These diachronic developments had not only reduced phonological distinctions in the roots, but had in the process also partially or wholly obliterated smaller flexional elements that differentiated words which had at one time been morphologically articulate (Lepsius 1861: 472, 492–6). Based on lexical comparison with other Tibeto-Burman languages such as Lepcha, Kuki-Chin and Tibetan, Wilhelm Grube arrived at the same conclusion (1881: 19–20). A century later, Søren Egerod eloquently reiterated this sinological view:

Quand le chinois apparaissait comme une langue écrite sur les bronzes ou dans de vieilles œuvres comme le *Shū Jīng*, nous n'avions plus de doute que nous ayons devant nous une langue dont la morphologie était développée, mais dont l'écriture était de telle nature que cette morphologie se cachait assez largement. On a continué d'écrire pendant très longtemps des expressions morphologiques différentes d'une racine avec un caractère unique. Ainsi, quand on lisait un texte, on suppléait la lecture par une interprétation de la langue écrite. (1972 [1967]: 101)⁶

Wilhelm Schott, another adherent of Klaproth's polyphyletic model, argued against both Turanian and Indo-Chinese. In a wonderfully worded letter now kept at the Royal Asiatic Society in London, Schott tried to persuade Brian Houghton to abandon Müller's Turanian theory. Likewise, in the proceedings of the Royal Academy in Berlin, Schott complained that the term *indo-chinesisch* was 'eine unpassende Benennung' because the three best known languages of Southeast Asia, Burmese, Vietnamese and Thai, were known to belong to three separate language families (1856: 161–2). Schott used the term 'Siam-sprachen' for the Daic or Kra-Dai languages, but he invented no term for the other two language families identified by Klaproth. Rather, somewhat diffidently, Schott resigned himself to the fact that people might go on using the term *indo-chinesisch*, but cautioned that those using the label ought not to adopt the uninformed monophyletic model that it represented.

Here history teaches us an important lesson. The English term 'Indo-Chinese', adopted in German as *indochinesisch*, with or without a hyphen, remained popular, and inexorably along with the catchy name came the model of genetic relationship

that it denoted. As a consequence, much subsequent scholarship either uncritically accepted the family tree or attacked the language family from within, only to end up belatedly with the same set of language families at the end of the 20th century that Klaproth had identified for this part of the world at the beginning of the 19th century.

Unfettered by the Indo-Chinese paradigm, Francis Mason recognized the Mon-Khmer-Kolarian or Austroasiatic family when he established the genetic relationship between the Munda languages of the Indian subcontinent and the Mon-Khmer languages of Southeast Asia (1854, 1860). By contrast, working within the monophyletic paradigm, Ernst Kuhn had to extricate Austroasiatic from Indo-Chinese to get 'zwei Hauptgruppen von Sprachen', one of which encompassed 'die Sprachen von Annam, Kambodscha und Pegu', whereas the other group lumped together 'die Sprachen von Tibet, Barma, Siam und China' (1883, 1889), to which Kuhn also added Karen and the languages of the Himalayas.

Subsequently, several tendencies conspired to take Chinese out of Tibeto-Burman and assign it to the wrong language family. Ignorance of Chinese historical phonology and widespread preconceptions about race led scholars like American philologist John Avery⁷ to treat Chinese as something outside of Tibeto-Burman (1885). At the same time, scholars of Indo-Chinese, unlike scholars who followed Klaproth, proved unable to distinguish between inherited and borrowed vocabulary in Thai. Konow and Grierson criticized the Indo-Chinese and Turanian views but adopted a cardinal legacy of its proponents by putting Chinese together with Daic or Kra-Dai into a 'Siamese-Chinese' family, distinct from 'Tibeto-Burman' (Grierson 1904, 1909). This bifurcation into a western and an eastern branch, which Kurt Wulff (1934) called 'das Tibeto-Barmanische' and 'das Siamesisch-Chinesische', became the hallmark of the Indo-Chinese model, shown in Figure 9.3. As long as the name Indo-Chinese remained in use, those who employed the term adopted the model it designated, e.g. Georg von der Gabelentz (1881), Emile Forchhammer (1882), August Conrady (1896), Berthold Laufer (1916).

Indo-Chinese was renamed 'sino-tibétain' by Jean Przyluski in 1924, and the new name gradually caught on. Finally, in the 1930s, Robert Shafer decided to take Daic out of Indo-Chinese, but on a pilgrimage to Paris he was convinced by

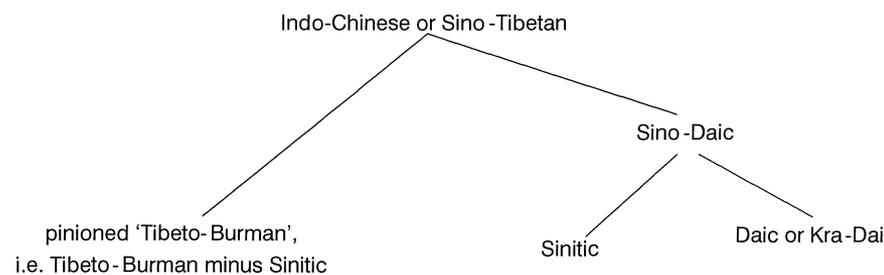


Figure 9.3 The Indo-Chinese or Sino-Tibetan theory: Daic or Kra-Dai has been excluded since the Second World War.

Maspero to leave Daic inside Sino-Tibetan (Shafer 1955: 97–8). So Paul Benedict was able to scoop Shafer by removing Daic in 1942 after he too had joined Kroeber's Berkeley project. Shafer patently rejected a bifurcation of the language family into 'Tibeto-Burman' and 'Siamese-Chinese'. Therefore, aside from Daic, which Shafer retained against his better intuitions, his Sino-Tibetan consisted of five divisions, i.e. Sinitic, Bodic, Burmic, Baric and Karenic. Benedict, however, stuck with the Indo-Chinese model which had been passed down from generation to generation, and after the excision of Daic the resultant tree effectively brought back the family to Klaproth's original Tibeto-Burman with one salient difference. The postulation of a reduced 'Tibeto-Burman' subgroup, from which Sinitic has been excised and which is coordinate with Sinitic under the top node, remains the sole defining trait of the Sino-Tibetan model.⁸

Sino-Tibetan, therefore, is essentially a subgrouping hypothesis that posits a pinioned 'Tibeto-Burman' taxon, as opposed to the originally conceived Tibeto-Burman family which I shall continue to call Tibeto-Burman proper. The 'Tibeto-Burman' of the Sino-Tibetanists encompasses all languages of the family other than Sinitic. Since these languages have never been shown to share any common innovation that would set them off collectively as a subgroup against and on par with Sinitic, the Sino-Tibetan hypothesis remains unsupported by evidence to date. Matisoff has continued to reproduce the Sino-Tibetan family tree as an article of faith (2000, 2003), but, when challenged to defend this subgrouping hypothesis, he has failed to adduce any shared innovation or compelling lexical evidence for pinioned 'Tibeto-Burman'.

Some subgrouping proposals are ambivalent with regard to a choice between Tibeto-Burman proper or Indo-Chinese, e.g. Shafer's Bodic or Burmic, in that these proposals could be subgroups within either model. This cannot be said for either Sino-Bodic or pinioned 'Tibeto-Burman'. Sino-Bodic essentially dates back to Klaproth's own observation that Tibetan appeared to be genetically closer to Chinese than either was to Burmese (1823a: 346, 356, 365). Additional evidence in support of the Sino-Bodic hypothesis was presented by Simon (1927–9), Shafer (1955, 1966, 1967, 1968, 1974), Bodman (1980) and myself (van Driem 1997). My coinage 'Sino-Bodic' reflects Shafer's view that the alleged affinity is between Sinitic and the nebulously delineated Bodic, not just between Sinitic and Bodish.⁹ Moreover, a complex relationship of borrowing may have existed between Chinese and languages such as Tibetan at various stages of their history, and this process may have been further complicated by a contact phenomenon described by Ferlus as 'hypercorrection by affected imitation', masking a layer of borrowings which has hitherto not been clearly identified in historical comparative studies (2003: 274).

Matisoff was able to eliminate only 12 of the 39 specific Sino-Bodic correspondences, namely 40, 48, 49, 56, 58, 60, 61, 64, 66, 69, 74 and 77 in Matisoff's numbering.¹⁰ A few more correspondences were unconvincingly challenged. For example, the alternative cognate set which Matisoff proposes for correspondence 75 is contestable, and his alternative explanation for correspondence 46 makes less semantic sense. Given the speciousness of some of Matisoff's etymologies

(e.g. 1992, cf. Sagart 1994b), his semantic sensibilities, as diagrammed in his 'metastatic flow charts' (e.g. 1978), are not always to be trusted. In addition to Sino-Bodic lexical isoglosses, my article presented Tibeto-Burman correspondences for which the phonological match with Sinitic is generally better for Bodic than for cognate forms from other branches of Tibeto-Burman.¹¹ In addition to leaving most of the Sino-Bodic evidence unassailed, Matisoff failed to address relevant evidence adduced by Shafer and Bodman.

So, in contradistinction to Sino-Tibetan, for which no evidence has ever been presented, lexical and morphological evidence warrants entertaining Sino-Bodic as a viable working hypothesis about the closest relatives of Sinitic within Tibeto-Burman. Stanley Starosta accepted Sino-Bodic and incorporated the hypothesis in his East Asian phylogeny, discussed below. Matisoff rails that the evidence for Sino-Bodic might be 'turning all our ideas about ST/TB subgrouping upside down' (2000: 366). Matisoff's histrionic reaction and strident tone must be seen as a sally not against Sino-Bodic *per se*, but against the threat which Sino-Bodic poses to Sino-Tibetan, the subgrouping hypothesis about pinioned 'Tibeto-Burman' that he inherited from his mentor Paul Benedict in 1968.

It has been suggested that perhaps the distinction between what is reconstructed as *a vs. *ə (or *ā vs. *a) in current versions of Proto-Sinitic might conceivably represent an ancient 'Sino-Tibetan' distinction lost in a merger which affected all 'Tibeto-Burman' languages, but this idea has not been pursued. Not all branches of Tibeto-Burman have been scrutinized in this regard, and ultimately such a conjecture cannot be sustained on the basis of an unwarranted limitation of the available evidence. A tentative cursory study by Jean Robert Ongenort has shown that whereas Old Chinese *a (or *ā) appears most often to correspond to an /a/ in modern Kiranti languages, the Tibeto-Burman vowel reflected by Old Chinese *ə (or *a) appears to have engendered a more complex pattern of vocalism in Kiranti (pers. comm., July 2005).

More importantly, even if the Old Chinese distinction were shown not to be reflected outside of Sinitic, then there is yet no way of knowing, given the present state of the art, whether the Sinitic distinction does not represent one of many innovations which define Sinitic as a branch of Tibeto-Burman. In light of correspondences between Kulung and Old Chinese long vowels, Tolsma previously raised the question whether Old Chinese long vowels are a Tibeto-Burman retention 'or that a sound change which yielded long vowels took place as early as the Old Chinese period' (1999: 497). Persistent misunderstandings about diachronic developments in Slavic accentuation are especially instructive in this regard (Kortlandt 2003). Czech vowels show a phonological length contrast, but the ontogeny of the distinction is complex. At the present state of our knowledge, even if the distinction were not to be shared with Kiranti, the most parsimonious explanation would be that the Old Chinese distinction between *a vs. *ə represents a split in Sinitic rather than a merger shared by all other Tibeto-Burman languages.

Another last straw for a drowning hypothesis to grasp at is held out by the idea that pinioned 'Tibeto-Burman' shares some lexical items not found in Sinitic.

However, each and every branch of Tibeto-Burman, including Sinitic, lacks reflexes of some common Tibeto-Burman roots. Gongduk, for example, resembles Chinese in lacking a reflex of the ubiquitous Tibeto-Burman root for 'pig', the most recently postulated reconstruction of which is still *p^wak (Benedict 1972: 217; Matisoff 2003: 662). Yet pork plays an important role in Gongduk culture just as it always has in Chinese cuisine. The diversity in vocabulary and grammar in Tibeto-Burman may not be as great as in Indo-European or Afroasiatic. Yet the Tibeto-Burman language family is not at all as cohesive a group as was once assumed.

Old Chinese represents an older stage of Sinitic, a phonologically innovative branch. So it is to be expected that the reconstructible Old Chinese syllabary should, because of its time depth, resemble other Tibeto-Burman languages more closely than do modern Sinitic languages. Yet the recent improved reconstructions by Baxter and Sagart differ dramatically from Karlgren's pioneering work and now make Old Chinese look like a very run-of-the-mill Tibeto-Burman language from the Himalayan perspective. The Sino-Tibetan view of Chinese as the odd man out is not just sustained by a lack of familiarity with recent breakthroughs in Sinitic reconstruction. More typically, this view is nourished by a lack of familiarity with languages of other branches of the family such as Gongduk, Hrusish or the Kho-Bwa cluster, all spoken in the Tibeto-Burman heartland closer to the language family's centre of gravity and all just as divergent from 'mainstream' Tibeto-Burman as are the modern Sinitic languages.¹²

It is natural to assume that the linguistic ancestors of Sinitic might have lost some of their original Tibeto-Burman lexicon on their long trek from the greater Himalayan region to the North China plain. Lured as they were by the riches of the advanced Neolithic civilizations along the Yellow River, it would also have been natural for them to adopt new vocabulary from the affluent pre-Tibeto-Burman resident populations of the North China plain. This migration may have taken place at the dawn of the Shāng dynasty, when common Tibeto-Burman had probably already broken up into the major branches attested today. At present, there is no evidence that the rest of the language family was still a unity at the time that Sinitic split off. Sino-Tibetan designates the abidingly incorrect Indo-Chinese construct in its most recent incarnation. The fact that there is no evidence for Sino-Tibetan does not diminish the fact that the hypothesis represents an intrinsically interesting proposition. Yet the theory which makes the least assumptions and is best supported by evidence is the default, and after nearly two centuries Klaproth's Tibeto-Burman is still the default hypothesis.

3. Grand monophyletic views: Sino-Austronesian

The old monophyletic views failed to correctly appraise the genetic position of Chinese. Turanian had generally been abandoned by the end of the 19th century, whereas Indo-Chinese still survives though it has been whittled down and renamed Sino-Tibetan. A twist in the history of linguistics is that new grand monophyletic models have been developed to genetically unite many of the languages of eastern

Eurasia and in the process define the genetic position of Chinese. Here three theories will be examined, i.e. Sino-Austronesian, Sino-Caucasian and East Asian. All three theories are fascinating and will no doubt continue to influence our conjectures about prehistory, as the evidence is accumulated, sifted and tested.

Sino-Austronesian is a new theory first presented at a conference in Texas in 1990. The Sino-Austronesian theory is an ongoing story which continues to unfold in fascinating and unexpected ways. In the first version of Sino-Austronesian, Sagart (1990, 1991, 1993) held that the evidence warranted entertaining the view that Sinitic is genetically related to Austronesian rather than, or more so than, to 'Tibeto-Burman'. The claim of a family comprising just 'Chinese plus Austronesian' was generally rejected, e.g. Blust (1995), Li (1995), Pulleyblank (1995) and Starostin (1995a, 1995b), but some, including myself, gave the intriguing evidence adduced by Sagart a fair hearing.

At the time, I speculated that the correspondences adduced by Sagart might be the residue of a contact situation between ancient Northern Tibeto-Burmans, i.e. Sinitic or Sino-Bodic peoples, and ancient Austronesians (van Driem 1998). I proposed that proto-Austronesians were the behind littoral cultures which lay south of the Yangtze delta such as the Hémùdù culture on Hángzhōu Bay in Zhèjiāng, the Dàpènkēng of Formosa, the Fùgùódūn of Quemoy and related Neolithic cultures of Fukien of the 5th and early 4th millennia BC. The contact situation between Proto-Austronesian and an ancient variety of Tibeto-Burman which accounted for Sagart's correspondences ensued upon the northward expansion of Proto-Austronesians from south of the Yangtze delta, giving rise to the Lóngshān interaction sphere which emerged in the 4th and 3rd millennia BC and connected coastal cultures from north to south, such as the Dàwènkǒu assemblage in Shāndōng, the Qīngliángǎng culture of northern Jiāngsū, and the Mǎjiābāng culture of the Yangtze delta.

The second version of Sino-Austronesian came to encompass 'Chinese plus Tibeto-Burman plus Austronesian' after a number of 'direct Proto-Austronesian-Proto-Tibeto-Burman comparisons not involving Old Chinese, or with better semantic agreement between Proto-Austronesian and Proto-Tibeto-Burman' led Sagart to concede that the facts now 'render less likely the possibility that the material shared by Old Chinese and Tibeto-Burman reflects a contact situation. They suggest that Tibeto-Burman languages may stand closer to Chinese (and to Proto-Austronesian) than I had originally assessed' (1994a: 303). In addition to reintroducing Tibeto-Burman into the equation, Sagart had improved his comparisons by replacing Otto Dempwolff's reconstruction of *Uraustronesisch*, taxonomically comparable to Malayo-Polynesian, with Robert Blust's proto-Austronesian reconstructions. Sagart also addressed relevant methodological issues (1995a, 1995b, 1995c).

The third and most recent incarnation of Sino-Austronesian (Sagart 2005a) is the most interesting and methodologically most rigorous. Li Fang-kuei's reconstruction of Old Chinese has been replaced with Sagart's own 1999 reconstruction. The comparanda now feature only Proto-Austronesian reconstructions in the accepted system of sound correspondences, and Sagart's comparisons

rigorously distinguish between etyma reflected at the Proto-Austronesian and the Malayo-Polynesian levels. In the process, the evidence in support of Sino-Austronesian has grown rather than diminished.

Sagart's Sino-Austronesian theory is now based on 75 lexical comparisons, 61 involving 'basic vocabulary' and 14 items of 'cultural vocabulary'. The Austronesian comparanda are taken from the Proto-Austronesian level or involve reconstructed 'Proto-East-Coast-Linkage'. The latter used to be something of a taxon within Austronesian, although the group has recently been abolished by Sagart's own 2004 revision of Austronesian phylogeny. Sagart's new Austronesian phylogeny, based on arguments advanced by Haudricourt (1956) and new insights into the time depth of Kra-Dai or Daic as a taxon (Ostapirat 2005), has both solved the 'Austro-Thai' problem and incorporated Kra-Dai into the Sino-Austronesian equation (Sagart 2002, 2004, 2005a, 2005b). For 69 out of the 75 correspondences, the Tibeto-Burman comparanda are reconstructed Old Chinese forms. For 45 of these 69 comparisons Sagart is able to adduce an additional cognate from another language, usually Tibetan or Burmese. In three instances, a Tibeto-Burman reconstruction by Peiros and Starostin (1996) is used, and in several cases the comparanda are taken from a modern language, e.g. Chepang, Lushai or Lepcha. Only six of the 75 comparisons involve a non-Sinitic form only, for which Sagart found no Old Chinese cognate.

Fourteen of the 75 items are cultural vocabulary and include items relating to cereal cultivation. Their special significance lies in the fact that two salient items relating to rice cultivation are uniquely shared by Tibeto-Burman and Austronesian, whilst Austronesian and Austroasiatic do not share this vocabulary (Sagart 2003a, 2005a). One of these correspondences, Austronesian *beRas 'husked rice' vs. Tibetan *hbras* 'rice', was first pointed out by Hendrik Kern (1889: 5). Whereas Kern believed that this correspondence reflected an early borrowing which indicated whence the ancestors of the Tibetans had first acquired rice, Sagart adduces the correspondence in support of a Sino-Austronesian phylum and adds the Old Chinese cognate 糲 ^hmə-rat-s. A second rice term is Austronesian *Sumay 'rice as food' vs. Old Chinese 米 ^amij? 'grain of cereal' and Garo *may* 'paddy'. Sagart also presents correspondences between Austronesian *beCeng '*Setaria*' vs. Old Chinese 稷 ^htsik and Austronesian *Numay '*Panicum*' vs. Old Chinese 麻 and 黍 ^amaj.

The Sino-Austronesian roots adduced to date reflect the proto-meanings: body hair, bone, brain, elbow, female breast, foot, head, palm of the hand, pus, mother, egg, horn or antler, leech, snake, worm, cloud or cloudy, earth, moon, salt, sunlight, water, wind, cave or hole, year, carry, chew, close or shut, come or go, short or cut off, dig, drown or disappear, fall, flow or water or river, follow, grasp or embrace, hold something in one's fist or hold something in one's mouth, lick, meet, open, put together, ruin or damage, scrape I, scrape II, sink, sleep, speak or say, think, vomit or spit, wash, gird, bent or crooked, broad, bent, ear, far, high or tall, hot, old or grown-up, sharp, thick, this, *Setaria*, *Panicum*, husked rice, paddy, chicken, cage or enclosure, net, broom, stopper or plug, to bury or tomb, loincloth or robe, plait or braid, shoot, hunt.

Sagart's thinking about genetic relationships has by no means remained static. He describes himself as 'one of the last doubters' that Chinese was even genetically related to Tibeto-Burman. So, when he finally accepted this genetic relationship, it was naturally Sino-Tibetan that he adopted, for this model maintained a safe distance between Sinitic and all its closest relatives. However, recently, Sagart has come to question the Sino-Tibetan paradigm espoused principally by Matisoff. Tibeto-Burman has most recently come to mean non-Sinitic for Sagart, who stresses that his 'use of the term should not' be construed to imply that he is 'presently convinced that it is a valid grouping' (2006). I submit that it is less misleading then to simply say 'non-Sinitic', since 'Tibeto-Burman' is used by believers in Sino-Tibetan to denote non-Sinitic languages as if they together formed a valid taxon. In all his previous work, Sagart too used the term 'Tibeto-Burman' explicitly in this meaning. Sagart's present non-acceptance of pinioned 'Tibeto-Burman' is an implicit disavowal of the Sino-Tibetan hypothesis that may indicate that he is well on the way to accepting the original Tibeto-Burman theory first propounded in Paris some 128 years before Sagart himself was born there. By the same token, Sagart's original name 'Sino-Austronesian' is to be preferred above the newer and unwieldy 'Sino-Tibetan-Austronesian', which incorporates the name of a hypothesis from which he has dissociated himself.¹³

At the same time, Sagart is uniting several of Klaproth's language families in ways that must be catching most scholars by surprise. Sagart's new Austronesian phylogeny, with his identification of Kra-Dai as a lower-level offshoot of a Muish

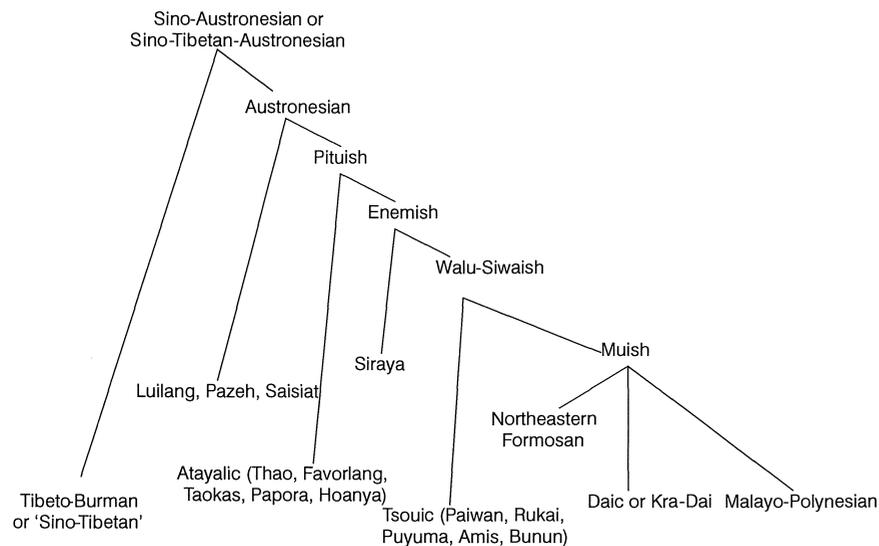


Figure 9.4 Sagart's Sino-Austronesian theory (2005), incorporating Sagart's major revision of Austronesian phylogeny (2004). Northeastern Formosan comprises Kavalan and Ketagalan. Under the Muish node, Northeastern Formosan is coordinate with the Formosan ancestor languages which gave rise to Kra-Dai and Malayo-Polynesian respectively.

ancestor language on Formosa, not only solves the Austro-Thai enigma, but also points the way towards a fundamental revision of the Austric problem.

Wilhelm Schmidt was the first to propose an Austric language family consisting of Austroasiatic and Austronesian, a later version of which even included Japanese (1906, 1930). Additional evidence in support of Austric was adduced by Kuiper (1948) and Reid (1994, 1999, 2005). August Conrady (1916, 1922) and Kurt Wulff (1934, 1942) proposed a mega-Austric superfamily consisting of Austroasiatic, Austronesian and Indo-Chinese, i.e. Kra-Dai and Tibeto-Burman. Another expanded Austric theory, Greater Austric, united Austroasiatic, Austronesian, Kra-Dai and Hmong-Mien (Blust 1996; cf. van Driem 2001: 298–302). Reid (2005: 150) is right to assess that:

With the accumulation of evidence presented by Sagart ... the concept of 'Austric' as a language family may eventually need to be abandoned in favour of a wider language family which can be shown to include both Austronesian and Austroasiatic languages but not necessarily as sisters of a common ancestor.

4. Grand monophyletic views: Sino-Caucasian

Whereas Sino-Austronesian is a new theory, Sino-Caucasian emerged from a long tradition of scholarship which sought genetic links between language isolates such as Basque and Burushaski, distant languages such as Chinese and Tibetan, and isolated families such as Yenisseeian and the languages of the Caucasus, e.g. Trombetti (1905, 1925), Bleichsteiner (1930), Bouda (1936, 1950, 1954, 1964). The chief current proponent of Sino-Caucasian is the late Russian linguist Sergei Starostin, Sagart's junior by five years.¹⁴ The four main branches of Sino-Caucasian are North Caucasian, Sino-Tibetan, Yenisseeian and Burushaski.

Even North Caucasian is itself not a universally accepted theory, but a genetic relationship proposed by Nikolai Trubetzkoy (1922) between West Caucasian, or Abkhazo-Adyghean, and East Caucasian. Evidence was adduced for this relationship by Georges Dumézil and later by various Soviet scholars. Most recently, Nikolaev and Starostin published a dictionary of reconstructed North Caucasian (1994). Two of the most interesting ingredients of the North Caucasian theory are the inclusion of the extinct Hattic language into West Caucasian, a hypothesis proposed at the beginning of the 20th century, and the inclusion of the extinct languages Hurrian and Urartean into East Caucasian, a theory proposed by Forrer (1919: 1040). Both hypotheses have been discussed elsewhere (van Driem 2001: 1057–60). Orël and Starostin have recently even added Etruscan to East Caucasian (1990).

Sino-Caucasian has undergone continual expansion, and the arguments in favour of the phylum are scattered throughout the literature, e.g. Starostin (1982, 1984, 1991, 1995a, 1995b, 2002), Nikolaev and Starostin (1984, 1994). Sino-Caucasian is just one leg of a phylogenetic centipede which unites all languages of the world within a single genetic phylum. The next higher node, Dene-Caucasian,

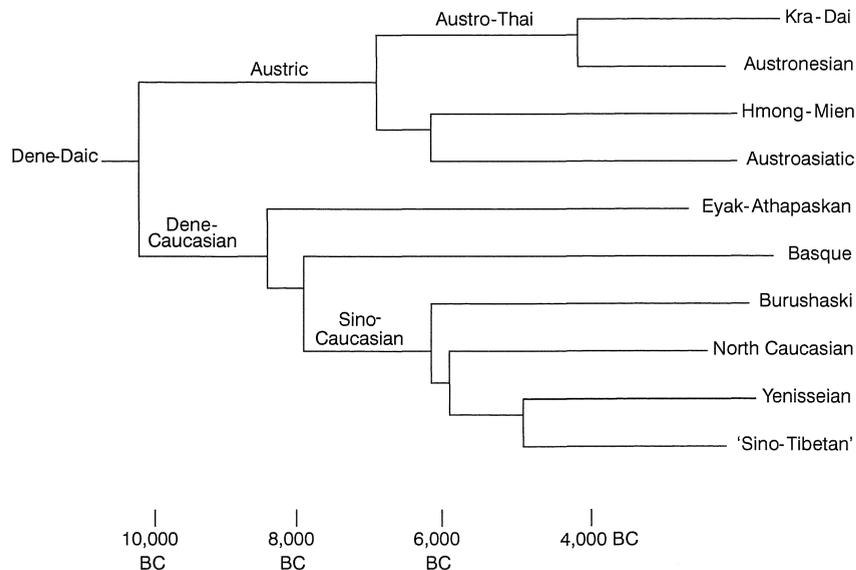


Figure 9.5 Starostin's Sino-Caucasian and Dene-Daic theories (2005). North Caucasian consists of West Caucasian, including Hattic, and East Caucasian, is taken to include Hurro-Urartean and Etruscan. The extinct languages Sumerian, Iberian and Pelasgian are also part of the equation. Starostin used the Chinese name Miáo-Yáo for Hmong-Mien.

comprises Basque and the Na-Dene languages (Starostin 1984, 1995a; Ruhlen and Starostin 1994). The treatment of the Basque material has been criticised by Trask (1994, 1995a, 1995b). Dene-Caucasian has been expanded to include extinct languages of the Iberian peninsula, about which hardly anything is known, as well as Sumerian and Pelasgian (Nikolaev 1991; Bengtson 1991).

The current state of the art in Sino-Caucasian comparative linguistics is posted on Starostin's webpage <ehl.santafe.edu>, as it appeared during the summer of 2005, where 1358 Sino-Caucasian etymologies were listed. Sino-Caucasian reconstructions are based on Starostin's reconstructed roots for North Caucasian, 'Sino-Tibetan', Yenisseeian and Burushaski. The Sino-Tibetan reconstructions correspond largely to those given in Peiros and Starostin (1996), which are based on five strategically chosen Tibeto-Burman languages, i.e. Old Chinese, Tibetan, Burmese, Jinghpaw and Lushai. Starostin's website has been strengthened by the inclusion of a Kulung dictionary provided by Gerard Tolsma, a Yamphu dictionary by Roland Rutgers and Limbu and Dumi dictionaries by myself.

In most cases, the Sino-Tibetan reconstructions in Peiros and Starostin are not reflected in all five languages, and in many cases they are supported by reflexes in only two of the five chosen languages. The same applies *mutatis mutandis* to the reconstructions posted on the website. This *modus operandi* is similar in principle to the assumption made at the Indo-European Etymological Dictionary (IED) in Leiden, whereby a form is judged to be reconstructible as a common

Indo-European root or process if the etymon in question is well reflected in any two out of twelve branches of Indo-European. The difference, of course, is that Indo-European is a language family with a well-understood history. Moreover, a modern Lushai form is not a reconstructed Mizo-Kuki-Chin etymon. So Peiros and Starostin's 'Sino-Tibetan' is somewhat analogous to a reconstruction of Indo-European based on Kurdish, French, English, Ardhamāgadhī and Norse runes.

Whenever a 'Sino-Tibetan' root is based just on reflexes in languages which according to a subgrouping hypothesis could belong to a single branch of Tibeto-Burman, such as Old Chinese, Tibetan and Kiranti as members of the hypothetical Sino-Bodic, the correspondences in question may not legitimate the reconstruction of a root at the Tibeto-Burman or 'Sino-Tibetan' level. The best analogue at present to the twelve branches of Indo-European is the model of the fallen leaves of the Tibeto-Burman tree depicted in Figure 9.2. Although a reconstruction of Proto-Kiranti consonants, for example, is available (Opgenort 2005), no reconstructions are available for most branches of Tibeto-Burman.

On the face of things, Starostin's 1358 reconstructions for Sino-Caucasian would seem to outweigh the 75 correspondences adduced for Sino-Austronesian by Sagart. However, only 130 of the 1358 Sino-Caucasian reconstructions are supported by reconstructions from all four putative member families, and only 847 additional correspondences involve reconstructed 'Sino-Tibetan' roots at all. Sino-Caucasian is not an established and generally accepted language family like Indo-European. Rather, the plausibility of Sino-Caucasian has yet to be demonstrated. So decisive evidence for Sino-Caucasian cannot be based on reconstructed etyma from only two or three of the purported constituent groups. What are we to make of the 64 Sino-Caucasian reconstructions supported only by a North-Caucasian reconstruction, the five Sino-Caucasian etyma supported by only a reconstructed 'Sino-Tibetan' root, the one postulated Sino-Caucasian root supported only by a common Yenisseeian reconstruction, and the one Sino-Caucasian root reflected only by Burushaski? Are these Sino-Caucasian roots posited merely to furnish comparanda at yet higher putative nodes such as Dene-Caucasian or Dene-Daic?

Some Sino-Caucasian correspondences are intriguing, such as the reconstruction *xGwV 'thou', synthesized from North Caucasian *ɛwV̄, Sino-Tibetan *K^wa-, Yenisseeian *kV-/ʔV̄k- ~ *gV-/ʔV̄g- and Burushaski *gu-/go- (record 241). An etymon, perhaps very much like Starostin's 'Sino-Tibetan' reconstruction *K^wa-'thou', is reflected both as an independent pronoun and in verbal agreement prefixes in different branches of Tibeto-Burman. For this reconstructed root, Starostin's 'etymological database' on the web gives only the purported Tibetan and Burmese reflexes, whereas the reconstruction would appear to be based on more than just Burmese and Tibetan. A problem with Starostin's etymological databases on the web is that they do not in fact render explicit either the empirical basis for the proposed reconstructions nor the process by which he arrives at them.

Another intriguing etymon Sino-Caucasian *=i-xGĀr- 'dry' is constructed on the basis of North Caucasian =iGwĀr, Sino-Tibetan *kār, Yenisseeian *qo(ʔ)r₁- ~ qo(ʔ)l- and Burushaski *qhar- (record 320). To this Sino-Caucasian etymon it is interesting to juxtapose Sagart's Sino-Austronesian reconstruction

*kaR 'dry', based on Sagart's reconstructed Proto-Austronesian root *-kaR 'dry', Old Chinese 乾 *kar 'dry' and Burmese *khân* 'dry up, evaporate, be exhausted (of a liquid)' (Sagart, pers. comm. 30 July 2005), whereby the Burmese final -n reflects an earlier final *-r (Matisoff 2003: 388). So, are both Sino-Caucasian and Sino-Austronesian reconstructions just disjointed parts of a bigger puzzle? Whatever the case may be, the sound laws connecting the Sino-Caucasian forms are not made explicit on the website, but some are detailed in earlier published work, e.g. Starostin (1984, 1991). Yet many Sino-Caucasian correspondences do not obey even these laws, and Starostin has invoked unspecified 'accentual factors' in the past to discount the frequent exceptions (1995a, 1995b).

Several examples taken at random are typical. Sino-Caucasian *HírkkV̄, glossed as 'male deer or goat', is extrapolated from the reconstructed North Caucasian root *whírĭV 'mountain goat', Sino-Tibetan *rjōk ~ *rjūk 'a kind of deer', Yenisiseian *ʔiʔx(V) 'male deer or billy goat' and Burushaski *har 'bull, ox' (record 66). This Sino-Caucasian root for 'deer' exists alongside four other Sino-Caucasian proto-forms for 'deer' (records 175, 472, 696 and 697) and yet another Sino-Caucasian root for 'goat', viz. *kwīʔnī, supported solely by the North Caucasian reconstruction *kwīʔnī ~ kwīʔnǎ ~ kwīʔǎ (record no. 1299). Equally unfathomable is the Sino-Caucasian reconstruction *=VʔwV̄ŋ 'go, travel', derived from North Caucasian *=VʔwVn, Sino-Tibetan *ʔwǎ (s-, -ŋ), Yenisiseian *hejV̄ŋ and Burushaski *né- (record 200).

More often than not,¹⁵ a Sino-Caucasian reconstruction is based on one or two reconstructed reflexes from the four proposed member families. Sino-Caucasian *HVIV, glossed as 'moon; burn(?)', is based solely on Sino-Tibetan *x̄'eIH, which in turn is supported by Old Chinese 燬 *x̄'ejʔ 'blazing fire' and a Proto-Kiranti root *wǎl¹⁶ (records 1338, 2656). Yet another Sino-Caucasian reconstruction *HVrV, likewise signifying 'burn', is based solely on Sino-Tibetan *rēw(H) (rec. 1252). Generally, Sino-Caucasian proto-forms rely most heavily on the North Caucasian reconstructions, which contain the most reconstructed segments to play with. In addition, proto-forms at various levels of reconstruction show much variation. Sino-Caucasian *=HixqwV̄, 'to bear, be born' is based on North Caucasian *=HiqwĀ(n), Sino-Tibetan *Ki(j) ~ Ke(j), Yenisiseian *kej- ~ *qej- ~ *gej- and Burushaski *-k 'children' (record 217). Sino-Caucasian *=HVžV̄ 'clear (of weather)' is based on North Caucasian *=Huž_Vn, Sino-Tibetan *Cǎj ~ *Cǎl, Yenisiseian *ʔēž- and Burushaski *cǎŋ ~ *cǎn, ~ *žǎŋ (record 42). Sino-Caucasian *xVx̄Hé 'hand, sleeve' is based on North Caucasian *x̄ěHē ~ *x̄ěHa 'sleeve' and Yenisiseian *xire 'arm', with the added caveat 'A very complicated picture: confusion of *kwīlʔi, *xq(w)ʔi, *xq(w)V̄ʔV̄ and *x̄V̄Hē' (record 980).

Semantics at the Sino-Caucasian level can often get a trifle vague. For example, there are seven etyma denoting 'a kind of tree', viz. record numbers 68, 252, 634, 983, 1155, 1306, 1315. There are eighteen Sino-Caucasian proto-forms signifying 'hair', viz. record numbers 130, 258, 263, 360, 554, 575, 603, 988, 1023, 1024, 1060, 1141, 1144, 1201, 1257, 1259, 1290, 1329. One of these is based solely upon, and is isomorphic with, the North Caucasian *čhwǎrǎ 'hair' (record 1290), whereas Sino-Caucasian *burV 'hair' is based solely on Burushaski *bur (record

1259). Out of the four Sino-Caucasian proto-forms denoting 'a kind of relative' (viz. records 108, 277, 284, 1027), Sino-Caucasian *q̄V̄r[H]V̄ is synthesized from North Caucasian *q̄ar[H]V 'cousin', Sino-Tibetan *K̄rij ~ *Kruj 'child-in-law', Yenisiseian *q̄ar₁- ~ *x̄ar₁- 'grandchild' and Burushaski *-rék 'sibling-in-law' (record 284).

There are five Sino-Caucasian roots denoting 'pus' (viz. records 95, 162, 760, 761, 907). The only one of these reflected in all four purported branches of Sino-Caucasian is the unwieldy *nēwxq̄wV̄, extrapolated from North Caucasian *nēwq̄ū, Sino-Tibetan *(s-)nuāk ~ *(s-)nuāŋ, Yenisiseian dǒ(?)kŋ and Burushaski *nagéi ~ *magéi 'boil, sore' (record 162). Sino-Caucasian *[b]VjV, glossed as 'an internal organ', appears to have been constructed on the basis of Sino-Tibetan *phe 'spleen' and Yenisiseian *b[aj]bVI 'kidney' (record 103). Three more Sino-Caucasian proto-forms denote 'an internal organ', viz. record numbers 354, 419, 1236. There are five reconstructed Sino-Caucasian roots meaning 'to laugh' (viz. records 16, 477, 880, 903, 957), and none are reflected in more than two of the four member families of this widespread family.

The time frame of the domestication of various cereals is called into question by two Sino-Caucasian agricultural terms, both glossed ambiguously as 'millet, rice'. Sino-Caucasian *λwīʔwV̄ has been constructed on the basis of the irregular North Caucasian root *λwīʔwV 'millet' and the shaky Sino-Tibetan *liwH ~ *liwH denoting some type of grain (record 590), whereas Sino-Caucasian *bōlǎwī is constructed from North Caucasian root *bōlǎwī ~ *bōnǎwī 'millet', Sino-Tibetan *phrē(s) 'rice' and Burushaski *bay 'millet' (record 733).

The notational intricacy of the ensemble of Starostin's reconstructions raises the question as to how much phonological complexity may plausibly be imputed to any putative proto-language. At the same time, some forms would appear to be attributable to a widespread tendency towards sound symbolism, a phenomenon recognized ever since Court de Gébelin (1774). For example, Sino-Caucasian *[p]ūHV̄ 'blow' is extrapolated from North Caucasian *pūHV, Sino-Tibetan *bū(-t), Yenisiseian *pV(j) and Burushaski *phu (record 280).

Grammatical etyma are at best vaguely supported. A Sino-Caucasian 'interrogative stem' *mV is based on a reconstructed North Caucasian interrogative stem *mV, an assumed but not really reconstructed Sino-Tibetan root *mV, an interrogative root *wi- ~ *we gleaned from Yenisiseian pronominal forms, and Burushaski *me- 'who' (record 426), but what are these comparanda precisely? The best reflected out of three Sino-Caucasian negative particle is *bV, ostensibly reflected in the reconstructions North Caucasian -bV, Old Chinese 不 *pə, Yenisiseian *-pun 'without, -less' and Burushaski *be 'not' (record 1187). There are two more, even shakier reconstructed Sino-Caucasian negative particles, viz. record numbers 1073, 1187. Some comparanda do not have much substance. The Sino-Caucasian verb 'to be', *ʔa, is based on a reconstructed North Caucasian auxiliary *=a ~ *=i, a poorly supported Sino-Tibetan locative or object marker *ʔǎ* ~ *x̄ǎ, an unexplained Yenisiseian reconstruction *ʔa and the Burushaski reconstruction *b-a- 'to be' (record 861).

5. Sino-Austronesian vs. Sino-Caucasian

How do Sino-Austronesian and Sino-Caucasian compare? The first difference involves the many degrees of freedom in Starostin's reconstructions as compared with Sagart's Sino-Austronesian. The comparanda in long-range comparisons are themselves reconstructions, and an element of subjectivity enters into the choice of reconstructions, which, at various levels, are usually Starostin's own. Given his stated aim of building a genealogical tree of all of the world's languages and the reduction of the number of nodes to common ancestors of particular language families, this multiple leeway in the choice of reconstructions cannot but afford ample room for the harmonization of phonological shape and meaning of constructed proto-forms, whether or not such a process is a conscious one. In the Sino-Austronesian comparison, by contrast, Sagart utilizes Blust's reconstructions for Austronesian along with just a few of his own. The semantics of Old Chinese forms is arguably as attested in the texts. Sagart's 1999 reconstruction of Old Chinese is largely corroborated by Baxter's reconstruction (1992, 1995), particularly where the rimes are concerned. Moreover, Sagart's reconstruction takes into account earlier reconstructions such as that of Jaxontov (1965), Lǐ Fāngguì (1971, 1974, 1976, 1983), Pulleyblank (1984, 1991), Zhèngzhāng Shàngfāng (1987) and Starostin (1989).

Starostin (1995a) once claimed to have found thirteen semantically precise Sino-Caucasian matches on Jaxontov's 33-word list. By contrast, Sagart's Sino-Austronesian material contains only seven semantically close matches on the Jaxontov list, i.e. including the numeral 'one' (Sagart 2005a). However, an average of between one and two phonological segments match per lexical comparison in Starostin's thirteen best correspondences, whereas an average of about three segments match phonologically in Sagart's seven correspondences. Calculations of this type involve a number of arbitrary decisions. Whereas an average of between three and four phonological segments per lexical comparison match in Sagart's overall list, the score is lower on the short list, simply because two of the seven items, viz. 'one' and 'this', consist of only two segments. More generally, however, this discrepancy in the number of phonological matches per adduced lexical comparison characterizes the entire corpus of correspondences adduced by Starostin <ehl.santafe.edu> and Sagart (2005a). Often enough, as in many of the examples extracted above from Starostin's website, only one phonological segment seems to match in a comparison. At present, therefore, Sagart's Sino-Austronesian would appear to come somewhat closer to attaining the rigour of the first sound laws formulated by Lambert ten Kate in 1710 and 1723 than does Starostin's Sino-Caucasian.

Another difference between the two theories of distant relationship is that several morphological processes have been found to be shared by Tibeto-Burman and Austronesian. No Sino-Caucasian shared morphology is in evidence, and most Sino-Caucasian grammatical morphemes are shaky. By contrast, the Tibeto-Burman nominalizing suffix *<-n>, intransitive prefix *<m-> and valency-increasing prefix *<s-> appear to be related to the Proto-Austronesian

nominalizing and goal focus marker *<-ən>, actor focus marker *<m- ~ -m-> and instrumental or beneficiary focus prefix *<Si-> respectively, all three morphemes being processes 'which form the backbone of Austronesian verbal morphology' (Sagart 2005a: 168–71). Sagart also proposes that the distributive marker *<-ar-> might be a morphological process shared by both families.

The Sino-Tibetan problem explained in the first half of this article presents a serious impediment to both Sino-Austronesian and Sino-Caucasian comparison, since both implicitly incorporate the Sino-Tibetan hypothesis and are thus built upon an unsupported assumption about the genetic position of Sinitic with respect to its closest relatives. The assumed veracity of the Sino-Tibetan paradigm compromises the validity of any long-range comparison involving Tibeto-Burman proper, but this problem can easily be remedied, at least in principle. Meanwhile, Sino-Tibetan continues to shape the reconstructions and the identity of correspondences and so compromise the evidence adduced for Sino-Austronesian and Sino-Caucasian. This affects both theories of distant relationship, but the problem is compounded in the case of Sino-Caucasian by the reliance on lexicostatistics.

The nodes in Starostin's genealogical tree of languages are dated by glottochronology as determined by lexicostatistics, based on the assumption of a fixed rate of change in core vocabulary over time, whereby lexical divergence is calculated by a neighbour-joining algorithm. Popular in Russia today, lexicostatistics was invented by Constantine Samuel Rafinesque (1831) in order to win a gold medal worth 1,000 francs in a competition held by the Société de Géographie in Paris¹⁷ to determine the origin of Asiatic negritos.

Yet for Tibeto-Burman linguistics the question as to whether Old Chinese was a pidgin or creole which arose when the linguistic ancestors of the Chinese first came to the Yellow River Valley at the dawn of the Shāng period will continue to haunt us. Whatever the prehistory of Sinitic may be, no shared feature has yet been shown to unite the rest of Tibeto-Burman as opposed to Sinitic. Moreover, lexicostatistical studies that once were meant to show Sinitic to be the first branch to split off characteristically ignored most branches of Tibeto-Burman shown in Figure 9.2. By contrast, Jaxontov's 1996 Tibeto-Burman phylogeny based on lexicostatistics, reproduced by van Driem (2003: 112–113), resembles Shafer's family tree in that Sinitic is just one of several branches of the language family. There is no bifurcation of the family into Sinitic and some truncated 'Tibeto-Burman' construct.

At the same time, Starostin stressed the importance of the hierarchical principle, which he attributed to Vladislav Markovič Illič-Svityč, who, in reconstructing Nostratic, compared entities taken to have existed at the same time depth. Illič-Svityč compared Proto-Altaic with Proto-Uralic, for example, and did not draw comparanda from disparate levels, such as an ancient tongue and a modern language. Yet the presumption of an unsupported and probably false hierarchy is the hallmark of the 'Sino-Tibetan' model. Reconstructions within this paradigm accord as much weight to reconstructed Old Chinese as to all other language data from the entire language family. Furthermore, Peiros and Starostin's 'Sino-Tibetan' reconstruction violates the hierarchical principle in basing itself entirely

on the comparison of Old Chinese, written Tibetan, written Burmese and modern Jinghpaw and Lushai. By the same token, if Sagart's new Austronesian phylogeny is correct, comparisons between 'Austro-Thai' and Austroasiatic violate the hierarchical principle as well. At the same time, Starostin's reconstructed Austroasiatic comparanda are not taken seriously by leading specialists in Austroasiatic and do not respect the accepted hierarchy of Austroasiatic phylogeny (cf. Diffloth 2005).

In this context, it is relevant to keep in mind that Old Chinese is not the 'oldest language' in the family. Old Chinese is not an entity comparable to, say, Latin, Greek and other extinct languages written in an alphabetic script. Old Chinese was written in an ideogrammatic script, in which symbols represented words and morphemes. Because of the antiquity of the written tradition, however, Old Chinese is also something more than just a reconstruction analogous to Proto-Romance. Scholars who conduct the useful exercise of reconstructing Proto-Romance on the basis of the attested modern tongues arrive at a system reminiscent of Latin, but the resultant construct is not Latin by any stretch of the imagination and lacks much of the morphology which is known to have characterized the common ancestral tongue (Mazzola 1976; Hall 1984). On the basis of Proto-Romance it would be difficult even to ascertain whether Latin was closer to Faliscan or to Oscan and Umbrian. Epistemologically, Old Chinese is not as much as a Tibeto-Burman analogue of Latin, nor is Old Chinese as little as a Tibeto-Burman analogue of Proto-Romance.

Old Chinese is a linguistic edifice founded upon reconstructed Middle Chinese and built with the rimes of the *Shī Jīng* 'Book of Odes', dating from between the 8th and 5th centuries BC, and the phonetic components in Chinese characters that were devised in the Shāng and Zhōu period, buttressed by refined philological arguments. Much phonological information on Old Chinese was lost, albeit not all of it irretrievably, when the script was unified during the Qín dynasty in the 3rd century BC. Much has yet to be learnt from original specimens of writing antedating this period.

Middle Chinese, the foundation upon which Old Chinese is built, is reconstructed on the basis of the comparison of modern Sinitic languages, traditionally known as 'Chinese dialects', Chinese loanwords which entered Vietnamese, Korean and Japanese, and the *Qièyùn*, a Táng dynasty dictionary published in 601 containing *fānqiè* spellings that specify the pronunciation of a character by two other ideograms, one representing the *zìmǔ* 'initial' and the other specifying the *yùnmǔ* 'rime'.

Coblin (2003) has soberingly reviewed the epistemological underpinnings of reconstructing older stages of Sinitic. Old Chinese is not the language spoken by the ancient Chinese, but a reconstructible syllabary. Yet the language spoken at the time was no doubt more than just a syllabary, as Lepsius mooted in 1861. Whichever recently reconstructed syllabary one prefers, Old Chinese now looks like a reconstruction of a Tibeto-Burman language and gives the lie to the Sino-Tibetan hypothesis.

Starostin's comparisons assume etymological identity, and he excludes look-alikes such as Sino-Tibetan *miǝŋ 'name' and Proto-Indo-European *(e)nomen- 'name', between which no system of correspondence obtains despite phonetic similarity. Yet the sound laws which unite 'Sino-Tibetan' and Sino-Caucasian as well as entities such as Dene-Daic are not made explicit. How are we then to know that the comparanda adduced in Sino-Caucasian comparisons are real, much less that the correct cognates have been identified in the purportedly related language families? How much of this construction is science, and how much of it is arcane? Much can be improved by making the sound laws and presumed regularities explicit, testable or open to scrutiny.

Long rangers often see scholars working in individual recognized language families as conservative and as hoarding their data. Yet scholars with greater and more detailed knowledge of individual languages and language groups are particular about getting the data correctly analysed and accurately represented. So the perceived difference in subcultures is more than just a sociological phenomenon but a question of methodological rigour. Taking the language family as a whole more seriously would inevitably lead to the removal of the 'Sino-Tibetan' bias and result in more credible reconstructions. In summary, the evidence for Sino-Caucasian appears tenuous, especially due to the shaky nature of some of the reconstructed 'Sino-Tibetan' comparanda. At the same time, it is significant, though not strictly a linguistic issue, that the Sino-Caucasian theory makes little sense of the archaeology or of the findings of population genetics to date.

The overall size of the empirical base in support of either Sino-Austronesian or Sino-Caucasian is not overwhelmingly vast. None the less, for reasons explained above, Sagart's 75 comparisons look more compelling than Starostin's 1358. Even so, Sagart's comparison notably excludes personal pronouns and numerals, which do not compare well, a fact which Sagart thinks is explicable in terms of 'far-reaching paradigmatic changes (analogy, politeness shifts involving deictics)' (2005a: 165). Sceptics may therefore still dismiss the selection of purported cognates as representing look-alikes or borrowings. Indeed, Starostin is inclined to dismiss Sagart's Sino-Austronesian correspondences as loans or to attribute them to a new Dene-Daic or Sino-Austrian node at an even greater time depth. My first and present inclination has been to attribute Sagart's data to an ancient contact situation which I have already described above. If in future the evidence involving shared morphology is borne out by more rigorous studies of Tibeto-Burman historical grammar, however, then a deep genetic relationship becomes more likely than an ancient contact situation.

Just as in the case of *indochinesisch*, after Schott in 1856 diffidently resigned himself to the fact that other scholars would continue using the term, so too today scholars who continue to use the term 'Sino-Tibetan' likewise continue to adhere to the theory of genetic relationship which the term designates. That is, they continue to speak of Tibeto-Burman in the pinioned rather than the proper sense, in contexts which presume the veracity of this catch-all subgroup as a genetic construct coordinate with Sinitic. Since there is no evidence for a unitary truncated

'Tibeto-Burman' subgroup coordinate with Sinitic, the term 'Sino-Tibetan' must be abandoned along with the phylogenetic model which it designates.

6. East Asian and future prospects

Finally, I shall turn to a theory which Stan Starosta proposed a year before he died in July 2002. The theory, called East Asian, proposes an ancient phylum encompassing Kra-Dai, Austronesian, Tibeto-Burman, Hmong-Mien and Austroasiatic. The ancient morphological processes shared by the families of this phylum are ostensibly an agentive prefix *<m->, a patient suffix *<-n->, an instrumental prefix <s-> and a perfective prefix *<n->. The East Asian word was disyllabic and exhibited a canonical structure CVCVC. The proto-homeland of the East Asian proto-language or Proto-East-Asian dialect continuum ('linkage') lay in the region laced by the Hàn, the Wèi and the central portion of the Yellow River in the period from 6500 to 6000 BC. Indeed, Starosta identified the Péilíngāng and Císhān neolithic with Proto-East-Asian.

Starosta envisaged the linguistic ancestors of the Austronesians as the first group to have split off of East Asian. This family spread to the coast and then down the eastern seaboard to establish the Hémùdù and Dàwènkǒu Neolithic cultures of 5000 BC, ultimately to cross over to Formosa. Much later, emerging from Formosa, one migration gave rise to the Malayo-Polynesian expansion to insular Southeast Asia, Oceania and parts of peninsular Southeast Asia, whereas another migration led back to the South China mainland, where it gave rise to Kra-Dai or Daic.

Back on the North China Plain, a second group split off and left the East Asian homeland to move south and settle along the Yangtze, where they shifted from millet to rice agriculture. These 'Yangtzeans' in turn later split up into the first Austroasiatic language communities, whom Starosta envisaged behind the Kūnmíng Neolithic of 4000 BC, and the Hmong-Mien, who later, according to Pulleyblank (1983), first burst into history in what is now Húběi and northern Húnán as the Chǔ polity (770–223 BC) which challenged the Eastern Zhōu.

Back in the central Yellow River basin, a third descendant group of East Asian remained. This third family was Tibeto-Burman. Starosta accepted the Sino-Bodic hypothesis and so rejected Sino-Tibetan.¹⁸ Tibeto-Burman in Starosta's conception split into Sino-Bodic, which he associated with the Yǎngsháo Neolithic of 5800 BC, and a branch which he called Himalayo-Burman, which he associated with the Dàdìwǎn Neolithic in Gǎnsù 6500 BC. Sino-Bodic split up into Sinitic and Bodic. Starosta appears to have relabelled Bodic 'Tangut-Bodish' because he mistakenly supposed Tangut to be more closely related to Bodish rather than to Qiāngic. Starosta's Himalayo-Burman split up into Qiāngic, Kāmarūpan and Southern Himalayo-Burman. Qiāngic is a recognized subgroup, which possibly includes Tangut. Southern Himalayo-Burman may presumably be taken to include groups such as Karen, Lolo-Burmese, Mizo-Kuki-Chin and perhaps Pyu. Kāmarūpan is a misleading 'hypothesis' introduced by Matisoff which groups together languages known not to constitute a genetic taxon (Burling 1999; van Driem 2001: 405–7).

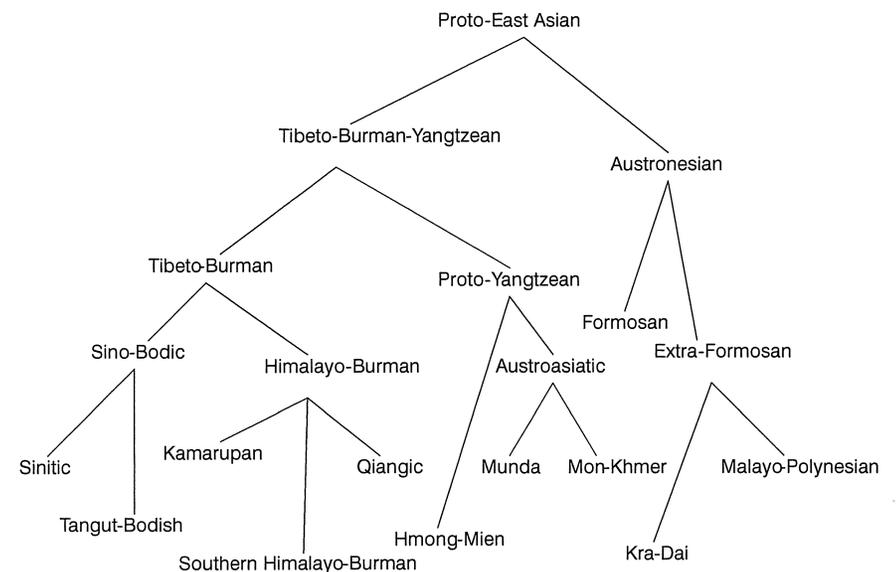


Figure 9.6 Starosta's Proto-East-Asian. This diagram faithfully represents Starosta's proposed East Asian phylogeny and corrects editorial errors which crept into his posthumously published tree diagram (2005: 183). The hypercorrect spelling 'Yangzi' has likewise been restored to the traditional English name 'Yangtze'.¹⁸

Starosta's theory basically proposes an agricultural dispersal of the type envisaged by Peter Bellwood and Colin Renfrew. The farming dispersal model is not problematic in straightforward cases such as the Polynesian colonization of hitherto uninhabited lands. However, this simplistic model is deficient for reconstructing linguistic intrusions and dispersals on continents, where population prehistory has been far more complex than the spread of agriculture reflected in the archaeological record. My qualified criticisms of the unqualified use of this hypothesis to argue the location of linguistic homelands can be consulted elsewhere (van Driem 2001: 423–6, 1004–21, 1051–65; esp. 2002: 238–9).

Although I do not currently subscribe to the East Asian theory any more than I do to Sino-Austronesian or Sino-Caucasian, I have attempted here to give both Sino-Austronesian or Sino-Caucasian a fair and sympathetic hearing. This final section sets the record straight about Starosta's intrinsically interesting hypothetical reconstruction of linguistic prehistory, particularly with regard to Tibeto-Burman and Sino-Bodic, since Starosta's theory somehow came out garbled in the posthumously published version. Starosta modestly concluded that the scenario which he sketched 'is almost certainly wrong in a number of points', but that 'its potential utility' lay 'in helping to focus scholars' efforts on particular specific questions, resulting in the replacement of parts of this hypothesis with better supported arguments' (2005: 194). It should come as no surprise if a good

number of Starosta's novel and insightful hunches were to be borne out by future research.

Notes

- 1 A polyphyletic view of linguistic stocks and language families arguably dates as far back as 1647. A fuller historical account of the rise of polyphyletic historical linguistic comparison has been provided elsewhere (van Driem 2001: 1039–51; 2005: 285–91).
- 2 In his quixotic attempts to reconcile the diversity which he observed with the monophyletic Turanian vision, Logan devised numerous *ad hoc* terms for real or imagined genetic ties between larger groups, e.g. 'Malagaso-Asonesian', 'Dravido-Asonesian', 'Tibeto-Ultraindian', 'Himalayo-Asonesian', 'Chino-Himalaic', 'Dravido-Australian', 'Ultra-Indo-Gangetic', 'Gangeto-Ultraindian'. None of these coinages was to be so enduring as Tibeto-Burman.
- 3 These first attempts at reconstruction inevitably suffered from major shortcomings and oversights and do not yet constitute reconstructions in the conventional historical linguistic sense, cf. Miller (1968, 1974), Sagart (2006).
- 4 Müller's writings on the topic are copious. I shall draw just one example from Klaproth on the distinction between ethnic and linguistic relationship: 'Es ist richtig zu sagen, die deutsche Sprache stammt von denselben Wurzeln ab als das *Sanskrit*, aber unsinnig darum das *Deutsche Volk* von den Hindu abzuleiten' (1823a: 43). Some scholars such as Huot agreed: 'L'opinion de M. Klaproth ne fait, selon nous, que confirmer notre opinion qui est celle de tous qui étudient la nature: que les langues ne peuvent que fournir des caractères incertains pour la classification des *espèces* ou des *racés* d'hommes' (Malte-Brun 1832: 1. 521), but this essential distinction was to be lost on many people.
- 5 In Figure 9.2, the Ērsū cluster is another name for 'Southern Qiāngic', and may in fact consist of several subclusters. Qiāngic is 'Northern Qiāngic', which is currently supposed to include the rGyal-rongic group recognised by Jackson Sun (Sūn Tiānxīn) and Huáng Būfán. In fact, the precise phylo genetic relationships between the diverse rGyal-rong languages, Ērgōng, Qiāng, Mǐ-ñiag (Mù yà), Tangut, Ērsū, Lǔsū, Tosu (Duōsū), Nà mùyì, Shǐ xīng, Guì qióng, Choyo (Quèyù), Zhābà and Prinmi (Pǔmǐ) have yet to be demonstrated. Whether or not the Qiāngic group which features prominently in Chinese scholarly literature is a valid clade has yet to be convincingly demonstrated, and there is a lot of work left to be done in Sìchuān and Yúnnán provinces.
- 6 By contrast, Matisoff's 'view from the Sinosphere' does not correspond to the insights of Sinologists but represents his self-confessed predilection to envisage the proto-language as endowed with Benedict's two proto-tones and structurally similar to Lahu, a language for which he professes great fondness (2000: 367)
- 7 Benedict's unusual treatment of Karen between 1972 and 1976, based mainly just on word order typology, may have been influenced by the view propounded by Avery at New Haven, Connecticut, that 'the position of the Karen dialects of British Burma is not yet settled, since they present features of both the isolating and agglutinating languages' (1885: xviii).
- 8 Well into the 1970s, Sino-Tibetanists still classified Daic or Kra-Dai as part of the Sino-Daic branch of Sino-Tibetan, e.g. Milner and Henderson (1965). General linguists still often continue to present Sino-Tibetan as a family comprising 'le chinois, le thaï, le tibétain et le birman', e.g. Malherbe (2001: 35).
- 9 Shafer pointed out: 'Bodish is genetically closer to Chinese than it is to Burmese. To anyone not led by the exotic appearance of Chinese characters to regard the language as a thing apart, this conclusion should not come as a surprise in view of geography and history' (1955: 97). His later discussion of the divisions extended the observation to Bodic as a whole.
- 10 The exhilaratingly productive search for Sino-Bodic evidence in Kiranti languages was abruptly curtailed when the member of the Himalayan Languages Project with whom I had undertaken to pursue this work fell chronically ill.
- 11 My article explicitly stated that the latter set of roots is reflected outside of Bodic, particularly in Brahmaputran, and Matisoff acknowledged that I stated this to be so, yet in the same article he insinuates that the latter cognate set too was adduced as representing exclusive Sino-Bodic isoglosses.
- 12 Just like British scholars in the 19th century, Jaxontov proposed a homeland in Sìchuān (1977). Subsequently, so did I (van Driem 1998). In their archaeological discussion of the Sìchuān homeland hypotheses, Aldenderfer and Zhang 'agree with van Driem that Sìchuan is a likely source for a Neolithic package' which gave rise to cultures on the Yellow River (2004: 39). Yet Aldenderfer and Zhang (2004: 37) appear to think that I do not include the mKhar-ro site near Chab-mdo or any other Tibetan archaeological sites in my model. The Tibetan archaeological site mKhar-ro or mKhar-chu, which I discuss at length (van Driem 2001: 430–1), is sinicized in the Chinese archaeological literature with characters that are correctly romanized as Kǎrùò, and which Aldenderfer and Zhang incorrectly transcribe as 'Karou'. Sites should be named properly in accordance with archaeological convention. Their misunderstanding again provides the context for my assertion that: 'Numerous artificial problems in Tibetan toponymy and cartography currently result from the practice of listing only the sinified version of Tibetan place names in Hànyǔ Pīnyīn romanisation without providing the real place names' (loc.cit.). Incorrect Hànyǔ Pīnyīn transcriptions merely exacerbate the problem. Aldenderfer and Zhang identify mKhar-ro or Kǎrùò as a colonial exponent of the Mǎjǐyāo neolithic in Gānsu, but their cursory familiarity with the literature leads them to think that they are the first to do so. In fact, a good number of Chinese archaeologists (e.g. Xìzàng etc. 1979; Ān 1992) had already identified mKhar-ro or Kǎrùò as a colonial exponent of the Mǎjǐyāo neolithic, and my model followed this consensus. Aldenderfer and Zhang do not differentiate between language spread by demic diffusion and language intrusion by colonial migration, and they inexplicably attempt to interpret 'Karou' as the result of demic diffusion from Sìchuān. Purely on linguistic grounds, Peiros's lexicostatistical classification based on the highest diversity of primary taxa purportedly indicates 'a possible location of the homeland in the territories south of the Himalayas', whereas the location of Sinitic could be 'easily explained as the result of later migration' (1998: 217). In Dec. 2004, at the 10th Himalayan Languages Symposium in Thimphu, I presented other arguments for a possible Himalayan homeland for Tibeto-Burman.
- 13 No doubt the acronym STAN will lead some to speculate that Sagart adopted the new name to commemorate the late Stanley Starosta, just as some have speculated that I named Sino-Bodic after the late Nicholas Cleaveland Bodman, who was one of its proponents before me. In fact, I only spoke with Bodman once in Lund in 1987, and Bodic is Shafer's old term for a hypothetical superordinate branch within the language family. Both the terms 'Bodish' and 'Bodic' contain the Tibetan word *Bod* 'Tibet'.
- 14 Sergei Starostin sadly passed away in Moscow at the age of 52 on 30 September 2005, just after this article had first been submitted for publication, several months after he had been awarded an honorary doctorate at Leiden.
- 15 In total, 331 Sino-Caucasian reconstructions are based only on North Caucasian and Sino-Tibetan reconstructions, 197 Sino-Caucasian reconstructions on correspondences between North Caucasian, Sino-Tibetan and Yenisseeian reconstructions, 163 Sino-Caucasian reconstructions on North Caucasian, Sino-Tibetan and Burushaski correspondences, 134 Sino-Caucasian reconstructed roots on North Causasian and Yenisseeian correspondences, 110 Sino-Caucasian roots on North Causasian and Burushaski correspondences, 86 Sino-Caucasian roots on Sino-Tibetan and Yenissieian

- correspondences, 57 Sino-Caucasian roots on North Caucasian, Yenisseian and Burushaski correspondences, 44 Sino-Caucasian reconstructions on Sino-Tibetan and Burushaski correspondences, 26 Sino-Caucasian reconstructions on Sino-Tibetan, Yenisseian and Burushaski correspondences, and 9 Sino-Caucasian reconstructions on correspondences between Yenisseian and Burushaski reconstructions.
- 16 Based on forms in Kiranti languages the names of which are misspelt as 'Kaling' (*recte* Khaling) and 'Tulung' (*recte* Thulung).
- 17 For a fuller historical account of the origin of lexicostatistics and the original mathematical models employed, see the relevant section in my paper for the Linguistic Society of Nepal (van Driem 2005).
- 18 However, the term 'Sino-Tibetan' appears in the posthumously published version of Starosta's article. Likewise, the tree diagram which was drawn up for Starosta posthumously misrepresents his proposed East Asian phylogeny for Tibeto-Burman or 'Sino-Tibetan'. The corrected tree diagram is given here as Figure 9.6.
- 19 The Mandarin Chinese for the Yangtze is Cháng Jiāng. The English name Yangtze derives from an older designation of a branch of the river in the Yangtze delta in Jiāngsū province downstream from Yángzhōu. This former branch of the river was named after a strategic ford Yángzǐ, the site of which no longer lies on the present course of the Yangtze.
- ## References
- Aldenderfer, M. and Zhang, Y. (2004) 'The prehistory of the Tibetan Plateau to the seventh century A.D.: Perspectives from China and the West since 1950', *Journal of World Prehistory*, 18(1): 1–55.
- Ān Zhīmǐn (1992) 'Neolithic communities in eastern parts of Central Asia', in A.H. Dani and V.M. Masson (eds) *History of Civilizations of Central Asia*, vol. 1, *The Dawn of Civilization: Earliest Times to 700 BC*, pp. 153–89, Paris: United Nations Educational Scientific and Cultural Organization.
- Avery, J. (1885) 'The Tibeto-Burman group of languages', *Transactions of the American Philological Association*, 16, appendix, Morning Session, Wednesday, July 8, 1885, pp. xvii–xix, New Haven, CT.
- Baxter III, W.H. (1992) *A Handbook of Old Chinese Phonology*, Berlin: Mouton de Gruyter.
- Baxter III, W.H. (1995) 'Ongoing Research: Old Chinese Version 1.1', paper presented at the Old Chinese Seminar held at Leiden University on 14 and 17 July 1995.
- Belyi, V.V. (1997) 'Rafinesque's linguistic activity', *Anthropological Linguistics*, 39(1): 60–73.
- Benedict, P.K. (1942) 'Thai, Kadai and Indonesian: A new alignment in southeastern Asia', *American Anthropologist*, 44: 576–601.
- Benedict, P.K. (1972) *Sino-Tibetan: A Conspectus*, Cambridge: Cambridge University Press.
- Benedict, P.K. (1976) 'Sino-Tibetan: another look', *Journal of the American Oriental Society*, 96 (2): 167–97.
- Bengtson, J.D. (1991) 'Notes on Sino-Caucasian ... Some Macro-Caucasian etymologies', in V. Sheveroshkin (ed.) *Dene-Sino-Caucasian*, pp. 67–172, Bochum: Universitätsverlag Dr N. Brockmeyer.
- Bleichsteiner, R. (1930) 'Die werschikisch-buruschkische Sprache im Pamirgebiet und ihre Stellung zu den Japhetensprachen des Kaukasus', *Wiener Beiträge zur Kulturgeschichte und Linguistik, Veröffentlichungen des Institutes für Völkerkunde an der Universität Wien*, 1: 289–331.
- Blench, R.M. and Spriggs, M. (1998) 'General introduction', in R.M. Blench and M. Spriggs (eds) *Archaeology and Language II: Correlating Archaeological and Linguistic Hypotheses*, pp. 1–19, London: Routledge.
- Blust, R.A. (1995) 'An Austronesianist looks at Sino-Austronesian', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 283–98, Berkeley, CA: University of California Press.
- Blust, R.A. (1996) 'Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology', in W.H. Goodenough (ed.) *Prehistoric Settlement of the Pacific*, pp. 117–140, Philadelphia, PA: American Philosophical Society.
- Bodman, N.C. (1980) 'Proto-Chinese and Sino-Tibetan: Data towards establishing the nature of the relationship', in F. van Coetsem and L.R. Waugh (eds) *Contributions to Historical Linguistics: Issues and Materials*, pp. 34–199, Leiden: E.J. Brill.
- Bouda, K. (1936) 'Jenisseisch-tibetische Wortgleichungen', *Zeitung der Deutschen Morgenländischen Gesellschaft*, 90: 149–59.
- Bouda, K. (1950) 'Die Sprache der Buruscho', *Eusko-Jakintza, Revue d'Études Basques*, 4: 37–50.
- Bouda, K. (1954) 'Burushaski Etymologien', *Orbis, Bulletin International de Documentation Linguistique*, 3(1): 228–30.
- Bouda, K. (1964) 'Burushaski Etymologien II', *Orbis, Bulletin International de Documentation Linguistique*, 13(2): 604–9.
- Burling, R. (1999) 'On "Kamarupan"', *Linguistics of the Tibeto-Burman Area*, 22(2): 169–71.
- Burling, R. (2004) *The Languages of the Modhpur Mandi (Garo)*, vol. 1, *Grammar*, New Delhi: Bibliophile South Asia.
- Coblin, W.S. (2003) 'The *Chiehyunn* and the current state of Chinese historical phonology', *Journal of the American Oriental Society*, 123(2): 377–83.
- Conrady, A. (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*, Leipzig: Otto Harrassowitz.
- Conrady, A. (1916) 'Eine merkwürdige Beziehung zwischen den austrischen und den indochinesischen Sprachen', in *Aufsätze zur Kultur- und Sprachgeschichte vornehmlich des Orients: Ernst Kuhn zum 70. Geburtstag am 7. Februar 1916 gewidmet von Freunden und Schülern*, pp. 475–504, Munich: Verlag von M. & H. Marcus.
- Conrady, A. (1922) 'Neue austrisch-indochinesische Parallelen', in *Asia Major: Hirth Anniversary Volume*, pp. 23–66, London: Robsthan and Company.
- Coupe, A.R. (2003) *A Phonetic and Phonological Description of Ao: A Tibeto-Burman Language of Nagaland, Northeast India*, Canberra: Pacific Linguistics, Australian National University.
- Court de Gébelin, A. (1774) *Monde Primitif, analysé et comparé avec le monde moderne, considéré dans l'Histoire Naturelle de la Parole; ou Grammaire universelle et comparative*, Paris: chez l'auteur, Boudet, Valleyre, Veuve Duchesne, Saugrain & Ruault.
- Cust, R.N. (1877) *Languages of the Indo-Chinese Peninsula and the Indian Archipelago*, London: Trübner for the Philological Society.
- Cust, R.N. (1878) *A Sketch of the Modern Languages of East India*, London: Trübner & Co.

- Diffloth, G. (2005) 'The contribution of linguistic palaeontology and Austroasiatic', in L. Sagart, R.M. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, pp. 77–80, London: Routledge Curzon.
- Driem, G. van (1997) 'Sino-Bodic', *Bulletin of the School of Oriental and African Studies*, 60(3): 455–88.
- Driem, G. van (1998) 'Neolithic correlates of ancient Tibeto-Burman migrations', in R.M. Blench and M. Spriggs (eds) *Archaeology and Language II*, pp. 67–102, London: Routledge.
- Driem, G. van (2001) *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, containing an Introduction to the Symbiotic Theory of Language*, 2 vols, Leiden: Brill.
- Driem, G. van (2002) 'Tibeto-Burman phylogeny and prehistory: Languages, material culture and genes', in P. Bellwood and C. Renfrew (eds) *Examining the Farming/Language Dispersal Hypothesis*, pp. 233–49, Cambridge: McDonald Institute for Archaeological Research.
- Driem, G. van (2003) 'Tibeto-Burman vs. Sino-Tibetan', in B. Bauer and G.-J. Pinault, (eds) *Language in Time and Space: A Festschrift for Werner Winter on the Occasion of his 80th Birthday*, pp. 101–19, Berlin: Mouton de Gruyter.
- Driem, G. van (2005) 'Sino-Austronesian vs. Sino-Caucasian, Sino-Bodic vs. Sino-Tibetan, and Tibeto-Burman as default theory', in Y.P. Yadava, G. Bhattarai, R.R. Lohani, B. Prasain and K. Parajuli (eds) *Contemporary Issues in Nepalese Linguistics*, pp. 285–338, Kathmandu: Linguistic Society of Nepal.
- Egerod, S. (1972) 'Les particularités de la grammaire chinoise', in J.M.C. Thomas and L. Bernot (eds) *Langues et techniques, nature et société: Études offertes en hommage à André Georges Haudricourt*, vol. 1, *Approche linguistique*, pp. 101–9, Paris: Éditions Klincksieck.
- Ferlus, M. (2003) 'On borrowing from Middle Chinese into Proto-Tibetan: A new look at the problem of the relationship between Chinese and Tibetan', in D. Bradley, R. LaPolla, B. Michailovsky and G. Thurgood (eds) *Language Variation: Papers on Variation and Change in the Sinosphere and in the Indosphere in Honour of James A. Matisoff*, pp. 263–75, Canberra: Pacific Linguistics, Australian National University.
- Forbes, C.J.F.S. (1878) 'On Tibeto-Burman languages', *Journal of the Royal Asiatic Society of Great Britain and Ireland*, NS 10: 210–27.
- Forchhammer, E. (1882) 'Indo-Chinese languages', *Indian Antiquary*, 11: 177–89.
- Forrer, Emil (1919) 'Die acht Sprachen der Boghazköi-Inschriften', *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, LIII: 1029–41.
- Gabelentz, H.G.C. von der (1881) *Chinesische Grammatik mit Ausschluss des niederen Stiles und der heutigen Umgangssprache*, Leipzig: T.O. Weigel.
- Genetti, C.E. (2003) *Tibeto-Burman Languages of Nepal: Manange and Sherpa*, Canberra: Pacific Linguistics, Australian National University.
- Grierson, G.A. (ed.) (1904) *Linguistic Survey of India*, vol. 2, *Mōn-Khmēr and Siamese-Chinese Families (Including Khassi and Tai)*, Calcutta: Superintendent of Government Printing, India.
- Grierson, G.A. (ed.) (1909) *Linguistic Survey of India*, vol. 3, part I, *Tibeto-Burman Family: Tibetan Dialects, the Himalayan Dialects and the North Assam Group*, Calcutta: Superintendent of Government Printing, India.
- Grube, W. (1881) *Die sprachgeschichtliche Stellung des Chinesischen*, Leipzig: T.O. Weigel.
- Hall, R.A. Jr. (1984) *Proto-Romance Morphology*, Amsterdam: John Benjamins.
- Haller, F. (2004) *Dialekt und Erzählungen von Themchen* (Beiträge zur tibetischen Erzählforschung, 14), Bonn: Vereinigung für Geschichtswissenschaft Hochasiens Wissenschaftsverlag.
- Hari, A.M. and Chhegu L. (2004) *Hyołmo-Nepālī-A□grejī Śabdakoś Yohłmo-Nepālī-English Dictionary*, Kathmandu: Central Department of Linguistics, Tribhuvan University.
- Haudricourt, A.-G. (1956) 'De la restitution des initiales dans les langues monosyllabiques: le problème du thai commun', *Bulletin de la Société de Linguistique de Paris*, 52: 307–22.
- Hildebrandt, K.A. (2003) 'Manange Tone: Scenarios of Retention and Loss in Two Communities', doctoral dissertation: University of California at Santa Barbara.
- Houghton, B. (1896) 'Outlines of Tibeto-Burman linguistic palaeontology', *Journal of the Royal Asiatic Society*, 1896: 23–55.
- Jacques, G. (2004) 'Phonologie et morphologie du japhug (rGyalrong)', thèse en vue de l'obtention du doctorat de linguistique, Université Paris VII – Denis Diderot, 9 Sept.
- Jaxontov, S.E. (1965) *Drevnekitajskij jazyk*, Moscow: Jazyki narodov Azii i Afriki.
- Jaxontov, S.E. (1977) 'Jazyki vostočnoj i jugovostočnoj Azii v pervom tysjačileti do našej èry', in N.N. Čebokсарov, M.V. Krjukov and M.V. Sofronov (eds) *Rannjaja ètničeskaja istorija i narody vostočnoj Azii*, Moscow: Nauka, pp. 98–108.
- Kate, Lambert Hermanszoon ten (1710) *Gemeenschap tussen de Gottische Spraeke en de Nederduytsche verতোont*, vol. 1, *By eenen Brief nopende deze Stoffe*; vol. 2, *By eene Lyste der Gottische Woorden, gelyklydyg met de onze, getrokken uyt het Gothicum Evangelium*; vol. 3, *By de Voorbeelden der Gottische Declination en Conjugation, nieulyks in haere Classes onderscheyden. Alles gericht tot Ophelderinge van den Ouden Grond van 't Belgisch*, Amsterdam: Jan Rieuwertsz.
- Kate, Lambert Hermanszoon ten (1723) *Aenleiding tot de Kennisse van het Verhevene Deel der Nederduytsche Sprake waer in Hare zekerste Grondslag, edelste Kragt, nuttelijkste Onderscheiding, en geregeldste Afeiding overwogen en nagespoort, en tegen het Allervoornaemste der Verouderde en Nog-levende Taelverwanten, als 't Oude Moeso-Gothisch, Frank-Duitsch, en Angel-Saxisch, benefens het Hedendaegsche Hoog-Duitsch en Yslandsch, vergeleken word*, 2 vols, Amsterdam: Rudolph & Gerard Wetstein.
- Kern, H. (1889) 'Het stamland der Maleisch-Polynesische volken', *Tijdschrift voor Nederlandsch-Indië*, 18(2): 1–9.
- Klaproth, J.H. von (1823a) *Asia Polyglotta*, Paris: A. Schubart.
- Klaproth, J.H. von (1823b) *Asia Polyglotta: Sprachatlas*, Paris: A. Schubart.
- Klaproth, J.H. von (1826) *Mémoires relatifs à l'Asie, contenant des recherches historiques, géographiques et philologiques*, 2 vols, Paris: Société Asiatique de Paris.
- Konow, Sten: See Grierson (1904, 1909).
- Kortlandt, F.H.H. (2003) 'Bad theory, wrong conclusions: M. Halle on Slavic accentuation', in J. Schaecken, P. Houtzagers and J. Kalsbeek (eds) *Dutch Contributions to the Thirteenth Annual Congress of Slavists, Ljubljana, August 15–21, 2003: Linguistics* (Studies in Slavic and General Linguistics, 30), pp. 237–40, Amsterdam and New York: Rodopi.
- Kuhn, E. (1883) *Ueber Herkunft und Sprache der transgangetischen Völker: Festrede zur Vorfeier des Allerhöchsten Geburts- und Namensfestes Seiner Majestät des Königs Ludwig II., gehalten in der öffentlichen Sitzung der Königlichen Akademie*

- der Wissenschaften zu München am 25. Juli 1881, Munich: Verlag der Königlichen Bayerischen Akademie.
- Kuhn, E. (1889) 'Beiträge zur Sprachenkunde Hinterindiens (Sitzung vom 2. März 1889)', *Sitzungsberichte der Königlichen Bayerischen Akademie der Wissenschaften* (Munich), Philosophisch-philologische Classe, 2: 189–236.
- Kuiper, F.B.J. (1948) 'Munda and Indonesian', in *Orientalia Neerlandica: A Volume of Oriental Studies, published under Auspices of the Netherlands' Oriental Society (Oostersch Genootschap in Nederland) on the Occasion of the Twenty-Fifth Anniversary of Its Foundation, May 8th, 1945*, pp. 372–401, Leiden: A.W. Sijthoffs Uitgeversmaatschappij.
- Lahaussais, A. (2002) 'Aspects of the Grammar of Thulung Rai, an Endangered Himalayan Language', doctoral dissertation, University of California at Berkeley.
- Laufer, B. (1916) 'The Si-hia language, a study in Indo-Chinese philology', *T'oung Pao*, 17: 1–126.
- Lepsius, C.R. (1861) 'Über die Umschrift und Lautverhältnisse einiger hinterasiatischer Sprachen, namentlich der Chinesischen und der Tibetischen', *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin* (for 1860): 449–96.
- Leyden, J.C. (1806) 'Plan for the Investigation of the Language, Literature, History and Antiquities of the Indo-Chinese Nations', 69-page manuscript held by the British Library, ADD. MSS 26,564; later published with changes as Leyden (1808).
- Leyden, John Casper (1808) 'On the languages and literature of the Indo-Chinese nations', *Asiatic Researches*, 10: 158–289.
- Li Fang Kuei (= Lǐ Fāngguì) (1971) 'Shàngǔ yīn yánjiū', *Tsing Hua Journal of Chinese Studies*, 9: 1–61.
- Li Fang Kuei (1974) 'Studies on Archaic Chinese' (trans. G.L. Mattos), *Monumenta Serica*, 31 (1974–5): 219–87.
- Li Fang Kuei (1976) 'Jǐge shàngǔ shēngmǔ wèntí', in *Zǒngtǒng Jiǎng Gōng Shìshì Zhōunián Lùnwén Jì*, pp. 1143–50, Taipei: Academia Sinica.
- Li Fang Kuei (1983) 'Archaic Chinese', in David N. Keightley (ed.) *The Origins of Chinese Civilization*, pp. 393–408, Berkeley, CA: University of California Press.
- Li, P.J. (1995) 'Is Chinese genetically related to Austronesian?', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 93–112, Berkeley, CA: University of California Press.
- Logan, J.R. (1852, 1853, 1854, 1855, 1857) 'Ethnology of the Indo-Pacific islands', *Journal of the Indian Archipelago and Eastern Asia* (Singapore), 6: 57–82, 653–88; 7: 20–63, 105–37, 186–224, 301–24; 8: 28–79, 200–61, 421–504; 9: 1–52, 162–272, 359–441, NS 1(2): 1–150.
- Logan, J.R. (1856) 'The Maruwi of the Baniak Islands', *Journal of the Indian Archipelago and Eastern Asia* (Singapore), NS 1(1): 1–42.
- Logan, J.R. (1858) 'The West-Himalaic or Tibetan tribes of Asam, Burma and Pegu', *Journal of the Indian Archipelago and Eastern Asia* (Singapore), NS 2(1): 68–114.
- Logan, J.R. (1859) 'Ethnology of the Indo-Pacific islands: The affiliation of the Tibeto-Burman, Mon-Anam, Papuanesian and Malayo-Polynesian pronouns and definitives, as varieties of the ancient Himalayo-Polynesian system; and the relation of that system to the Dravido-Australian', *Journal of the Indian Archipelago and Eastern Asia* (Singapore), 3(1): 65–98.
- Malherbe, M. (2001) 'La planète des langues', *Science et Avenir*, special issue, 'Quelle langue parlait-on il y a 100 000 ans? La langue d'Homo erectus' (Dec. 2000–Jan. 2001), 28–35.

- Malte-Brun, C. (1832) *Précis de la géographie universelle, ou description de toutes les parties du monde sur un plan nouveau, d'après les grandes divisions naturelles du globe: Nouvelle édition, revue, corrigée, mise dans un nouvel ordre et augmentée de toutes les nouvelles découvertes par J.J.N. Huot*, 6 vols, Brussels: Th. Lejeune (posthumous).
- Mason, F. (1854) 'The Talaing language', *Journal of the American Oriental Society*, 4: 277–89.
- Mason, F. (1860) *Burmah, its people and Natural Productions, or Notes on the Nations, Fauna, Flora and Minerals of Tenasserim, Pegu and Burmah*, Rangoon: Thomas Stowe Ranney.
- Matisoff, J.A. (1978) *Variational Semantics in Tibeto-Burman: The 'Organic' Approach to Linguistic Comparison*, (Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics, 6), Philadelphia: Institute for the Study of Human Issues.
- Matisoff, J.A. (1992) 'A key etymology', *Linguistics of the Tibeto-Burman Area*, 15(1): 139–43.
- Matisoff, J.A. (2000) 'On "Sino-Bodic" and other symptoms of neosubgroupitis', *Bulletin of the School of Oriental and African Studies*, 63(3): 356–69.
- Matisoff, J.A. (2003) *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibeto-Burman Reconstruction*, Berkeley, CA: University of California Press.
- Mazzola, M.L. (1976) *Proto-Romance and Sicilian*, Amsterdam: John Benjamins.
- Miller, R.A. (1968) 'Review of Robert Shafer, *Introduction to Sino-Tibetan*, Parts I and II', *Monumenta Serica*, 27: 398–435.
- Miller, R.A. (1974) 'Sino-Tibetan: Inspection of a Conspectus', *Journal of the American Oriental Society*, 94(2): 195–209.
- Milner, G.B. and E.J.A. Henderson (eds) (1965) *Indo-Pacific Linguistic Studies*, 2 vols, Amsterdam: North-Holland Publishing Co.
- Müller, F.M. (1855) *Languages of the Seat of War in the East, with a Survey of the Three Families of Language, Semitic, Arian, and Turanian* (frontispiece title: *Max Müller's Survey of Languages*, 2nd edn), London: Williams & Norgate.
- Nikolaev, S.L. (1991) 'Sino-Caucasian languages in America', in V. Sheveroshkin (ed.) *Dene-Sino-Caucasian*, pp. 42–66. Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Nikolaev, S.L. and Starostin, S.A. (1984) 'Severnokavkazskie jazyki i ix mesto sredi drugix jazykovyx semej Perednej Azii', in *Lingvističeskaja rekonstrukcija i drevnejsaja istorija vostoka*, vol. 3, *Jazykovaja situacija v Perednej Azii v X-IV tysjačeljetjix do našej èry*, pp. 26–34, Moscow: Nauka, Glavnaja Redakcija Vostočnoj Literatury.
- Nikolaev, S.L. and Starostin, S.A. (1994) *A North Caucasian Etymological Dictionary*, Moscow: Asterisk Publishers.
- Opgenort, J.R. (2004) *A Grammar of Wambule: Grammar, Lexicon, Texts and Cultural Survey of a Kiranti Tribe of Eastern Nepal*, Leiden: Brill.
- Opgenort, J.R. (2005) *A Grammar of Jero, with a Historical Comparative Study of the Kiranti Languages*, Leiden: Brill.
- Orël, V.È. and Starostin, S.A. (1990) 'Etruscan as an East Caucasian language', in V. Sheveroshkin (ed.) *Proto-Languages and Proto-Cultures: Materials from the First International Interdisciplinary Symposium on Language and Prehistory, Ann Arbor, Michigan, November 1988*, pp. 60–8, Bochum: Universitätsverlag Dr. N. Brockmeyer.
- Ostapirat, W. (2005) 'Kra-Dai and Austronesian: Notes on phonological correspondences and vocabulary distribution', in L. Sagart, R.M. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, pp. 107–31, London: RoutledgeCurzon.

- Pallas, P.S. (1786, 1789) *Linguarum Totius Orbis Vocabularia Comparativa Augustissimae cura collecta: Sravnitel'nye Slovari Vsex Jazykov" i Naręčij sobrannye Desniceju Vsevysočajšej Osoby*, 2 vols, St Petersburg: Johannes Carolus Schnoor.
- Peiros, I. (1998) *Comparative Linguistics in Southeast Asia*, Canberra: Pacific Linguistics.
- Peiros, I. and Starostin, S.A. (1996) *A Comparative Vocabulary of Five Sino-Tibetan Languages*, 6 fascicles, Parkville: University of Melbourne, Department of Linguistics and Applied Linguistics.
- Plaisier, H. (2005) *A Grammar of Lepcha*, Leiden University: doctoral dissertation.
- Pott, August Friedrich (1833, 1836) *Etymologische Forschungen auf dem Gebiete der Indo-Germanischen Sprachen, mit besonderem Bezug auf die Lautumwandlung im Sanskrit, Griechischen, Lateinischen, Littauischen und Gothischen*, 2 vols, Lemgo: Meyersche Hof-Buchhandlung.
- Przyluski, J. (1924) 'Le sino-tibétain', in A. Meillet and M. Cohen (eds) *Les Langues du Monde*, pp. 361–84, Paris: Librairie Ancienne Édouard Champion.
- Pulleyblank, E.G. (1983) 'The Chinese and their neighbours in prehistoric and early historic times', in D.N. Keightley (ed.) *The Origins of Chinese Civilization*, pp. 411–66, Berkeley, CA: University of California Press.
- Pulleyblank, E.G. (1984) *Middle Chinese: A Study in Historical Phonology*, Vancouver: University of British Columbia Press.
- Pulleyblank, E.G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver: University of British Columbia Press.
- Pulleyblank, E.G. (1995) 'Comments from Pulleyblank', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 325–35 Berkeley, CA: University of California Press.
- Rafinesque, Constantine Samuel (1831) 'Mémoires sur l'origine des nations nègres, ou Introduction à l'histoire des nègres indigènes d'Asie, d'Afrique Polynésie, Amérique et Europe par C.S. Rafinesque, professeur des Sciences historiques et naturelles à Philadelphie', manuscript kept at the American Philosophical Society in Philadelphia [cited by Belyi 1997].
- Reid, L.A. (1994) 'Morphological evidence for Austric', *Oceanic Linguistics*, 33(2): 323–44.
- Reid, L.A. (1999) 'New linguistic evidence for the Austric hypothesis', in E. Zeitoun and P. Jen-kuei Li (eds) *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, pp. 1–30, Taipei: Academia Sinica.
- Reid, L.A. (2005) 'The current status of Austric: a review and evaluation of the lexical and morphosyntactic evidence', in L. Sagart, R. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*, pp. x–y, London: RoutledgeCurzon.
- Ruhlen, M. and Starostin, S.A. (1994) 'Proto-Yeniseian reconstructions, with extra-Yeniseian comparisons', in M. Ruhlen (ed.) *On the Origins of Languages: Studies in Linguistic Taxonomy*, pp. 70–92, Stanford, CA: Stanford University Press.
- Sagart, L. (1990) 'Chinese and Austronesian are genetically related', paper presented at the 23rd International Conference on Sino-Tibetan Languages and Linguistics, 5–7 Oct., University of Texas at Arlington.
- Sagart, L. (1991) 'Chinese tones from Austronesian final consonants', in M. Ratliff and E. Schiller (eds) *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society*, pp. 367–79, Tempe, AZ: Arizona State University.
- Sagart, L. (1993) 'Chinese and Austronesian: Evidence for a genetic relationship', *Journal of Chinese Linguistics*, 21(1): 1–62.
- Sagart, L. (1994a) 'Proto-Austronesian and Old Chinese evidence for Sino-Austronesian', *Oceanic Linguistics*, 33(2): 271–308.
- Sagart, L. (1994b) 'Discussion note: A reply to J. A. Matisoff's 'a key etymology'', *Linguistics of the Tibeto-Burman Area*, 17(1): 167–8 [published 1996].
- Sagart, L. (1995a) 'Some remarks on the ancestry of Chinese', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 195–223, Berkeley, CA: University of California Press.
- Sagart, L. (1995b) 'Comments from Sagart', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 337–7, Berkeley, CA: University of California Press.
- Sagart, L. (1995c) 'Questions of method in Chinese-Tibeto-Burman comparison', *Cahiers de Linguistique, Asie Orientale*, 24(2): 245–55.
- Sagart, L. (1999) *The Roots of Old Chinese*, Amsterdam: John Benjamins.
- Sagart, L. (2002) 'Sino-Tibeto-Austronesian: an updated and improved argument', Paper presented at the 9th International Conference on Austronesian Linguistics at Australian National University, Canberra, 11 January 2002.
- Sagart, L. (2003a) 'The vocabulary of cereal cultivation and the phylogeny of East Asian languages', *Bulletin of the Indo-Pacific Prehistory Association*, 23 (Taipei Papers, vol. 1): 127–36.
- Sagart, L. (2003b) 'Sources of Middle Chinese manner types: Old Chinese prenasalized initials in Hmong-Mien and Sino-Tibetan perspective', *Language and Linguistics*, 4(4): 757–68.
- Sagart, L. (2004) 'The higher phylogeny of Austronesian and the position of Tai-Kadai', *Oceanic Linguistics*, 43(2): 411–44.
- Sagart, L. (2005a) 'Sino-Tibetan-Austronesian: an updated and improved argument', in L. Sagart, R.M. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, pp. 161–76, London: RoutledgeCurzon.
- Sagart, L. (2005b) 'Tai-Kadai as a subgroup of Austronesian', in L. Sagart, R.M. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, pp. 177–81, London: RoutledgeCurzon.
- Sagart, L. (2006) 'Review of Matisoff (2003)', *Diachronica*, 23(1): 206–23.
- Schmidt, W. (1906) 'Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentral-Asiens und Austronesiens', *Archiv für Anthropologie, Neue Folge*, V: 59–109.
- Schmidt, W. (1930) 'Die Beziehungen der austrischen Sprachen zum Japanischen', *Wiener Beiträge zur Kulturgeschichte und Linguistik, Veröffentlichungen des Institutes für Völkerkunde an der Universität Wien*, I: 239–51.
- Schott, Wilhelm (1856) 'Über die sogenannten indo-chinesischen Sprachen insonderheit das Siamische', *Abhandlungen der Königlichen Akademie der Wissenschaften zu Berlin aus dem Jahre 1856*, Philosophisch-historische Klasse: 161–79.
- Shafer, R. (1955) 'Classification of the Sino-Tibetan languages', *Word, Journal of the Linguistic Circle of New York*, 11: 94–111.
- Shafer, R. (1966, 1967, 1968, 1974) *Introduction to Sino-Tibetan*, parts I, II, III and IV, Wiesbaden: Otto Harrassowitz.
- Simon, W. (1927–9) 'Zur Rekonstruktion der altchinesischen Endkonsonanten', *Mitteilungen des Seminars für Orientalische Sprachen an der Friedrich-Wilhelms-Universität zu Berlin*, 30(1): 147–67, 175–204; 32(1): 157–228.

- Starosta, S. (2005) 'Proto-East-Asian and the origin and dispersal of languages of East and Southeast Asia and the Pacific', in L. Sagart, R.M. Blench and A. Sanchez-Mazas (eds) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, pp. 182–97, London: RoutledgeCurzon.
- Starostin, S.A. (1982) 'Praenisejskaja rekonstrukcija i vnešnje svjazi enisejskix jazykov: Opyt rekonstrukcii praenisejskoj zvukovoj sistemy', in E.A. Alekseenko, I.I. Goxman, V.V. Ivanov and V.N. Toporov (eds) *Ketskij sbornik: Antropologija, ètnografija, mifologija, lingvistika*, pp. 144–237, Leningrad: Nauka.
- Starostin, S.A. (1984) 'Gipoteza o genetièeskix svjazjax sinotibetskix jazykov s enisejskimi i severnokavkazskimi jazykami', in *Lingvistièeskaja rekonstrukcija i drevnejšaja istorija vostoka: Tezisy i doklady konferencii*, vol. 4, *Drevnejšaja jazykovaja situacija v vostočnoj Azii*, pp. 19–38, Moscow: Izdatel'stvo 'Nauka'.
- Starostin, S.A. (1989) *Rekonstrukcija drevnekitajskoj fonologièeskoj sistemy*. Moscow: Izdatel'stvo 'Nauka'.
- Starostin, S.A. (1991) 'On the hypothesis of a genetic connection between the Sino-Tibetan languages and the Yeniseian and North Caucasian languages', in V. Sheveroshkin (ed.) *Dene-Sino-Caucasian*, pp. 12–42, Bochum: Universitätsverlag Dr N. Brockmeyer.
- Starostin, S.A. (1994) 'The reconstruction of Proto-Kiranti', paper presented at the 27ème Congrès International sur les Langues et la Linguistique Sino-Tibétaines, Centre International d'Études Pédagogiques à Sèvres, 14 Oct.
- Starostin, S.A. (1995a) 'Old Chinese vocabulary: A historical perspective', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 225–51, Berkeley, CA: University of California Press.
- Starostin, S.A. (1995b) 'Comments from Starostin', in W.S.-Y. Wang (ed.) *The Ancestry of the Chinese Language* (Journal of Chinese Linguistics Monograph Series, 8), pp. 393–404, Berkeley, CA: University of California Press.
- Starostin, S.A. (1995c) 'Srvnitel'nyj slovar' enisejskix jazykov', in S.A. Starostin (ed.) *Ketskij sbornik: Lingvistika*, pp. 176–315, Moscow: Izdatel'skaja Firma 'Vostoènaja Literatura' Rossijskoj Akademii Nauk.
- Starostin, S.A. (1996) 'Word-final resonants in Sino-Caucasian', *Journal of Chinese Linguistics*, 24(2): 281–311.
- Starostin, S.A. (1998) 'Hurro-Caucasica', in T.M. Nikolaeva (ed.) *Politropon: k 70-letiju Vladimira Nikolaeviča Toporova*, pp. 90–9, Moscow: Izdatel'stvo 'Indrik'.
- Starostin, S.A. (1999a) 'Comparative-historical linguistics and lexicostatistics', in V. Shevoroshkin and P.J. Sidwell (eds) *Historical Linguistics and Lexicostatistics*, pp. 3–50, Melbourne: Association for the History of Language.
- Starostin, S.A. (1999b) 'Methodology of long-range comparison', in V. Shevoroshkin and P.J. Sidwell (eds) *Historical Linguistics and Lexicostatistics*, pp. 61–6, Melbourne: Association for the History of Language.
- Starostin, S.A. (2002) 'A response to Alexander Vovin's criticism of the Sino-Caucasian theory', *Journal of Chinese Linguistics*, 30(1): 142–53.
- Starostin, S.A. (2005) 'Current state of long-range linguistic taxonomy', paper presented on 9 June at Leiden University on the occasion of the conferral of an honorary doctorate upon him.
- Strahm, E. and Maibaum, A. (2005) *Jirel-Nepālī-A□grejī Šabdakoš Jirel-Nepālī-English Dictionary*, Kathmandu: Centre for Nepal and Asian Studies, Tribhuvan University.
- Tolsma, G. (1999) 'Internal reconstruction and comparative evidence of the long vowels in Kulung', in Y.P. Yadava and W.W. Glover (eds) *Topics in Nepalese Linguistics*, pp. 495–7, Kathmandu: Royal Nepal Academy.

- Trask, R.L. (1994) 'Basque: The search for relatives', *Dhumbadji!*, 2(1): 3–54.
- Trask, R.L. (1995a) 'Basque: The search for relatives (part 2)', *Dhumbadji!*, 2(2): 3–18.
- Trask, R.L. (1995b) 'Basque and Dene-Caucasian', *Mother Tongue*, 1: 3–82, 172–201.
- Trombetti, A. (1905) *L'unità d'origine del linguaggio*, Bologna: Libreria Treves di Luigi Beltrami.
- Trombetti, A. (1925) *Le origini della lingua basca* (Memoria presentata alla Reale Accademia delle Scienze dell'Istituto di Bologna nella Sessione del 24 Novembre (1923) Bologna: Cooperativa Tipografica Azzoguidi.
- Trubetzkoy (i.e. Trubeckoj), N.S. (1922) 'Les consonnes latérales des langues caucasiennes septentrionales', *Bulletin de la Société Linguistique de Paris*, 23(3): 184–204.
- Turin, M. (2005) *A Grammar of Thangmi*, Leiden University: doctoral dissertation.
- Wāng Fēng (2004) 'Language contact and language comparison: The case of Bái', doctoral dissertation, City University of Hong Kong.
- Watters, D.E. (2002) *A Grammar of Kham*, Cambridge: Cambridge University Press.
- Watters, D.E. (2004) *A Dictionary of Kham: Taka Dialect (A Tibeto-Burman Language of Nepal)*, Kirtipur: Central Department of Linguistics, Tribhuvan University.
- Witsen, N. (1692) *Noord en Oost Tartarye, ofte Bondig Ontwerp van eenige dier Landen en Volken, welke voormaels bekend zijn geweest, beneffens verscheide tot noch toe onbekende, en meest nooit voorheen beschreven Tartersche en Nabuurige Gewesten, Landstreeken, Steden, Rivieren, en Plaetzen, in de Noorder en Oostelykste Gedeelten van Asia en Europa*, 2 vols, Amsterdam: François Halma (2nd impr, 1705).
- Wulff, K. (1934) *Chinesisch und Tai: Sprachvergleichende Untersuchungen* (Det Kongelige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, 20/3), Copenhagen: Levin & Munksgaard.
- Wulff, K. (1942) [posthumous] *Über das Verhältnis des Malay-Polynesischen zum Indochinesischen*, Copenhagen: Munksgaard.
- Xizàng Zizhìqū Wénwù Guǎnlǐ Wěiyuán Huì [Managing Committee for Cultural Affairs of the Autonomous Region of Tibet] (1979) 'Xizàng Chāngdū Kǎruò yízhǐ shìjué jiǎn bào' (Brief report of an initial excavations of the site at mKhar-ro, Chab-mdo, Tibet), *Wénwù*, 1979(9): 22–8.
- Zhèngzhāng Shàngfāng (1987) 'Shàngǔ yùnmǔ xitǒng hé sǐdèng jièyīn shēngdiào de fāyuán wèntí', *Wēnzhōu Shīyùàn Xuébào (Shèhuì kēxué bǎn)*, 4: 67–90.