

THE SOCIOLINGUISTIC SITUATION OF
SIKKIM'S INDIGENOUS LANGUAGES AND
THE ORIGINS OF THE PEOPLES OF SIKKIM

GEORGE VAN DRIEM
University of Bern

THE TRANS-HIMALAYAN LANGUAGE FAMILY

One type of evidence about our prehistory comes from languages and linguistic relationships. This type of evidence involves identifiable language families and the branching patterns of language family trees. The indigenous languages of Sikkim all belong to the Trans-Himalayan language family, first identified by Julius von Klaproth (1823) in Paris as the language family comprising Tibetan, Burmese, Chinese, Garo and other 'trans-Gangetic languages', the latter being the Tibeto-Burman languages of Nepal, Sikkim, Bhutan and northeastern India.

In the course of the 19th and 20th centuries, a rival model arose under the name 'Indo-Chinese', later relabelled *sino-tibétain* by Jean Przyluski (1924). Indo-Chinese or 'Sino-Tibetan' encompassed numerous unrelated languages, and so had to be whittled down. This tree model has, moreover, always been defined by a phylogenetic presumption for which no historical linguistic evidence has ever been adduced. The neutral name Trans-Himalayan designates the language family as a whole, whatever model of phylogeny may ultimately emerge from the historical linguistic evidence in due course. The label 'Sino-Tibetan', on the other hand, historically came to denote a specific bifurcate language family tree model that grouped all non-Sinitic languages into a single taxon without supporting historical linguistic evidence (Orlandi 2021).¹

Today's state of the art in comparative linguistics supports von Klaproth's Trans-Himalayan model,² and Sino-Tibetanists, still unable to ad-

¹ In face of the vested interests of certain older scholars in the field and their hold on certain journal referees, Orlandi's moral courage and sense of rectitude despite his vulnerability as a young scholar must be greatly appreciated and respected in contributing such an important piece, for, to use Voltaire's words in reference to the literary career of Bernard le Bouvier de Fontenelle, in the case of this talented scholar it may likewise be observed that: 'Il vit combien il est dangereux d'avoir raison dans des choses où des hommes accrédités ont tort' (1803: 235).

² The history of the field is told in van Driem (2014, 2019).

duce historical linguistic evidence for their phylogenetic beliefs, have today sought recourse to lexicostatistics.³ As previously observed, our conclusion must therefore be: ‘There is no such language family as Sino-Tibetan’ (van Driem 2011: 30). As a consequence, von Klaproth’s original Trans-Himalayan language family model has today prevailed.⁴

As an alternative to any of the empirically unsupported family trees for the language family,⁵ I proposed the Fallen Leaves model (van Driem 2001b). The various trees hitherto proposed remain premature until the historical phonology and grammar of the language family have been adequately understood to justify them. In recent years, Nathan Hill has largely single-handedly effectuated a sorely needed ‘methodological reorientation of the study of Trans-Himalayan languages towards the paragon of Indo-European historical linguistics’ (2019: 257). The work of younger scholars has also begun to contribute towards changing the field (e.g. Gerber & Grollmann 2018, Bodt 2019, 2021, 2023, forthcoming). Fallen Leaves comprises recognised linguistic subgroups, some of which have

³ Bayesian models are mathematical tools termed ‘phylogenetic’ because they generate tree diagrams for any set of data, even if no tree structure obtains between these data. As a tool, Bayesian maths should not be confused with methodology. The methodology employed in the now infamous studies by Zhang *et al.* (2019) and Sagart *et al.* (2019) is lexicostatistics on the basis of precious little lexical material. At a conference on the ancestry of the languages and peoples of China, held at Jinán University in Canton (Guǎngzhōu) in May 2017, one member of the gang of four who went on to author the lexicostatistical study (Zhang *et al.* 2019) showed me, with what struck me as astonishing candour, how the strategic choice of the limited set of vocabulary items typically used in Bayesian lexicostatistical studies could greatly affect the outcome. In particular, he showed me that a certain selection of lexemes could skew the tree to put Kiranti at the top, since he knew that the result would titillate me, but by the same token another selection could tilt the tree back in another direction and so yield a very different outcome. Elsewhere (van Driem 2021, 2022), I have provided detailed critiques of the new ploy of resorting to lexicostatistics whenever Sino-Tibetanists are confronted with the absence of historical linguistic evidence in support of their tree model.

⁴ In sharp contrast to the name ‘Trans-Himalayan’ for the language family first discerned by Julius von Klaproth, Robbeets’ coinage ‘Trans-Eurasian’ for Altaic essentially designates Poppe’s original 1960 phylogenetic model of the Altaic language family, which Robbeets adopted half a century later unchanged, but has ever since striven to rebrand.

⁵ I drew one such tree myself (van Driem 1997), based on the archaeological record and linguistic impressionism (cf. van Driem 2005). Though my 1997 tree was supported by the archaeological consensus at the time, this interpretation of the archaeological record now requires revision (cf. van Driem 2021: 152–153). Like my 1997 tree, the similar tripartite trees proposed by Bradley (1997, 2002, 2012) and DeLancey (2021) are essentially based on the same linguistic impressionism as Shafer’s tentative divisions, discussed in the following section.

been validated and some of which have not, arranged in a heuristically useful diagram, which can be modified periodically as the historical linguistic study of Trans-Himalayan languages progresses and allows a tree structure to emerge (Figure 1). Each representation of the Fallen Leaves diagram is therefore liable to be updated.



FIGURE 1: The Fallen Leaves model, an agnostic heuristic framework showing the currently recognised Trans-Himalayan linguistic subgroups. As explained in the main text, the label ‘Bodish’ had been replaced with ‘Bodic’ (cf. van Driem 2021: 194), whilst West Himalayish and Tamangic are in future likely to vanish from the diagram, being subsumed within Bodic as ‘West Bodic’ and ‘South Bodic’ respectively (Bodt 2023).

A LEGACY OF LABELS AND NOMENCLATURE

Nathan Hill’s brand of Trans-Himalayan historical linguistics has effectively replaced the methodologically flawed ‘Sino-Tibetan’ approach and its faulty tree models. Yet the thinking of earlier scholars survives as a legacy which continues to mould today’s discourse. Robert Shafer once divided the language family into six divisions, viz. Sinitic, Bodic, Karenic, Burmic, Baric and Daic. In 1938, just months before Paul Benedict joined Alfred Kroeber’s ‘Sino-Tibetan Philology’ project at Berkeley,

Shafer had recognised that Daic was not a member of the language family. Yet his boss Alfred Kroeber and the Parisian sinologist Henri Maspéro compelled him to retain Daic as a sixth division (Shafer 1955: 97–98, van Driem 2001b: 344).⁶ As Shafer understood, and as von Klaproth (1823) had recognised long before, Daic represents a distinct language family, today called Kra-dai. Shafer's Sinitic and Karenic have survived as two out of many Trans-Himalayan subgroups identified in the Fallen Leaves model.

By contrast, Shafer's Bodic, Burmic and Baric each constituted a medley of disparate subgroups (1944, 1953, 1954, 1955, 1966, 1967, 1968, 1974). None of these three exploratory heuristic 'divisions' has yet been borne out by historical linguistic evidence, thus rendering the three terms essentially defunct labels. Yet some of Shafer's working hypotheses persist as labels still in use today, e.g. Bodish, Digarish, Midzuish, West Himalayish. Following the established nomenclatural tradition in Indo-European studies, Shafer proposed 'a logical system of nomenclature', so that 'the ending *-ic* denotes a main division of a family, as Sinitic...', and 'the ending *-ish* indicates a sub-group of one of the main divisions of the family' (1941: 58).

Hill (2019) broke with established nomenclatural tradition, for Hill's Burmish encompasses his Burmic. Hill's Burmish (2019: 51–52) is not equivalent to Shafer's Burmish 'section', but instead roughly equivalent to Shafer's 'Burma Branch' (1955: 103).⁷ Within his Burmish group, Hill has repurposed the label 'Burmish' to designate a lower-order subgroup comprising Burmese, Achang and Chintang (Xiāndǎo).

BODISH AND BODIC

In Shafer's classification, Bodish was one of the subgroups of Bodic. Both coinages are based on the Tibetan word for Tibet, *Bod*. In Shafer's terminology 'Bodish proper' was part of 'Bodish', which in turn was part of 'Bodic'. The way in which Shafer used the terms changed over time, and the use of the terms has continued to morph ever since. For example, today rGyalrongic, Tamangic, Tshangla are regarded as independent subgroups, distinct from 'Bodish'.

⁶ In addition to the occasional autobiographical reference, sometimes buried in a footnote, Shafer (1963) once published a few pages about himself in Louvain.

⁷ Shafer's Burmish 'section' comprised Lolo-Burmese and rGyalrongic, whereas his Burmic 'division' encompassed Burmish and a medley of subgroups which Shafer presumed bore close genetic affinity to Burmish, e.g. Nungish, Kachinic, Mru, Luish, Kuki-Chin.

Within Bodish, Shafer distinguished West, Central, South and East Bodish. Shafer had no information about the languages of Bhutan, and his understanding of the languages and the geography of Tawang was deficient. As a consequence, Shafer's 'East Bodish' formed part of his 'Bodish proper', within which he subsumed all languages which he believed to have derived from 'Old Bodish', of which Classical Tibetan is the literary exponent. Shafer's (1954) error with respect to 'East Bodish' was recognised by Michael Aris (1979a, 1979b).

In sequel to Aris' elucidations, Shafer's 'East Bodish' was reinterpreted by van Driem (1998, 2001b), Bielmeier (2004, 2018) and by Hill (2010) as representing a group of related languages that were descendant not from Old Bodish, but from a sister language of Old Bodish. The resultant heuristic model conceived of Shafer's West, Central and South Bodish as constituting 'Bodish proper' or just plain 'Bodish'. The redefined Bodish consisted of Dzongkha, Drenjongke and Tibetan, including all 'Tibetan dialects' or 'Tibetic languages' as far to the west as Baltistan in Pakistan-occupied Kashmir and as far to the northeast as Amdo in Chinese-occupied Tibet.

Long before Shafer, the Bodish languages were traditionally referred to in Tibetology as 'Tibetan dialects' or, in German, as *tibetische Dialekte*. Bielmeier (2004, 2018) followed this well-established convention. Nicolas Tournadre (2014a, 2014b) introduced the term *langues tibétiques* or 'Tibetic languages' in respect of the great linguistic diversity between the often mutually unintelligible languages subsumed under this group. Gawne and Hill (2017) used the term 'Tibetan languages', which as I have pointed out (van Driem 2019), has for historical and political reasons been perceived as problematic by speakers of Dzongkha and by some scholars as well. Most recently, Hill (2019) has sensibly proposed the continued straightforward use of the term 'Bodish' for the languages which derive directly from 'Old Bodish' or Old Tibetan.

Closely related to Bodish or 'Bodish proper' are the languages that have come to be denominated as 'East Bodish', i.e. Mangde, Bumthang, Khengkha, Kurtöp, Dzala, Dakpa, Chali. There are several dialects of Mangde, a language which is consequently often identified by a loconym, such as the dialect of Phobjikha, for example. There are several additional 'East Bodish' or Bodic language communities in the 'Nyamnyang མཉམ་མཉམ་མཉམ་ *mÑam-smyañ* river valley,⁸ where the Chinese colonial forces infam-

⁸ In a map of the area prepared in 1913 by surveyor Henry Treise Morshead and captain Frederick Marshman Bailey, the 'Nyamnyang མཉམ་མཉམ་མཉམ་ *mÑam-smyañ* river is anglicised as 'Nyamjang', a spelling retained in some modern maps. Bodt recorded the local pro-

ously made illegal incursions from Chinese-occupied Tibet into Indian territory in 1962.

Whilst the river is named 'Nyamnyang, the valley is locally known as Pangchen སུང་ཆེན་ *sPañ-chen* 'large flatland'. Administratively speaking, Bodt reports that the area used to be called Pangchen Dingdr'û སུང་ཆེན་གླིང་ལྗོངས་ *sPañ-chen lDiñ-drug* 'six divisions of the large flatland' (Bodt 2014: 208),⁹ and informally referred to as Chemithang རྗེ་མའི་ཐང་ *Bye-maḥi-than* 'sand flat' [tɕʰemithaŋ], a toponym which now often appears in the misleadingly transmogrified anglicised spelling 'Zemithang'.

This previously unresearched and linguistically complex portion of Tawang and West Kameng in the west of Arunachal Pradesh, which Aris (1979b) called the 'Monyul corridor', was extensively investigated for the first time by Tim Bodt, who published a detailed and highly valuable study on the area (Bodt 2014). In the various settlements of the 'Nyamnyang valley and surrounding hill tracts, a number of distinct languages are spoken. In addition to *Dzala 'mat* 'Dzala language' spoken in Bhutan, *Dakpa ket* 'Dakpa language' is spoken as three distinct lects, viz. Dakpa in Tibet, the Dakpa of Dakpaneng (i.e. the Lumla region and southern Trashiyangtse) and the Dakpa spoken in the Tshosum region of Tawang. Dakpa is usually called *Monket* 'Mon language', and in fact members of the language community do not usually refer to themselves as 'Dakpa', but as 'Monpa'.

The Tibetan term མོན་ *Mon* or Chinese 蠻 *Mán* have historically been used as a cover term to designate various often entirely distinct ethnolinguistic groups whose settlements straddle the Himalayas or lie along the southern flanks of the Himalayas. Despite its pejorative connotation in Tibetan and Chinese, the term is used by the members of many of the language communities thus designated to denote themselves, and this also happens to be the case with the Dakpa language community (van Driem 2001b: 472–473, 914–918). In addition to Dzala and Dakpa, *Pangchenpa 'mat* 'tongue of the people of the large flatland' is spoken in the 'Nyamnyang river valley, with 'Lepo', being spoken across the

nunciation of the toponym as [ɲammjaŋ] and [ɲamɲaŋ]. In the Bhutanese Dakpa village of Khinyel, Bodt recorded the pronunciation [ɲamzaŋ], and it is upon this local Bhutanese Dakpa pronunciation that the spelling in the Morshead and Bailey map was evidently based. The name of the lateral tributary known as the 'Namkha river གན་མ་ལ་ཚུ་ *gNam-kha-chu* today appears in an unsatisfactory Roman spelling on most modern maps. Bodt took me through the 'Nyamnyang valley in October and November 2013.

⁹ This language community lies outside of Bhutan. Yet these names are rendered here in Roman Dzongkha because of the proximity to Bhutan and because Roman Dzongkha provides an unambiguous phonological rendering of an actual Bodish pronunciation.

border in Tibet, representing a distinct subvariety. Although this language forms part of the Dzala-Dakpa cluster, *Pangchenpa 'mat* shows incomplete mutual intelligibility with Tawang Dakpa or *Monket*.

Meanwhile, Nathan Hill has stressed that, whereas ‘Bodish proper’ is defined by numerous unique shared phonological innovations, these ‘Tibetan sound changes ... do not affect the East Bodish languages’ (Hill 2019: 21).¹⁰ To our present state of knowledge, East Bodish therefore does not, in fact, constitute a valid subgroup as defined by shared innovations which unite these languages as a coherent taxon. Rather, ‘East Bodish’ represents what is left when ‘Bodish proper’ has split off. Therefore, ‘East Bodish’ can no longer be taken to designate a linguistic subgroup, but merely labels a set of closely related languages that are not derived directly from Old Tibetan. At this moment, it is therefore defensible to say that there is no such linguistic subgroup as East Bodish. The historical linguistic evidence for this stance is provided and discussed in great detail by Tim Bodt’s study of ‘East Bodish’ or Bodic languages in this issue of the *Bulletin*, in which he presents the most important and well-informed contribution on the topic.

To alleviate the terminological ambiguity, I proposed repurposing Shafer’s defunct label ‘Bodic’ to designate the taxon comprising both Bodish and the medley of languages which has conventionally been called ‘East Bodish’ (van Driem 2019). In this way, we can refer to all these languages instead simply as Bodic, within which Bodish constitutes a subset defined by well-documented shared innovations. The Bodic languages outside of the Bodish subgroup have until now traditionally been called ‘East Bodish’, and the label may continue to serve as a term of convenience to designate the Bodic tongues other than the languages of the Bodish subgroup, explicitly without making the implicit claim that East Bodish languages constitute a coherent subgroup within Bodic. Our current state of knowledge is that the East Bodish languages together with ‘Bodish proper’ represent a polyphyletic set of language subgroups.

In the present issue of the *Bulletin*, Bodt (2023) adduces evidence for a redefined Bodic and discusses whether this redefined Bodic can be validated as a linguistic subgroup. Bodt examines and evaluates the evidence to assess whether or not East Bodish can be validated and established as a linguistic subgroup and comes to an assessment empirically based on historical comparative linguistic evidence. Therefore, for the sake of argument, the fallen leaf formerly labelled ‘Bodish’ has been relabelled

¹⁰ Jacques (2012) has contributed an internal reconstruction of Tibetan stem alternations.

‘Bodic’ in the Fallen Leaves diagram. Moreover, the repurposed use of the label ‘Bodic’, *sensu stricto*, as in this updated Fallen Leaves diagram presented here, does not preclude the future use of the term in a sense closer to Shafer’s original conception, i.e. *sensu lato* also encompassing West Himalayish and Tamangic (Shafer 1950, 1951, van Driem 2001b). Elaborate evidence for this hypothetical subgrouping is now presented for the first time by Bodt in this issue of the *Bulletin*.

Like Bodic *sensu stricto*, the infelicitously named ‘West Himalayish’ may represent a polyphyletic set or, in any event, a set of languages the internal phylogeny of which is not understood (Ramirez 2021). Within Shafer’s original hypothetical Bodic branch *sensu lato*, a geographically inspired renaming of constituent subgroups could be undertaken, whereby the unaesthetically named ‘West Himalayish’, which has traditionally been assumed to include the extinct Zhangzhung language (Shafer 1957, Haahr 1968, Takeuchi & Nishida 2009, Jacques 2009, Martin 2010, 2013, 2016) could be given with the more aesthetic label West Bodic.

Tamangic would remain Tamangic, which might encompass languages such as Kaike and Ghale (cf. van Driem 2011). Alternatively, however, Bodt in this issue of the *Bulletin* proposes the term ‘South Bodic’ for a Tamangic that has been repositioned phylogenetically within a redefined Bodic *sensu lato* within the Trans-Himalayan language family. Based on the historical comparative linguistic evidence which Bodt has mustered, he then repositions the Dzala-Dakpa cluster, which I identified as a coherent subgroup fifteen years ago (van Driem 2007), as a distinct taxon. In other words, Bodt has effectively dismantled ‘East Bodish’ and replaced this polyphyletic catchall with a redefined ‘East Bodic’ alongside a separate Dzala-Dakpa cluster. However, these two taxa have now been arranged within Bodt’s redefined Bodic *sensu lato* in three different phylogenetic configurations, reflecting his three hypotheses for which he has adduced and discussed the empirical underpinnings.

SOCIOLINGUISTIC SITUATION OF THE BODISH LANGUAGES

The languages which derive from ‘Old Bodish’ or ‘Old Tibetan’ spread from the Yarlŭng ཡར་ལུང་སྐད་སྡེ་ [jɑː.lũː]¹¹ *Yar-kluŋs* valley, whence the Tibetans

¹¹ The pronunciation of Tibetan toponyms in Central Tibet is provided in Roman Tibetan and also between square brackets in the notation of International Phonetic Association, based on the variety of Central Tibetan spoken in Shikátsé (Haller & Haller 2007). Roman Tibetan is a phonological transcription which phonemically represents spoken Central Tibetan in Roman script. This system of representing the living spoken Tibetan in Roman script forms the topic of a contribution to an upcoming issue of the *Bulletin*

began expanding in the 7th century AD (van Driem 2001b: 829–846). The tumuli housing the tombs of the early historical Yarlŭng kings are situated in the ancient horn province of Yóru གཡོ་རུ [jörü] *g.Yo-ru*, in the district of Chóngcê རྩོད་རྒྱལ་ [te^höŋteë] *hPhyoñ-rgyas*, where the tomb of king Sóngtsén Kampó སློང་བཙན་རྒྱལ་པོ་ [sōŋtsë kəmpō] *Sroñ-btsan sGam-po* is famously located in the grave field of Murá མུ་ར་ [murā] *Mu-ra* in the Yarlŭng valley (Tucci 1950, Richardson 1963). According to Tibetan tradition, the ten historical kings of the Yarlŭng dynasty were preceded by a lineage of 32 prehistoric Yarlŭng chieftains, and indeed the Tibetan tumulus tradition appears to have begun in the 4th century AD. Mapping of the burial mound sites from the Tibetan imperial period has shown the highest concentration of tumuli in the province of Úrú རུ་རུ་ [ürü] *dBu-ru*, especially along the Kyícú སྐྱིད་ཅུ་ [kjīteū] *Skyid-chu*¹² upstream from the Chokáng རྩོད་ཁང་ [te^hokā] *Jo-khañ*, where Sóngtsén Kämpo established his capital (Hazod 2007, 2013, Kriz & Hazod 2020).

Tumuli likewise abound in the horn provinces of Yóru གཡོ་རུ [jörü] *g.Yo-ru* and Yérú གཡས་རུ་ [jērū] *g.Yas-ru*, whereas the distribution of burial mounds is sparser in the horn province of Rulâ རུ་ལག་ [ruldā] *Ru-lag*. Yet the distribution of ancient tumuli extends into adjacent regions beyond the original four Tibetan ‘horn’ provinces (Kriz & Hazod 2020). Rulâ རུ་ལག་ *Ru-lag*, the youngest of the Tibetan horn provinces, was incorporated into the Tibetan empire in the 8th century AD (Uray 1960) in the aftermath of the conquest and assimilation of the Zhangzhung kingdom in the 7th century AD (van Driem 2001a). For a sense of geographical perspective, the ancestral Tibetan heartland surrounding the tumuli in the district of Chóngcê རྩོད་རྒྱལ་ [te^höŋteë] *hPhyoñ-rgyas* lies just 150 km north of the northeastern Bhutanese town of Trashiyangtse.

The zone of the original Rupshí རུ་བཞི་ *Ru-bzi* [rupɕi] ‘four horn provinces’, comprising Phò བོད་ [p^hö] *Bod* or historical Tibet proper before Tibetan imperial expansion, constitutes an elongated region along the Tsángpō གཙང་པོ་ [tsāŋpō] *gTsañ-po*. This large sliver of territory, consti-

of Tibetology. In contrast to the first experimental version of Roman Tibetan, in which the low register tone was marked by a grave accent (van Driem 2021: iv), the updated version of Roman Tibetan leaves vowels in low register tone orthographically unmarked, whilst the high register tone is marked by an acute accent, the high falling tone by a circumflex accent and the low falling tone by a grave accent. Sikkimese toponyms are rendered in Roman Drenjongke (Namgyal & van Driem 2022), and place names in the Chumbi Valley, Bhutan and the ‘Nyamnyang river valley are rendered in Roman Dzongkha (Tshering & van Driem 2019).

¹² also སྐྱི་ཅུ་ *sKyī-chu*, and ultimately derived from the clan name ལྷི་ *Kyi* (Sørensen & Hazod 2007: 17–27).

tuting the original Tibet, lay in its entirety to the south of the Heavenly Lake, Námtsó གནམ་མཚོ་ [nām̥tsō] *gNam-mtsho*, and was contiguous with the modern territories of Bhutan and Sikkim. In the west, the original Tibet or Phò བོད་ [pʰh̥ɔ̃] *Bod* extended from the part of the Tibetan plateau north of the Nepalese district of Rasuvā eastward to the portion of the Tibetan plateau that lies to the north of the district of Upper Siang in the Indian state of Arunachal Pradesh (Hazod 2009).

After the era of conquest and martial expansion, Tibet during the imperial period extended for over 2,600 km from Baltistan in the west to the town of Yá-ngá གཡལ་གར་ [jāŋā] *g.Yag-rña* on the eastern Tibetan frontier, currently known to the Chinese as Yǎ'ān 雅安 in the redrawn province of Sìchuān, which swallowed up a lot of Tibetan territory in the aftermath of the Chinese invasion and colonial occupation of Tibet in 1950 (van Walt van Praag 1987, van Walt van Praag & Boltjes 2020). From the 7th to the 9th century, martial conquest by Tibetan imperial forces likewise disseminated Tibetan language and culture over 1,350 km to the north and northeast of the Yarlûng valley. The languages which denominated by Shafer as 'South Bodish' appear to comprise a coherent cluster within Bodish and include Drenjongke in Sikkim, *J'umbi kha* in the ལུམོ་ J'umo or Thromó ལྷོམོ་ [tʃʰm̥o] *Gro-mo* valley,¹³ Dzongkha in western Bhutan and Cho-ca-nga-ca-kha in the Kurichu valley.¹⁴ The

¹³ Siiger reported that the Chumbi valley 'used to be Lepcha territory' (1967, I: 44), and a number of native Lepcha toponyms in the Lepcha ritual texts analysed and translated by Siiger & Rischel (1967, II) are located in or near the Chumbi valley. Today a village located at 27°28'04"N and 88°54'39"E goes by the name of ལུམི་ *Chu-hbi*, which the Chinese colonial occupiers have sinicised as Mandarin *Chūnpéi* 春培. Per Kjeld Sørensen has suggested to me that this toponym, if not a new settlement bearing a *post hoc* adaptation of the anglicised 'Chumbi' [< Dz. ལུམོ་ལའི་ *J'umbi*, the adjectival form of ལུམོ་ J'umo, the Dzongkha name for the Chumbi valley, known as ལུམོ་ Gy'umo in Drenjongke] (Namgyal & van Driem 2022: 22), could be connected to Lepcha ལུམོ་ *cubi* 'abode of high snows' (cf. Mainwaring 1898: 81, 255–256, Tamsang 1994: 306).

¹⁴ Shafer's 'South Bodish Unit' (1955: 101) actually contained 'Upper and Lower Gro-mo', and he listed both 'Sikkimese' and 'Dandzongka', but not Dzongkha. It appears that Shafer included a reference to Dzongkha as 'Dru', which, however, he listed under his 'Central Bodish Unit' (1955: 100). I added both Dzongkha and Cho-ca-nga-ca-kha to South Bodish (van Driem 1998, 2001b), and this enumeration is followed by Tournadre (2014b) and Bielmeyer *et al.* (2018: 44), although Tournadre incorrectly adds Dur Brokkat and Brokpa to his 'Southern section' (2014b: 122). A discussion of South Bodish languages other than Cho-ca-nga-ca-kha is provided by Namgyal & van Driem (2022). Early linguistic work on the development of Dzongkha as a written language is discussed by Nado (1982), van Driem (1998), Tshering & van Driem (2019).

areas where these languages are spoken lie at a distance of just 200 to 400 km from the historical Tibetan homeland.

An understanding of both the modern geopolitical situation of Tibet and adjacent regions where Bodic languages are spoken is immediately germane to the language endangerment position and future prospects of the modern Bodic languages. The great degree of dialectal diversity between the Bodish languages in Sikkim and Bhutan and across the vast expanse of the Tibetan plateau, stretching from the inland sea known as the Blue Lake or Tshó Ngónpó མཚོ་སྤོན་པོ་ [tʰɔ̃ ɲɔ̃pɔ̃]¹⁵ *mTsho sNon-po* in the northeast to Baltistan in Pakistan-occupied Kashmir in the west, reflects centuries of local language evolution since the Tibetan imperial expansion that began in the 7th century. The Bodish linguistic dispersal took place at the height of Tibetan power, at a time when China had been reduced to a Tibetan tributary state. In 763, when the Táng government once had the temerity to withhold tribute to the Tibetan court, the Tibetan imperial army captured the Chinese capital at Cháng'ān, and for a spell even installed the Chinese brother-in-law of king Thrisóng Tetsén ལྷི་སྤོང་ཧེ་སྤོང་ཧེ་བཅོན་ [tʂʰɔ̃sɔ̃ tɕsɛ̃] *Khri-sroñ lDe-btsan* (reignat 755–797) as the emperor of China.¹⁶

In 1244, Tibet came under Mongol suzerainty but remained an independent polity preserving its own legal system. Instead, the Mongol state entered into a *mchod-yon* མཚོན་ཡོན་ 'preceptor-patron' relationship with Tibet, in which the clerical leader of Tibet, who during this period was the *Sa-skya Pañdi-ta* ས་སྐྱུ་པཎྌི་ཏཾ, figured as the *mchod-gnas* མཚོན་གནས་ 'object of veneration', i.e. court chaplain or *purohita*, and the secular head of the Mongol state served as the *yon-bdag* ཡོན་བདག་ 'lay patron supporting the propitiation of the preceptor'. This notion of a preceptor-patron relationship would endure for centuries as a modality of international relations in the Tibetan Buddhist world order. Meanwhile, after suffering over four decades of military onslaughts, China definitively fell to the Mongols in 1279. Tibet threw off Mongol suzerainty in 1354, and China threw off the Mongol yoke in 1368.

¹⁵ The Oirat Mongol name *Köke Nur* 'blue lake', which widely appears in older Western atlases as 'Koko Nor', and the Chinese name *Qīnghǎi* 青海 'blue sea' are both modelled directly after the Tibetan.

¹⁶ Useful accounts of Tibetan history include Richardson (1962), Stein (1959, 1962), van Driem (2001b), Kapstein (2006), van Schaik (2011). In a related vein, Ardussi (1977), Aris (1979a), van Driem (2001b) contain relevant accounts of Bhutanese history and also cite numerous other valuable sources on Bhutanese history in their bibliographies.

China was ruled by the native Míng dynasty, but in 1644 China came under the rule of the Manchu Qīng dynasty. The Qīng government asserted suzerainty over Tibet in 1720, but Manchu hegemony was in practice a loose relationship between the two states. The Manchu government in Peking readopted the *mchod-yon* ‘preceptor-patron’ relationship with Tibet, that had previously been cultivated under the Mongols, but now with the successive Dalai Lamas functioning as *mchod-gnas* ‘objects of veneration’, and the Qīng emperors as their *yon-bdag* ‘lay propitiators’. Significantly, China declined to come to the aid of Tibet during the first Nepalese invasion of Tibet (1788–1789). In response to the second Nepalese invasion of Tibet (1791–1792), the Qīng government did acquiesce to coming to Tibet’s aid, but during the Nepal-Tibetan war (1855–1856) China again neglected to assist Tibet. Consequently, by provision of the peace treaty signed at Kathmandu in March 1856, Nepal officially replaced China as the protector of Tibet, and Nepal subsequently maintained a resident at Lhasa and enjoyed duty-free trade with Tibet.

The Chinese revolution led to the overthrow of the Manchu dynasty in February 1912 and the establishment of the Republic of China. On the 12th of March 1912, the government of sovereign Tibet renounced all ties with imperial China. The complexity of the jostling for control over territory in the Sino-Tibetan borderlands and the power play which unfolded between the Tibetan and Chinese governments at this time is well illustrated by the study of Jagou (2019). After the Second World War, Chinese communist insurgents exploited their country’s weakness to overrun the mainland in 1949, at which time the legitimate government of China fled to Taiwan. A communist dictatorship seized control in Peking, and the new polity which designated itself the People’s Republic of China emerged as the world’s newest and most aggressive colonial power, invading East Turkestan on the 12th of October 1949 and invading sovereign Tibet on the 6th of October 1950. Increasing Chinese interference and belligerence ultimately forced the Dalai Lama, who served both as the head of state and the spiritual leader of Tibet, to flee in March 1959. Since 1959, the legitimate government of Tibet has been headquartered at Dharamsala in the Indian state of Himachal Pradesh, where the Tibetan government in exile maintains its temporary seat to the present day.

A good understanding of modern Tibetan history and recent developments in Sino-Tibetan relations is indispensable to an accurate portrayal of the language endangerment situation of Tibetan and future prospects for the survival of Drenjongke and Dzongkha. After the Dalai Lama was forced into exile, Chinese troops of the ‘People’s Liberation Army’ in

occupied Tibet slaughtered over 87,000 Tibetans, and over 1.2 million Tibetans perished in the years of brutal oppression which followed. Operating from within occupied Tibet, Chinese colonial forces crossed the Indo-Tibetan frontier and invaded India on the 20th of October 1962. In 1965, the Chinese colonial government redrew the so-called provincial boundaries in order to obliterate the international border between Tibet and China in a bid to erase the Tibetan nation from collective memory. The imperialist venture in Tibet orchestrated by the Chinese Communist Party has also unleashed the demographic weapon which McGranahan (2019) calls ‘Chinese settler colonialism’.

For seven months, starting from the 2nd of March 1969, the People’s Republic of China waged war against the Soviet Union along the border between Chinese-occupied Manchuria and the Soviet Far East. On the 19th of January 1974, China invaded the Paracel Islands, which had for centuries formed part of Vietnamese fishing waters, and on the Southeast Asian mainland the People’s Republic of China launched a war against Vietnam on the 17th of February 1979. On the 14th of March 1988, China again assaulted Vietnam, invading the South Johnson Reef. In 2016, China invaded outlying archipelagoes of the Philippines and commenced its illegal occupation of the Spratlys and Scarborough Shoal in violation of international law.

In addition to numerous other such belligerent territorial encroachments by China against neighbouring countries, a few of which are discussed elsewhere (van Driem 2021), within the Chinese-occupied territories the Chinese colonial government has pursued the eradication of Tibetan language and culture in Tibet, Mongolian language and culture in Inner Mongolia and Cantonese language and culture in Hong Kong.¹⁷

¹⁷ Beginning in 2015 with the kidnapping of the book publisher Guì Mínhǎi 桂敏海, a Swedish citizen of Chinese extraction, Peking began systematically to violate all the provisions of the Joint Declaration of 19 December 1984 signed by the United Kingdom and the People’s Republic of China. In order to obfuscate their extraterritorial kidnapping on Thai soil and incarceration of a Swedish national, the Chinese police played with the name of the victim, changing 桂敏海 Guì Mínhǎi to 桂民海 Guì Mínhǎi, and then contesting the identity of the kidnap victim. The complete breach of trust by the Chinese Communist Party with regard to the Joint Declaration culminated in the ruthless imposition of totalitarian rule over Hong Kong in 2019 and the brutal oppression of the people of Hong Kong, including imprisoning opposition candidates and protesters and the incidental practice of execution by organ procurement in breach of the dead donor rule (Robertson & Lavee 2022, Davies 2022). On the language front, Mandarin has been imposed upon the people of Hong Kong, and Choi Yuk-lin, Secretary of Edu-

Available documentation attests to the planned eradication not only of Uighur culture and language in Chinese-occupied East Turkestan, but even to the gradual government-orchestrated extermination of Uighur people by the Chinese Communist Party dictatorship in Peking in concentration camps the likes of which have not been seen since the days that the German government was led by the National Socialist German Workers' Party (Chin 2021, Pompeo 2021). Whilst seeking systematically to erase the language and cultures of the native populations in its colonies, China also exploits the native populace and extracts lithium, rare earths and mineral resources from occupied Tibet, East Turkestan and Inner Mongolia.

The aim of Chinese Communist Party policies is to expunge Tibetan language and culture from Chinese-occupied Tibet, destroying Tibetan architecture, libraries and ancient writings, institutions and enslaving and humiliating the people. In 1995, the Chinese Communist Party kidnapped the 6-year-old boy who had been recognised as the 11th incarnation of the Paṅ-chen Lama, causing this important figure in Tibetan Buddhism to join the ranks of the Tibetans who continue to vanish without a trace as *desaparecidos* by the hand of the Chinese colonial government. In their Tibetan Political Prisoner Database (TPPD), the Tibetan government in exile has been able to document over one hundred writers, songwriters and artists who have been incarcerated surreptitiously as political prisoners by the Chinese colonial forces. At the time of writing, beginning from the 29th of April 1998, 159 Tibetans in Chinese-occupied Tibet and 10 Tibetans living in exile have individually committed suicide by self-immolation in protest against the brutal oppression of the Chinese occupying forces. Most self-immolations have been committed in eastern Tibet outside of the so-called 'Tibetan Autonomous Region' created by the Chinese colonial administration (ICT 2022). In 2007, the ostensibly secular Chinese Communist Party presumed to legislate 'Management Measures for the Reincarnation of Living Buddhas in Tibetan Buddhism'. Today, the nation of Tibet ranks as the least free country or territory in the world, sharing the spot with South Sudan and Syria, ranking below even China and North Korea (Repucci & Slipowitz 2022).¹⁸

cation in Hong Kong, presaged in July 2022 that Cantonese must die out as a language of education, to be replaced by Mandarin.

¹⁸ In the context of the moribund Tibetan culture and a Tibetan language in the throes of death resulting from the sustained colonial onslaught of ruthless Chinese occupiers, Warner (2022) describes Tibetan clothing fashion and trends in Tibetan music today as ultimately futile displacement activities in response to what he calls 'the end of Tibet'.

In sequel to years of measures taken by the Chinese colonial government to discourage the use of the Tibetan language in occupied Tibet, in March 2022 apps in Tibetan language were blocked on phones and hand-held devices. At the same time, the Chinese colonial government began kidnapping children between four and six years of age from their parents, and today over one million kidnapped Tibetan children are being held in mandatory boarding pre-schools, where they are made not to speak Tibetan, taught Mandarin and indoctrinated in the belief system of the Chinese Communist Party (Tsomo *et al.* 2022, Campbell 2023, Feng 2023, Office of the High Commissioner for Human Rights 2023). In all government schools at the township and village level, Mandarin, the language of the colonisers, is used as the sole medium of instruction, whilst Tibetan, the native language of the country, is offered only as a subject. Since Mary Robinson's visit in 1998, China has blocked the United Nations High Commissioner for Human Rights from visiting Tibet in order to conceal the gorier details of the gross violations against the human rights of the Tibetan people perpetrated systematically by the Chinese colonial government.

American corporate concerns likewise play a highly destructive role in abetting the illegal occupation of Tibet and obliteration of Bodish language and culture. Looking north from any high point in Dolakhā district in Nepal, Melungtse མེ་ལུང་རཏེ *Me-luñ-rtse* stands at 7181 metres in neighbouring Tibet and can be seen to tower above all other summits of the Rolwaling Himal. The second highest peak Gaurī Śaṅkar, sacred to the goddess Pārvatī and the god Śiva, stands in Dolakhā district at 7134 metres. Between the two peaks, the name of the river valley, Melungchu མེ་ལུང་ཅུ *Me-luñ-chu*, preserves the same native toponym. Melung མེ་ལུང *Meluñ* is the village council area, known previously as a village *pañcāyat*, on the Nepali side of the border, and the contiguous area on the Tibetan side is likewise named Melung མེ་ལུང *Me-luñ*. It is unclear whether the place name is originally a native Tamang toponym or has some now lost etymology in the local variety of Tibetan, but the genuine place name is retained in both these two languages and is also used in Nepali. By contrast, the supposed Mandarin name which Wikipedia and Google Maps promote instead of the genuine mountain name Melungtse མེ་ལུང་རཏེ *Me-luñ-rtse* represents a toponymical fiction that has been newly concocted by the Chinese occupiers. Yet this 7181 metres tall example is but one small case in point.

In terms of surface area, Tibet is the tenth largest country in the world. Yet Google, an American multinational technology company, toponymi-

cally obliterates Tibet from the face of the earth. Google Maps supports the Chinese Communist Party's drive to eradicate Tibetan language and obliterate Tibetan culture by replacing native Tibetan place names with fake Mandarin toponyms dictated by the Chinese colonial administration. In this way, Google not only effaces the entire country of Tibet from the map of Asia but furthermore aids the Chinese occupying forces in Tibet in the eradication of Tibetan cultural identity through the artificial Pinyinisation of Tibetan place names at variance with the linguistic facts of Tibetan phonology and in violation of long-standing Tibetological conventions. The entire exercise serves to mask the remaining Tibetan place names that Google and the Chinese colonial administration have not yet managed to expunge completely.

Outside of Chinese-occupied Tibet, the sociolinguistic situation of Bodish languages differs greatly from the dismal reality in Chinese-held territories. Parts of historical Tibet are located in India and Nepal. In the western Indian Himalayas, three distinct phonologically extremely conservative Bodish dialects are spoken in *sBal-ti-yul* སྐལ་ཏི་ཡུལ་, which since the 1840s has been known as Baltistan. This portion of historical Tibet, which currently lies in Pakistan-occupied Jammu and Kashmir, comprises the districts གངས་ཆེ་ *Gaŋs-che*, རོང་མངོ་ *Roŋ-mdo*, མཁར་མང་ *mKhar-man*, སྐར་མངོ་ *sKar-mdo* and སེས་གར་ *Sis-gar*,¹⁹ where over 90% of the population are of Tibetan ethnicity. The three Bodic dialects are collectively referred to as Balti སྐལ་ཏི་ *sBal-ti*. The Balti language is used as a medium of instruction in roughly half of the schools in Baltistan, according to Muhammad Raza Ghalib, particularly in the more rural areas where Balti is used only because the teachers are weak in their command of Urdu and English. Balti language itself is not taught as a subject in the schools in Baltistan, and the 'Balti' spoken by teachers in the classroom is described by Raza Ghalib as a 'mixture of Urdu, Persian and English'. The Tibetan script or ཡི་གེ་ *yi-ge* is not taught in any of the schools in Baltistan, but two native Balti scholars organised free classes in Tibetan script during the winter months of 2018 and 2019. The Baltistan Student Federation སྐལ་ཏི་ཡུལ་ལི་སྐྱོབ་ལྷན་ཁག་གི་མཉམ་འབྲེལ་ *Bal-ti-yul-li Slob-phrug-kun-gyi mNam-ḥbrel*, which describes itself as 'a socio-educational organisation struggling [to] keep interest of educational, political and constitutional rights of Gilgit', conducts no organised activities to promote the Balti language or the Tibetan script.

¹⁹ The original Balti orthography of Shigar remains uncertain at this time, with several vying orthographies having been suggested. Raza Ghalib reports that the local toponym is a subject of ongoing investigation.

Whereas Islam reached Sindh in the 8th century, Baltistan was converted to Islam only in the 14th century (Iqbal 2018). The nearby region of Nūristān, formerly Kāfiristān, in neighbouring Afghanistan was only forcibly converted to Islam from their ancient native religious tradition in 1895. Pakistan, which was established in 1947, has occupied Baltistan and some adjacent parts of Jammu and Kashmir since 1948. The constitution of Pakistan declares Islam to be the state religion, and, although the document guarantees freedom of religion ‘subject to law, public order and morality’, mob lynchings of anyone perceived to have offended Islamic sensibilities are common, and people accused of alleged blasphemy are routinely officially executed or mobbed and murdered. Religious oppression and the violent persecution of religious minorities by Islamists are rampant (Office of International Religious Freedom 2022).

Balti advocates of the use of the native Tibetan script in Baltistan are therefore understandably careful to stress that ‘there is no link between religion and script’ (Raza Ghalib 2015), and in the Balti context today this statement appears to ring true. Rather, the use of the Tibetan script is tied historically with the mother tongue of the people and so experienced by many in Baltistan as a characteristic attribute of their ethnic identity, even though a mastery of the script is not yet common.

Inside India, Nepal, Sikkim and Bhutan, speakers of Bodish language enjoy linguistic rights and basic human freedom. The particulars of the sociolinguistic situation vary from place to place, but many commonalities can be observed. In most schools, the main medium of instruction is usually not a Bodic language, but instead Nepali, Hindi or English. In government schools in Sikkim and Bhutan, and also in private schools in Ladakh and Nepal, Classical Tibetan and the local Bodish language may be taught as subjects, whether this language be Dzongkha, Drenjongke or Ladakhi. Any restrictions tend to be of a socio-economic nature. In order to enhance career prospects, social standing and economic opportunity, of their children, parents and pupils tend to choose English in Bhutan, Nepali, Hindi and English in Sikkim, Nepali and English in Nepal, and Urdu and English in Ladakh.

In Ladakh, schools teach in Ladakhi medium until 5th grade, after which there is a transition to Urdu and English as the medium of instruction, but Ladakhi or *Bhoṭī* may continue to be chosen as an elective subject in higher years (Komissaruk 2021). The term *Bhoṭī* is historically just the Hindi word for Tibetan, *Bhoṭ* merely being the Sanskrit rendering of Phò ལྷོ [p^hɔ̃] *Bod* ‘Tibet’. Likewise, the Nepali terms *Bhoṭe*, *Bhoṭiyā* and *Bhuṭiyā* have historically been used to denote the tongue or the mem-

bers of any Tibetan language community, whichever Bodish tongue they happen to speak. The Tibetic lects or *Bhoṭī* languages stretching through the Himalayas from Baltistan to Bhutan and beyond represent distinct tongues with only a limited degree of mutual intelligibility between them. Consequently, the Hindi term *Bhoṭī* as a language designation suffers from all the same problems as the English term 'Tibetan', usually necessitating the use of apt local names such as Ladakhi, Drenjongke, Dzongkha and so forth for the particular language in question. In practice, the language taught as an elective subject in Ladakh is often Chöke ཅོ་སྐད་ *Chos-skad* 'liturgical language', i.e. Classical Tibetan, both because of the importance of the vast body of literature written in Classical Tibetan and because of the dearth of learning materials in Ladakhi. Yet great progress has been made in the development of Ladakhi pedagogical materials thanks to the efforts of the Students' Educational and Cultural Movement of Ladakh (SECMOL) since its establishment in 1988.

In Nepal, Durbar High School was established by Jaṅg Bahādur Rāṅā in 1853 as a school for the elite in Kathmandu. In 1951, the government school system was established by His Majesty's Government of Nepal and expanded on a national scale from 1971. Nepali has always served as the medium of instruction, and English is taught in schools as a mandatory subject. Like Japan, Thailand, Bhutan and Afghanistan, Nepal was never colonised by a foreign power, and consequently, except for in elite and private schools, the level of English instruction has historically not been very high. Indeed, English has not until recently begun to become as important in Nepal as it is, for example, in India. Only in recent years has the popularity of English amongst the younger generation in Nepal become so widespread as to begin to pose a threat, with English actually having replaced Nepali as the medium of instruction in some private schools.

After the abolition of the *pañcāyat* system and the introduction of a multi-party system in April 1990 by His Majesty King Virendra, language rights for all became enshrined in law. In practice, instruction in native mother tongues is limited to private schools and impeded by either a lack or a scarcity of suitable learning materials and a low economic incentive to finance mother tongue education by native language communities. By contrast, Chöke ཅོ་སྐད་ *Chos-skad* 'liturgical language' has always been available as a language of instruction for centuries in Buddhist monasteries throughout Nepal. Today in parts of Nepal where Bodish languages are natively spoken, such as Humla, Mugu, Mustang and Tsum, instructions in Tibetan is in private schools, which sometimes take the form of

boarding schools outside of the native language area, such as the Great Compassion Boarding School and Himalayan Buddhist Academy in Pokharā, where almost all of the youth of Mustang goes to school.

In Sikkim, every government school and Eklavya Model Residential School (EMRS) offers Drenjongke or 'Bhutiya language' as an elective. Officially and also often in practice, the medium of instruction is English in schools. However, most teachers teach in Nepali, which also happens to be one of the official languages of the state and serves *de facto* as the main language of the state since the Kingdom of Sikkim was annexed by India in 1975. Classes in Drenjongke, Lepcha and Limbu are offered as elective subjects in school. The Nepali immigration set into motion by the British in the 19th century led to the descendants of migrants from Nepal outnumbering the native peoples of Sikkim. As a consequence, in addition to Drenjongke, Lepcha and Limbu, school courses are today additionally offered in nine other Tibeto-Burman languages which the migrant ancestors of some young Sikkimese brought with them from Nepal.



FIGURE 2: The distribution of language in Bhutan: The East Bodish language in Bhutan are Mangde, Bumthang, Khengkha, Kurtöp, Dzala, Dakpa, Chali, and East Bodish furthermore includes the languages of Tawang and the Nyamnyang valley that have been researched by Tim Bodt and are discussed above.

The three native Sikkimese languages, Drenjongke, Lepcha and Limbu, were introduced in 2000 as major subjects in Bachelor's programmes in Sikkim under the aegis of North Bengal University, each deemed to represent a 'major Indian language'. With the establishment of Sikkim University in 2007, the language programmes were offered as three-year programmes followed by an honours course. Since 2021, the three native Sikkimese languages have been offered up to the Master's level, and in principle Drenjongke, Lepcha and Limbu are available as Ph.D. topics.

Bhutan, like Nepal, was never colonised by a foreign power. Nonetheless, English has by choice of the Royal Government of Bhutan been accorded a paramount role in formal education as the medium of instruction throughout the kingdom. Dzongkha is a mandatory subject, but only in the traditional monastic schools and in Simtokha Rigzhung School, established in Thimphu in 1961, is Dzongkha used as the principal medium of instruction, with English as a mandatory subject. None of the other languages of Bhutan are taught in formal education, but many Bhutanese acquire a rudimentary command of Tshangla, and all Bhutanese tend to have at least some command of Dzongkha, the national language of the kingdom.

Tibetan under the Hindi name *Bhoṭī* is taught in schools in parts of Arunachal Pradesh inhabited by Monpa language communities. Teachers are trained at the Central Institute of Himalayan Culture Studies in West Kameng. These *Bhoṭī* textbooks are based directly on the Tibetan textbooks developed by the Central Tibetan Administration for use in the Tibetan Children's Village (TCV) schools all over India and Nepal. Officials and administrators have only gradually begun to cotton on to the fact that teaching Tibetan, under whatever name, to speakers of Monpa does not constitute mother tongue education any more than would the teaching of the Bengali language to speakers of Konkani. Similarly, Classical Tibetan, which serves as a language of liturgy amongst the Monpa too, and the 'standard' variety of spoken modern Central Tibetan chosen by the legitimate government of Tibet in exile at Dharamsala are likewise two distinct languages. Intentions at all levels are good, but more linguistic information ought to be communicated to the relevant decision makers to obviate simplified understandings of rather complex linguistic realities.

In summary, Bodish languages are less under threat on the southern flank of the Himalayas than on the Tibetan plateau, where a hostile occupying power seeks to stamp out Tibetan culture and ethnic identity and so too even drive the Tibetan language into extinction. Along the south-

ern flank of the Himalayan chain, however, global and local socio-economic factors cause members of the younger generation to seek fluency in English and also to prioritise major regional languages, such as Nepali and Urdu, above their native Bodish tongue. Ironically, a major hurdle for speakers of all Bodish languages is literacy, precisely because of the richness and cultural dominance of the Classical Tibetan literary tradition. The orthography is antiquated and reflects an earlier stage of Bodish, whereas all modern Bodish languages, both on the Tibetan plateau and along the southern flank of the Himalayas, have undergone centuries of independent phonological development and grammatical evolution, each becoming a new language of its own.

Faced with the inordinate difficulty of learning the archaic spelling, voices of the younger generation and within the governments of Bhutan and Sikkim were raised in support of orthographic form. The first wave of orthographic reform in the 1970s in both Sikkim and Bhutan was haphazard in nature, rather than systematic and phonologically based. These efforts consequently led to numerous novel *ad hoc* spellings, creating new inconsistencies, alongside the retention of archaic spellings for most of the vocabulary. The result therefore merely exacerbated the spelling problem from the pedagogical point of view in both Bhutan and Sikkim. At the behest of the Royal Government of Bhutan, both Roman Dzongkha and Phonological Dzongkha were developed and introduced (van Driem 1991, 1992, 1994, 1998, Tshering & van Driem 2019), and under the auspices and sponsorship of the government of Sikkim both Roman Drenjongke and Phonological Drenjongke were developed (Namgyal & van Driem 2022). In Thimphu, Ratra Drukpa has organised workshops to train people in the use of Roman Dzongkha and Phonological Dzongkha, but time will tell whether or not these easy-to-learn and consistent spelling systems become widely adopted in future.²⁰

LOCATING THE BODIC, LEPCHA AND LIMBU HOMELANDS

The homelands of the Bodic languages, Lepcha and Limbu present linguistic questions which have not just a spatial but also a temporal dimension. Bodic, as defined here, comprises the languages of the Bodish

²⁰ At the 1st Tibetan Language Linguistic Forum, organised at Nánkāi University in Tiānjīn in August 2016, young Tibetans from various dialect areas of Tibet shared with me their curiosity about Roman Dzongkha and Phonological Dzongkha because they to were daunted by the difficulty of learning the spelling of the written language. However, each Bodish language or Tibetic lect may require its own regional orthography, just as Dutch, German and Bernese each have their own spelling systems.

subgroup and the languages traditionally subsumed under the label East Bodish, which are likely to form a polyphyletic set or, at least, a set of languages the internal phylogeny of which has not yet been understood. In the opening paragraphs of the preceding section, we situated the geographical locus of Old Bodish in space and time in the Yarlûng valley some time before the 7th century AD, perhaps as early as the beginning of the Christian era or before.²¹ As already noted, this historical Bodish linguistic homeland lies just 150 km north of the northeastern Bhutanese town of Trashiyangtse. The next relevant set of information is the geographical distribution of modern East Bodish language communities, as illustrated in the map in Figure 2.

The East Bodish languages in Bhutan are Mangde, Khengkha, Bumthang, Kurtöp, Dzala, Dakpa and Chali. Also subsumed under the label East Bodish are the languages of Tawang and the 'Nyamnyang valley that have been researched by Tim Bodt and that have been discussed above. Bodic, as defined here, comprises both Bodish and the other Bodic languages, designated by the cover term 'East Bodish', and, given the geographical distribution of the Bodic languages in time and space, the most parsimonious hypothesis for their point of origin is that the Bodic linguistic homeland must have lain somewhere in the region between the Yarlûng, Kurichu, Mangde and Tawang river valleys.

The next consideration, however, is the presence of other native language communities in Bhutan, such as the Lhokpu in southwestern Bhutan, the Black Mountain Mönpa in the Black Mountains, the Gongduk in south-central Bhutan and the Tshangla in eastern Bhutan. The range of the Lhokpu formerly extended further north in western Bhutan, and the range of the Black Mountain Mönpa is likewise known to have been greater than it is today. The same applies to the Gongduk, who explicitly identify themselves as the aboriginal འདྲུང་ *gDuñ* populace of central Bhutan. There is no question that the Tshangla represent a populous native group of eastern Bhutan. Since some of these groups are also thought of historically as representing earlier inhabitants by the speakers of Bodic languages in Bhutan, it is logical to posit a geographical locus for Bodic just north of the southern flanks, and this is where a dot has been placed on the map in Figure 3, representing the locus of Bodic at some indeterminate point in the past that we may conjecturally position at the first half of the 1st or maybe the second half of the 2nd millennium BC.

²¹ Chamberlain (2015) has attempted to relate Tibetan dialectal geography to riverine watersheds, but the presumed dispersal of Bodish in his discussion is posterior in time to the points in time and in space posited here.



FIGURE 3: Geographical distribution of Trans-Himalayan subgroups, each dot representing the historical centre of one of the 41 linguistic subgroups. Each dot represents not a language but a linguistic subgroup, each of which may comprise between one to several dozen languages. The underlying map has been provided by the New Himalaya Decolonial Atlas at Cloudy Bay or, in Māori, Te Koko a Kupe. According to the cartographers, the borders of Tibet, as shown, represent the historical national frontiers before the illegal occupation of the country by Chinese colonial forces in 1950. The Cloudy Bay map also delineates the borders of East Turkestan, which was an independent republic from 1933 to 1934 and again from 1944 to 1949, before being subjugated again by Chinese colonial forces. The legitimate Tibetan government has been headquartered at Dharamsala since 1959, and the government-in-exile of East Turkestan is headquartered in Washington. Also indicated are the borders of Southern Mongolia, which likewise seeks independence from Chinese rule. The Cloudy Bay cartographers have drawn the northwestern border of Nepal in conformity with the Treaty of Sugaulī of 1816 concluded between the Kingdom of Nepal and the East India Company. British maps drafted in the years after the treaty depicted the Kālī river, which crosses the Tibetan border at Limpiyādhurā, as the border between the Kingdom of Nepal and East India Company territory. The boundaries on this map are representational only and purport neither to be accurate nor to imply endorsement by the author or the publisher.

In Figure 3, the geographical centre of gravity of the language family as a whole can be seen to lie within the arc of the eastern Himalayas and the Indo-Burmese borderlands. Out of 41 Trans-Himalayan subgroups, 29 branches are found either exclusively or predominantly south of the Himalayan divide within the Indian subcontinent, viz. Tamangic, Newaric, Kiranti, Lepcha, Digarish, Lhokpu, Midžuish, Chepangic, Magaric, Tani, Siangic, Raji-Raute, Tshangla, Kho-Bwa, Ao, Zeme, Angami-Pochuri, Karbi, Brahmaputran, Mru, Gongduk, Hrusish, Black Mountain, Dhimalish, Tangkhul, Meithei, Pyu, Karenic and Mizo-Kuki-Chin. Seven Trans-Himalayan linguistic subgroups are found to the north and east of the Himalayas, viz. Bái, Tūjiā, rGyalrongic, Qiāngic, Ěrsūish, Nàic and Sinitic. Five branches of the Trans-Himalayan family are represented by language communities distributed both along the northern side and on the southern flank of the Himalayas, viz. Bodish, Lolo-Burmese, Nungish, West Himalayish and Kachinic.

Whereas all of the Bodic languages are represented by a single dot in Figure 3, so too the Lepcha language by itself constitutes a Trans-Himalayan linguistic subgroup in its own right and is represented by a dot of its own. As Saul Mullard (2011: 5-9) has argued, even on the basis of the previously prevailing interpretation of the archaeological record, the linguistic ancestors of the Lepcha are likely to have settled in Sikkim by 5,000 BC. Today, the combined evidence of the ethnolinguistic phylogeography of Trans-Himalayan language communities and the genetic phylogeography of Y-chromosomal lineages associated with the Trans-Himalayan populations now push the probable date of Lepcha habitation in Sikkim even further back into the past (van Driem 2021).

By the same token, of course, no population at such a distant time in the past spoke Lepcha. Rather, the ancient population inhabiting Sikkim millennia ago spoke some ancient Trans-Himalayan tongue which over millennia would locally evolve into Lepcha. We may call this ancient tongue proto-Lepcha, and the best hope at arriving at an understanding of this lost tongue lies in historical linguistic comparison in combination with internal linguistic reconstruction of Lepcha. Robert Andrew Dermot Forrest (1962) identified a large number of Austroasiatic cognates in Lepcha, which led to the hypotheses that the Lepcha language may either have arisen when a Trans-Himalayan tongue absorbed many Austroasiatic loans through close contact, or when a Trans-Himalayan language was adopted by an indigenous Austroasiatic population. Sprigg (1982) pointed out that Lepcha is, unsurprisingly, replete with Drenjongke loans after centuries of strong contact influence, leading to Bodish

borrowings having replaced native Lepcha lexical material, which could ultimately have been of either Trans-Himalayan or Austroasiatic provenance. Whilst recognising the Bodish influence on the Lepcha lexicon, Bodman (1988) established that Lepcha itself was genetically a Trans-Himalayan tongue with either an Austroasiatic adstratum or substrate.

Since the dawn of the Holocene, many ancient hunter-forager populations ancestral to modern Trans-Himalayan language communities settled all along the long arc of the eastern Himalayas and sub-Himalayan mountain tracts in the lush jungles of its sparsely populated southern flank. So too the ancient population ancestral to the Lepcha probably inhabited the area that today is Sikkim from this time. The relative frequencies of the Y-chromosomal haplogroup O1b1a1a (M95) in particular Trans-Himalayan language communities of the Indian subcontinent (Sahoo *et al.* 2006, Reddy *et al.* 2007, Gazi *et al.* 2013, van Driem 2021) suggest that a subset of the paternal ancestors of certain Trans-Himalayan populations, e.g. the Lepcha, the Mizo and perhaps certain Bodo-Koch communities, were men who may have spoken antique languages of the Khasian or Palaungic branches of the Austroasiatic language family, and who were absorbed and assimilated into Trans-Himalayan language communities.

Based on the comparison of Lepcha with Karlgren's reconstruction of Old Chinese and the historical phonology of Lepcha, which featured the evolution of the ancient Trans-Himalayan *s*-prefix into post-consonantal palatal offglides, Nicholas Bodman (1973, 1980) argued for a close relationship between Lepcha and Sinitic and even proposed the name 'Sino-Himalayan' to label this relationship. As Bodman explained to me at the conference dedicated to Trans-Himalayan languages held at the University of Lund in October 1988, this term 'Sino-Himalayan' could be used either to label a hypothetical taxon comprising Lepcha and Sinitic or as a label for the entire language family.

In Bodman's prescient view of how the field of Trans-Himalayan linguistics would develop, the eastern Himalayan area represented the centre of gravity of the language family. In 1998, Ilia Peiros (1998: 217) advanced the idea that the Tibeto-Burman homeland lay in the sub-Himalayan regions of the eastern Himalayan arc. When Peiros (2004) expressed this view at an international conference in Geneva, a few of the participants who at the time still adhered to the empirically unsupported and now obsolete Sino-Tibetan paradigm then chiefly propagated from Berkeley and Peking greeted this thesis with incredulity or even scoffed at the idea, but not the eminent Russian scholar Sergej Anatol'evič Starostin,

who was also in attendance in Geneva and who had in fact espoused the same view since 1994.

At an international conference at Sèvres, Starostin (1994) presented a new view of the language family which he termed Sino-Kiranti. Based on the Limbu grammar (van Driem 1987), he stressed the archaic nature of Kiranti and shared commonalities with reconstructed Old Chinese (Starostin 1989). On the blackboard, he drew a trifurcating tree of the Sino-Kiranti family splitting into Sinitic, Kiranti and ‘Tibeto-Burman’ and an alternative bifurcating model with Sino-Kiranti, splitting into Kiranti and ‘Sino-Tibetan’. The two drawings were defiantly presented as a challenge to our thinking.

When I subsequently proposed the subgroup Sino-Bodic (van Driem 1997), the underlying assumptions were that Limbu was a member of the Kiranti group and that the Kiranti languages were somehow part of the nebulous and ponderous construct which Shafer had labelled Bodic. In view of the data adduced, the 1997 subgroup proposal should have been called Sino-Kiranti, rather than Sino-Bodic. However, the denomination ‘Sino-Kiranti’ had already been taken, with Starostin having used the label three years earlier to designate two alternative proposals for the overall structure of the language family as a whole. Frederik Kortlandt therefore suggested to me that I chose the label ‘Sino-Bodic’ instead. In retrospect, it would have been more accurate to have repurposed the label ‘Sino-Kiranti’.

At the same time, the unity of Kiranti as a valid linguistic subgroup was first called into question by Werner Winter (1986), and these doubts are pursued in my discussion of the isoglosses which separate the Limbu dialects from the Rai languages and the intermediate status of Yakkha, Chulung and Āṭhpaharīya, which may represent Rai languages which have undergone Limbu influence through their geographical proximity and long-standing ties of intermarriage (van Driem 2001: 664, 719 *et passim*). Gerber & Grollmann (2018) have pursued the issue of whether Kiranti constitutes a valid subgroup. Figure 3 still represents Kiranti with a single dot, that has been situated at a locus between Limbuwan and the region that constitutes the patchwork quilt of diverse Rai language communities. Future research may necessitate the use of two loci on the map to represent two distinct subgroups in this region.

The Limbus view themselves as the original inhabitants of the Tamor watershed, and the Lepchas as the native pre-Bodish denizens of both the Rangit and Teesta watersheds. With the sparse demography of the past times, the ranges of the Limbus and Lepchas could not just have met but also overlapped, as is in fact the case with many language communi-

ties throughout the Himalayas even today. Lepcha settlements have historically been reported as far to the west as in what today is Ilām district, where no extant ancient Lepcha settlements are found today. Some Limbus claim that their range once extended to the Teesta.

Limbu and Lepcha have been residing in their present areas for all of recorded history, and linguistic phylogeography informs us that their immediate linguistic ancestors may have inhabited these same areas for millennia. Recorded history informs us that the ancestors of the Drenjongpa have been settled in Sikkim for centuries, where they have lived in harmony and extensively intermarried with the Lepcha and, to a lesser extent, with the Limbu. At a much more remote point in time, the linguistic ancestors of the Bodic language communities are likely to have lived north of the Himalayas in the area between Sikkim, Bhutan and the Yarlung valley.

In considering the proposals of a close genetic relationship between Lepcha and Sinitic, proposed by Bodman, or Limbu and Sinitic, as proposed by Starostin, the geographical distribution of Trans-Himalayan linguistic subgroups in Figure 3 prompts some reflection. Hungarian is spoken in the heart of Europe. Yet its closest linguistic relatives are Khanty and Mansi, which together form the Ugric branch of Uralic. Other than Samoyed, Ugric represents the most easterly branch of the Uralic language family. Nonetheless Hungarian happens to be the westernmost Uralic language. Similarly, within the Iranian branch of languages, Ossetic is phylogenetically an Eastern Iranian language, much more closely related to Pashto in eastern Afghanistan than to Western Iranian languages such as Kurdish and Persian. Yet Ossetic, spoken in the north Caucasus, happens to be the westernmost Iranian language. An analogous situation is likely to obtain with respect to Sinitic in light of the close affinity proposed with Lepcha and Limbu.

In Figure 3, the dot representing Sinitic in the far northeast marks an outlier, the result of an ancient population movement of migrants out of the Trans-Himalayan linguistic homeland in the eastern Himalayan arc towards the Yellow River basin, which flourished as a comparatively affluent region in Neolithic times. An intermediate position on the map is marked by the dot representing Tüjjiā, a language community whose linguistic ancestors must have straggled along the same trail eastward from the Trans-Himalayan heartland. As the linguistic ancestors of the Chinese migrated to the remote northeast, this founding Sinitic language community was buttressed by the winds of language change and contact influence. What today is China was already inhabited by language com-

munities speaking tongues belonging to unrelated linguistic phyla, such as Hmong-Mien, Altaic and Yenisseian. The resultant contact situations led Sinitic to acquire creoloid structural traits and undergo extensive lexical replacement of ancestral Trans-Himalayan vocabulary by borrowings from the language communities that were already inhabiting the Neolithic cultures of the Yellow River basin.

Because of the presence of Altaic language communities in modern times, scholars have in the past often conjectured that the inhabitants of the Yellow River basin were speakers of an ancient Altaic language (He *et al.* 2019, Chaubey & van Driem 2020). However, Edwin George Pulleyblank (1962) argued that the early inhabitants of the Yellow River were ethnolinguistically related to the people whom the Chinese called 匈奴 Xiōngnú, recorded historically as living north of the Chinese in the 3rd century BC. Pulleyblank (1962, 1966, 1983, 2000) developed the case for the Yenisseian ethnolinguistic identity of the Xiōngnú, and his theory was adopted by those scholars most knowledgeable of Yenisseian languages, viz. Dul'zon (1964), Krejnovič (1968), Doerfer (1973), Jaxontov (1986) and Vovin (2000).

Subsequently, Vovin *et al.* (2006) identified the historically recorded Xiōngnú autonym 羯 **Kjet* with the Yenisseian ethnonym Ket. Finally, Gāo (2013, 2021) has adduced Sinitic lexemes with possible Yenisseian etymologies. It therefore appears plausible that the earliest Sinitic or Old Chinese arose when the language of Trans-Himalayan immigrants was adopted by resident Yenisseian language communities, leading to creoloid features observed in Chinese as compared with the Tibeto-Burman languages still spoken within the Trans-Himalayan homeland. Similarly, lexical and grammatical traits of the Brahmaputran languages have been interpreted by DeLancey (2014) as evidence that these languages likewise underwent a process of creolisation in the distant past, making their apparent divergence similarly a secondary effect. In terms of ancient contact situations and their linguistic effects, the upheaval of a long migration to the remote Yellow River basin may had a counterpart in the turbulent migratory history of the Brahmaputran alluvial plain. By contrast, the linguistic ancestors of the Limbu and Lepcha enjoyed living in languor, tucked away in the lush refuges of high alpine valleys in a portion of the eastern Himalayan arc which lay well to the lee of migratory upheaval over time.

FROM PHENOTYPICAL IMPRESSIONISM TO MOLECULAR GENETICS

The fascination with the highly divergent phenotypes of our fellow man is attested by the somatological descriptions of various peoples in the writings of Herodotus, Vergil, Strabo, Diodorus, Xenophanes and Manilius. Snowden (1970, 1989) studied the wide range of expressions in Greek and Latin texts to describe the skin colour, physiognomy, stature and physical attributes of Egyptians, Colchians, Ethiopians and other peoples than Greeks and Romans in antiquity, as well as the depictions of other ethnicities in Graeco-Roman art and sculpture.

No doubt the interest in each other's appearance and the penchant for taking note of the phenotypical peculiarities of our conspecifics has been an abiding inclination of human beings from the time of our distant australopithecine ancestors. The first modern scientific attempt to classify humans phenotypically into 'races' was undertaken in 1684 by François Bernier, who published this first classificatory schema of mankind in the *Journal des Sçavans*. The history of the rise of 'race' in scholarly thinking until the demise of the notion of race in the face of molecular genetic findings at the end of the 20th century is told in *Ethnolinguistic Prehistory* (van Driem 2021). The book dissects the history of the very idea of a so-called 'Mongoloid race', tracing the origins of this notion back to Königsberg in the year 1774.

What's in a name? This question is a necessarily rhetorical one because the choice of names and labels does indeed to a large extent shape perceptual reality. In the popular imagination, the idea survives amongst many people in Sikkim and northeastern India that their linguistic or immediate genetic ancestors came from Mongolia. To begin with, the Mongolic languages, being members of the Altaic language family, are unrelated to the Sinitic languages, which, like Limbu, Lepcha and Drenjongke, belong to the Trans-Himalayan language family. Furthermore, we all have numerous ancestral lineages, not just one line of descent, and our linguistic ancestors and our biological forebears need not have been the same people.

The dots in Figure 3 assign conjectural geographical loci to the oldest reconstructible stages of the respective Trans-Himalayan linguistic subgroups at some time in the past, perhaps a few millennia after the dawn of the Holocene. Languages change at a relatively fast pace, and vast stretches of time lie beyond the linguistically reconstructible past. When we go far back enough in time, all of our ancestors came from Africa. Global mitochondrial phylogeography also shows our ancestral emergence from Africa with great clarity (Oppenheimer 2012). Everywhere in

the world, the mitochondrial landscape tends to be much older than the paternal lineages. By contrast, Y-chromosomal phylogeography tends to be younger and often correlates well with the geographical spread of language families except in a number of salient rare cases.

When we take such a long view of time, the question of being indigenous or native to a place in terms of ancestry becomes preposterous, but people more usually define their ethnic identity in historical time or in terms of traceable provenance at shallower time depths. The centre of linguistic diversity of the Trans-Himalayan subgroups, as shown in Figure 3, puts the lie to the notion that the linguistic ancestors hailed from China, let alone Mongolia. From the linguistic point of view, Lepcha and Limbu are completely distinct and divergent subgroups. Yet both Lepcha and Limbu share the distinction that scholars such as Bodman and Starostin have perceived close links between them and Sinitic, although the proposed phylogenetic propinquity remains hypothetical in both cases.

When the evidence of Y-chromosomal phylogeography is examined, the language communities yield molecular evidence that provides independent corroboration of the reconstruction of the past already developed on the basis of linguistic evidence. The paternal lineages O2 (M122) and O2a2b1 (M134) are characteristic molecular markers for communities speaking Trans-Himalayan language, where in some cases the paternal haplogroup O2a2b1 (M134) may be borne by up to 100% of the men of a particular Tibeto-Burman language community. As detailed in *Ethnolinguistic Prehistory* (van Driem 2021), ancient DNA studies support the hypothesis bearers of these paternal lineages introduced Trans-Himalayan language to the Yellow River basin, where they met and assimilated bearers of the “Yenisseian” paternal lineages Q1a1 (M120), Q3a (M324) and Q3a3 (P201), although the paternal lineage N (M231) was also found (Zhao *et al.* 2011, 2014, 2015, Huang & Li 2017, Cui *et al.* 2020).

Ancient DNA identified as representing the denizens of the Xiōngnú empire, dating from between 209 and 98 AD, shows an amalgam of lineages, in descending order of frequency, Q1a (F1096), R1a (M420), C2b (F1067) and a medley of other paternal haplogroups, including the earliest case of an O haplogroup this far north. The early and late mediaeval sites from Mongolia likewise show more than a dozen paternal lineages, with the relative proportion of haplogroup Q1a dwindling, and the proportion of “Altaic” C2b increasing before and during the mediaeval period (Jeong *et al.* 2020, Lee *et al.* 2023: s1a). Assuming the applicability of the Father Tongue correlation, the ancient DNA evidence lends support to a Yenisseian linguistic substrate on the North China plain.

ETHNOCENTRISM AND OTHER PITFALLS TO AVOID

In Paris, I once playfully asked my young companion, who had recently won a national beauty pageant, whether he thought that Paris might be the centre of the known universe, as I often contend in jest. The nonplussed male model countered by chastising me for apparently not knowing that there were other planets in the solar system, even other stars and entire solar systems and even other galaxies. With a slightly exasperated edge in his voice, the young Mr. France instructed me that 'Paris est évidemment le centre du monde, mais pas de l'univers entier!' [Paris is obviously the centre of the world, but not of the whole universe!].

In the 1990s at Siem Reap, Gérard Diffloth once showed me an amazing book written in 1892 by Henri-Nicolas Frey, a Corsican colonel of the *Infanterie de Marine* who served in various French colonies in Africa, Oceania and Asia, including Tonkin, and ultimately attained the rank of major general. Frey (1892) claimed to have proved that all human languages derived from Vietnamese, and on a map in his book he illustrated how the world had been peopled, depicting lines of migration all emanating from Tonkin in French Indochina across the face of the entire planet.

Frey's grand hypothesis formed part of his own eclectic understanding of the many racial theories that were current in his day. In order to explain correspondences between the mammalian fossil records of India and Madagascar, Philip Lutley Sclater (1864: 219), in a carefully worded passage, proposed the name Lemuria to designate a former continent that later broke up into the Indian subcontinent, Madagascar, Africa and part of the Americas. Based on evidence from the natural history of mammalian evolution, Sclater had in 1864 essentially proposed the previous existence of a continent that today, with our modern understanding of tectonic plate theory, is termed Gondwanaland.

Nine years later, Ernst Heinrich Haeckel, though well versed in geological chronology as it was understood in his time, fell prey to anachronism when he imagined that Sclater's Lemuria represented 'die wahrscheinliche Wiege des Menschengeschlechts, das hier sich vermuthlich zuerst aus anthropoiden Affen hervorbildete' [the probable cradle of the human race, which presumably first developed here from anthropoid apes] (1873: 321). Blundering into even greater anachronisms, which mixed events at vastly different time depths, Frey (1892: 136b) presented his own rendition of Haeckel's map (1873: 689), whilst arguing that the actual site of Lemuria was actually Tonkin, using heterogeneous arguments such as the distribution of the orang utan and random linguistic

chance resemblances, known as ‘look alike’, which he, of course, was able to find between the most disparate languages of the world.

Strangely, this very same monogeneticist theory about the peopling of the world from Tonkin is still espoused today by the Vietnamese journalist Hà Văn Thù, who was born at Thái Bình in 1944, twelve years after the multiply decorated major general Frey died at Menton in the Alpes-Maritimes. Hà’s books have newly appeared in English translation (Hà 2020, 2021a, 2021b). Liam Kelley (2020) writes that Hà is merely one of the louder exponents of this ethnocentric view of prehistory, that is currently widespread in Vietnam as well as amongst the Vietnamese diaspora. Contrary to what Kelley supposes, however, the ‘centrality’ of this ethnocentric view amongst the Vietnamese is not new.

Hà Văn Thù is, in fact, the principal proponent of this view, and he found his inspiration directly in the writings of colonel Frey, whose 1892 monograph he repeatedly cites. This now popular strand of modern Vietnamese lore therefore stems directly from the Corsican colonel. In his work, Hà Văn Thù sees Frey’s theory as corroborated by the genetic studies of Stephen Oppenheimer, Chuán-Chào Wáng and Huī Lǐ. Hà Văn Thù also seeks inspiration and corroboration in the ethnological writings of Nguyễn Đình Khoa, who sees Vietnam as the cradle of the Mongoloid and Australoid ‘races’, whose mixed progeny purportedly spread throughout the world.

Like Frey in 1892, Hà Văn Thù today believes that all languages of the world can be derived from Vietnamese. Famously, the Brabantian scholar Goropius Becanus (1569) propounded the theory that all of the languages of the world derived from Dutch. However, Goropius Becanus was born at Gorp near Hilvarenbeek in 1519, years before Sigismundus Gelenius wrote his *Lexicon Symphonum* in Basel in 1537 and some time before Indo-European historical linguistics was first developed in Leiden between 1597 and 1647.

Now that historical linguistics has grown over centuries into a full-fledged discipline with a sophisticated instrumentarium, it smacks of a certain quaint *ringardise* today to be espousing Frey’s linguistic theory of Vietnamese representing the mother of all languages. So compelling can be our inclination toward ethnocentrism. Another common pitfall is anachronism, for the temptation to indulge in an ethnocentric view need not only be linguistic or spatial, but may also be temporal, as when we project our ethnic identity onto the historical past or even into prehistory.

At the time of the Buddha, there were no such languages as English or French, and there were no English or French people. By the same token, when the Buddha walked the earth, there were no such languages

as Limbu, Lepcha, Nepali or Drenjongke, and at that time there were no people who went by names such as Limbu, Lepcha, Nepali or Drenjongpa. To project modern identities onto the past leads to the common error of anachronism. The paintings on the walls of the caverns of Lascaux do not represent early French naïve art, nor was the Buddha born in the Federal Democratic Republic of Nepal. Neither did Pocohontas live in the United States of America.

Modern citizens in England and in Sikkim may feel comfortable with the idea that no Englishmen or Drenjongpas existed at the time of Alexander the Great. Yet in the People's Republic of China, the official scholarly narrative espoused by the Communist Party of China seeks to project a Hân Chinese identity anachronistically onto the past. When studying the scientific literature, readers should be aware that Chinese scholars fall into this pitfall in their writings, not just because of political doctrine, but also because this is a perennial blind spot. Chinese archaeologist Kwang-chih Chang (1983) therefore knowingly warned his countrymen against the anachronisms that arise from affixing the label 'Chinese' to archaeological cultural assemblages or peoples of the distant past.

Chang stressed that the ancient polities Xià, Shāng and Zhōu in the Yellow river valley are likely to have represented ethnolinguistically distinct populations. The ancient cultures on what today is the North China Plain were not necessarily peopled by populations directly ancestral to today's Hân Chinese. Modern national and ethnic identities only arose or were invented in the course of recent historical time. In ancestral terms, none of us are fully native to our native countries. Because the past took such an awfully long time, none of us are truly sons of the soil except perhaps in the short term of our lifetime compounded by some number of generations in the past. Pure ethnic groups do not exist.²²

²² Even the Japanese, who rightfully stress their cultural distinctness, and who have also in the past been wont to pride themselves on their 'racial' distinctness, are not sons of the soil. My invited lecture entitled 'Who are the Japanese, and where do the Japanese come from?', given at the workshop 'Human Evolution in Eurasia elucidated through Genetics, Archaeology and Linguistics' hosted by the National Institute of Genetics at Mishima on the 17th of March 2017, detailed the tripartite origins of Japanese populations (van Driem 2017). Gyaneshwer Chaubey and I juxtaposed the substance of the 2017 Mishima talk on the tripartite origins of Japanese to the contrasting case of Munda languages (Chaubey & van Driem 2020). The three waves of peopling that gave rise to the modern Japanese people are elaborate in *Ethnolinguistic Prehistory* (van Driem 2021). A new study on ancient DNA has corroborated our model of the tripartite origins of the Japanese people with some supplementary findings (Cooke *et al.* 2021).

REFERENCES

- Ardussi, John Albert. 1977. *Bhutan Before the British: A Historical Study*. Canberra: Australian National University Ph.D. thesis, October 1977.
- Ardussi, John Albert, Anna Balikci Denjongpa and Per Kjeld Sørensen. 2021. *The Royal History of Sikkim: A Chronicle of the House of Namgyal as Narrated in Tibetan by Their Highnesses Chogyal Thutob Namgyal and Gyalmo Yeshe Dolma, Based upon the Preliminary Translation by Kazi Dawasamdub, Corrected, Supplemented and Thoroughly Revised*. Chicago: Serindia Publications.
- Aris, Michael. 1979a. *Bhutan, the Early History of a Himalayan Kingdom*. Warminster: Aris & Phillips Ltd.
- Aris, Michael. 1979b. 'Notes on the history of the Mon-Yul corridor', pp. 9-20 in Michael Aris and Aung San Suu Kyi, eds., *Tibetan Studies in honour of Hugh Richardson: Proceedings of the International Seminar of Tibetan Studies*. Warminster: Aris & Phillips Ltd.
- Bielmeier, Roland. 2004. 'Shafer's proto-West Bodish hypothesis and the formation of the Tibetan verb paradigms', pp. 395-412 in Anju Saxena, ed., *Himalayan Languages, Past and Present*. Berlin: Mouton de Gruyter.
- Bielmeier, Roland, et al. 2018. *Comparative Dictionary of Tibetan Dialects, Volume 2: Verbs*. Berlin: Mouton de Gruyter.
- Bodman, Nicholas Cleaveland. 1973. 'Some Chinese reflexes of Sino-Tibetan s- clusters', *Journal of Chinese Linguistics*, 1 (3): 383-396.
- Bodman, Nicholas Cleaveland. 1980. 'Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship', pp. 34-199 in Frans van Coetsem and Linda R. Waugh, eds., *Contributions to Historical Linguistics: Issues and Materials*. Leiden: E.J. Brill.
- Bodman, Nicholas Cleaveland. 1988. 'On the place of Lepcha in Sino-Tibetan: A lexical comparison', *Linguistics of the Tibeto-Burman Area*, 11 (1): 1-26.
- Bodt, Timotheus Adrianus. 2014. 'Ethnolinguistic survey of westernmost Arunachal Pradesh: A fieldworker's impressions', *Linguistics of the Tibeto-Burman Area*, 37 (2): 198-239.
- Bodt, Timotheus Adrianus. 2023. 'East Bodish revisited', *Bulletin of Tibetology*, 53 (2): 47-199.
- Bodt, Timotheus Adrianus. [forthcoming]. *Proto-Western Kho-Bwa: Reconstructing the Past of a Small Indigenous Community* (Academia Sinica Language and Linguistics Monographs). Taipei: Academia Sinica.
- Bodt, Timotheus Adrianus. 2021. 'The Duhumbi perspective on Proto-Western Kho-Bwa onsets', *Journal of Historical Linguistics*, 11 (1): 1-59.
- Bodt, Timotheus Adrianus. 2019. 'The Duhumbi perspective on Proto-Western Kho-Bwa rhymes', *Die Sprache*, 52 [2016 / 2017] (2): 141-176.
- Bradley, David. 1997. 'Tibeto-Burman languages and classification', pp. 1-71 in David Bradley, ed., *Papers in Southeast Asian Linguistics № 14*. Canberra: Pacific Linguistics.
- Bradley, David. 2002. 'The subgrouping of Tibeto-Burman', pp. 73-112 in Christopher Beckwith, ed., *Mediaeval Tibeto-Burman Languages*. Leiden: Brill.
- Bradley, David. 2012. 'Tibeto-Burman languages of China', Vol. IV in Rint Sybesma et al., eds., *Encyclopaedia of Chinese Languages and Linguistics*. Leiden: Brill.
- Campbell, Charlie. 2023. 'China's residential schools separate a million Tibetan children from their families, U.N. says', *Time*, 7 February 2023.
- Chamberlain, Brad. 2015. 'Linguistic watersheds: A model for understanding variation among the Tibetic languages', *Journal of the Southeast Asian Linguistics Society*, 8: 71-96.

- Chang Kwang-chih. 1983. 'Sandai archaeology and the formation of states in ancient China: Processual aspects of the origins of Chinese civilization', pp. 495-522 in David N. Keightley, ed., *The Origins of Chinese Civilization*. Berkeley: University of California Press.
- Chin, Josh. 2021. 'Leaked documents detail Xi Jinping's extensive role in Xinjiang crackdown: Records of internal speeches released by U.K.-based panel show how China's leader set the blueprint for vast campaign of forcible assimilation against Uyghurs, other groups', *Wall Street Journal*, 30 November 2021.
- Cooke, Niall P., Valeria Mattiangeli, Lara M. Cassidy, Kenji Okazaki, Caroline A. Stokes, Shin Onbe, Satoshi Hatakeyama, Kenichi Machida, Kenji Kasai, Naoto Tomioka, Akihiko Matsumoto, Masafumi Ito, Yoshitaka Kojima, Daniel G. Bradley, Takashi Gakuhari and Shigeki Nakagome. 2021. 'Ancient genomics reveals tripartite origins of Japanese populations', *Science Advances*, 7, article № eabh2419.
- Cui, Yinqiu, Fan Zhang, Pengcheng Ma, Linyuan Fan, Chao Ning, Quanchao Zhang, Wei Zhang, Lixin Wang and Martine Robbeets 2020. 'Bioarchaeological perspective on the expansion of Trans-Eurasian languages in Neolithic Amur River basin', *Evolutionary Human Sciences*, 2020 (2): e15.
- Davis, Michael C. 2022. 'Hong Kong: How Beijing perfected repression', *Journal of Democracy*, 33 (1): 100–115.
- DeLancey, Scott. 2014. 'Creolization in the divergence of Tibeto-Burman', pp 41–70 in Thomas Owen-Smith and Nathan Hill, eds. *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Tibeto-Burman Area*. Berlin: Mouton.
- DeLancey, Scott. 2021. 'Classifying Trans-Himalayan (Sino-Tibetan) languages', pp. 207-223 in Paul Sidwell and Mattias Jenny, eds., *The Languages and Linguistics of Mainland Southeast Asia*. Berlin: Mouton de Gruyter.
- Doerfer, Gerhardt. 1973. 'Zur Sprache der Hunnen', *Central Asiatic Journal*, 17 (1): 1–50.
- van Driem, George. 1987. *A Grammar of Limbu*. West Berlin: Mouton de Gruyter.
- van Driem, George. 1991. གཞུང་འབྲེལ་རྫོང་ཁོང་གི་སྐད་ (འཛམ་གླིང་རྒྱབ་ཕྱོགས་པའི་ཡིག་ཐོག་) འབྲི་ལུགས་ལམ་སྟོན། *Guide to Official Dzongkha Romanisation*. Thimphu: Royal Government of Bhutan.
- van Driem, George. 1992. རྫོང་ཁོང་གི་བརྗོད་སྟོན་པའི་གཞུང་། *The Grammar of Dzongkha*. Thimphu: Royal Government of Bhutan.
- van Driem, George. 1994. 'The phonologies of Dzongkha and the Bhutanese liturgical language', *Zentralasiatische Studien*, 24: 36-44.
- van Driem, George. 1997. 'Sino-Bodic' *Bulletin of the School of Oriental and African Studies*, 60 (3): 455-488.
- van Driem, George. 1998. རྫོང་ཁོང་གི་ཡིག་ཐོག་། *Dzongkha*. Leiden: School of Asian, African and Amerindian Studies.
- van Driem, George. 2001. 'Zhangzhung and its next of kin in the Himalayas', pp. 31-44 in Yasuhiko Nagano and Randy LaPolla, eds., *New Research on Zhangzhung and Related Himalayan Languages* (Bon Studies 3, Senri Ethnological Reports 19). Osaka: National Museum of Ethnology.
- van Driem, George. 2001. *Languages of the Himalayas: An Ethnolinguistic Handbook of the Greater Himalayan Region, containing an Introduction to the Symbiotic Theory of Language* (2 volumes). Leiden: Brill.
- van Driem, George. 2005. 'Tibeto-Burman vs. Indo-Chinese: Implications for population geneticists, archaeologists and prehistorians', pp. 81-106 in Laurent Sagart, Roger Blench and Alicia Sanchez-Mazas, eds. *The Peopling of East Asia: Putting Together the Archaeology, Linguistics and Genetics*. London: Routledge Curzon.

- van Driem, George. 2007. 'Dzala and Dakpa form a coherent subgroup within East Bodish, and some related thoughts', pp. 71-84 in Roland Bielmeier and Felix Haller, eds., *Linguistics of the Himalayas and Beyond*. Berlin: Mouton.
- van Driem, George. 2011. 'Tibeto-Burman subgroups and historical grammar' *Himalayan Linguistics*, 10 (1): 30-39.
- van Driem, George. 2014. 'Trans-Himalayan', pp. 11-40 in Nathan Hill and Thomas Owen-Smith, eds., *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter.
- van Driem, George. 2017. 'Who are the Japanese, and where do the Japanese come from?', invited lecture presented at the workshop *Human Evolution in Eurasia elucidated through Genetics, Archaeology and Linguistics*, hosted by the National Institute of Genetics at Mishima, 17 March 2017.
- van Driem, George. 2019. 'Review: Lauren Gawne and Nathan Wayne Hill (eds.): *Evidential Systems of Tibetan Languages* (Trends in Linguistics Studies and Monographs, 302.) 472 pp. Berlin: Walter de Gruyter, 2017. £91. ISBN 978 3 11 046018 6. doi: 10.1017/S0041977X19000181', *Bulletin of the School of Oriental and African Studies*, 82 (1): 190-192.
- van Driem, George. 2021. *Ethnolinguistic Prehistory: The Peopling of the World from the Perspective of Language, Genes and Material Culture*. Leiden: Brill.
- van Driem, George. 2022. 'The creoloid origins of Chinese', *Bhāṣā: Journal of South Asian Linguistics, Philology and Grammatical Traditions*, 1 (2): 239-261.
- Dul'zon, Andrej Petrovič. 1964. *Očerki po grammatike ketskogo jazyka, I*. Tomsk: Izdatel'stvo Tomskogo Universiteta.
- Feng, John. 2023. 'China's plan to assimilate Tibet', *Newsweek*, 14 March 2023.
- Forrest, Robert Andrew Dermot. 1962. 'The linguistic position of Róng (Lepcha)', *Journal of the American Oriental Society*, 82 (3): 331-335.
- Frey, Henri-Nicolas. 1892. *L'annamite, mère des langues: Communauté d'origine des races celtiques, sémitiques, soudanaises et de l'Indo-Chine*. Paris: Librairie Hachette.
- Gāo Jīngyī [高晶一]. 2013. 'Huns and Xiongnu identified by Hungarian and Yeniseian shared etymologies', *Central Asiatic Journal*, 56: 41-48.
- Gāo Jīngyī [高晶一]. 2021. 'Huns and Xiongnu Identified by Hungarian and Yeniseian shared etymologies II', *Central Asiatic Journal*, 64 (1-2): 57-73.
- Gazi, Nurun Nahar, Rakesh Tamang., Vipin Kumar Singh, Ahmed Ferdous, Ajai Kumar Pathak, Mugdha Singh, Sharath Anugula, Pandichelvam Veeraiah, Subburaj Kadarkaraisamy, Brijesh Kumar Yadav, Alla G. Reddy, Deepa Selvi Rani, Syed Saleheen Qadri, Lalji Singh, Gyaneshwer Chaubey, Kumarasamy Thangaraj. 2013. 'Genetic structure of Tibeto-Burman populations of Bangladesh: Evaluating the gene flow along the sides of the Bay of Bengal', *Public Library of Science PLoS One*, 8 (10): e75064.
- Gerber, Pascal, and Selin Grollmann. 2018. 'What is Kiranti? – A critical account', *Bulletin of Chinese Linguistics*, 11: 88-152.
- Goropius Becanus, Johannes [Jan van Gorp van der Beke]. 1569. *Origines Antwerpianæ sive Cimmericum Becceselana novem libris complexa* (esp. Liber Quintus 'Indoscythica'). Antwerpen: Christophorus Plantinus.
- Hà Văn Thù. 2020. *The Formation Process of the Origin and Culture of the Viet People*. San Jose, California: Nhân Ảnh Publishing House.
- Hà Văn Thù. 2021. *Out of Vietnam Explore into the World*. San Jose, California: Nhân Ảnh Publishing House.
- Hà Văn Thù. 2021. *Rewriting Chinese History*. San Jose, California: Nhân Ảnh Publishing House.

- Haeckel, Ernst Heinrich. 1873. *Natürliche Schöpfungsgeschichte. Gemeinverständliche wissenschaftliche Vorträge über die Entwicklungslehre im Allgemeinen und diejenige von Darwin, Goethe und Lamarck im Besonderen*. Berlin: Georg Reimer.
- Haller, Felix Anton, and Chungda Haller. 2007. *Einführung in das moderne Zentraltibetische auf Basis des Dialektes von Shigatse, westliches Zentraltibet (Tsang)*. Bern: unveröffentlichtes Manuskript.
- Hazod, Guntram. 2007. 'The grave on the 'cool plain': On the identification of 'Tibet's first tomb' in Nga-ra-thang of 'Phyong-po', pp. 259–283 in Birgit Kellner, Helmut Krasser, Horst Lasic, Michael-Torsten Much and Helmut Tauscher, eds., *Pramāṇa-kīrtiḥ: Papers Dedicated to Ernst Steinkellner on the Occasion of his 70th Birthday* (Wiener Studien zur Tibetologie und Buddhismuskunde 70.1 & 70.2). Wien: Arbeitskreis für Tibetische und Buddhistische Studien.
- Hazod, Guntram. 2009. 'Part III: Imperial Central Tibet: An Annotated Cartographical Survey of its Territorial Divisions and Key Political Sites', pp. 161–232 in Brandon Dotson, ed., *The Old Tibetan Annals: An Annotated Translation of Tibet's First History*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Hazod, Guntram. 2013. 'The plundering of the Tibetan royal tombs: An analysis of the event in the context of the uprisings in Central Tibet of the 9th/10th century', pp. 85–115 in Christoph Cüppers, Robert Mayer and Michael Walter, eds., *Tibet after Empire: Culture, Society and Religion between 850-1000* (Proceedings of the Seminar held in Lumbini, Nepal, March 2011). Lumbini: Lumbini International Research Institute.
- Hill, Nathan Wayne. 2010. 'An overview of Old Tibetan synchronic phonology', *Transactions of the Philological Society*, 108 (2): 110–125.
- Hill, Nathan Wayne. 2019. *The Historical Phonology of Tibetan, Burmese and Chinese*. Cambridge: Cambridge University Press.
- Huáng Yùnzǐ and Lǐ Huī. 2017. 'Genetic and linguistic evidences for the Xiōngnú-Yenisseian hypothesis', *Man In India*, 97 (1): 95–110.
- Hyslop, Gwen. 2014. 'Waves across the Himalayas: On the typological characteristics and history of the Bodic subfamily of Tibeto-Burman', *Language and Linguistics Compass*, 8 (6): 243–270.
- International Campaign for Tibet (ICT). 2022. *Self-Immolation Fact Sheet* (as updated on 6 April 2022). Washington: International Campaign for Tibet.
- Iqbal, Zafar. 2018. 'Emerging history of Baltistan', in *Diverse and Adventurous History of Hunza*. Karachi: North Books.
- Jacques, Guillaume. 2009. 'Zhangzhung and Qiangic languages', *Senri Ethnological Studies*, 75: 121–130.
- Jacques, Guillaume. 2012. 'An internal reconstruction of Tibetan stem alternations', *Transactions of the Philological Society*, 110 (2) 212–224.
- Jagou, Fabienne. 2019. 'The dispute between Sichuan and Xikang over the Tibetan Kingdom of Trokyap (1930s–1940s)', pp. 337–362 in Stéphane Gros, ed., *Frontier Tibet: Patterns of Change in the Sino-Tibetan Borderlands*. Amsterdam: Amsterdam University Press.
- Jaxontov, Sergej Evgenievič. 1986. 'Jazyki severnyx sosedej Kitaja v pervom tysjačeljetija do n.č.', pp. 123–134, Vol. 1 in *Istoriko-kul'turnnye kontakty narodov altajskoj jazykovoj obščnosti: Tezisy dokladov XXIV sessii postojannoju međunarodnoju altaističeskoju konferencii (PIAC)*. Moscow: Institute of Oriental Studies.
- Jeong, Choongwon, Ke Wang, Shevan Wilkin, William Timothy Treal Taylor, Bryan K. Miller, Jan H. Bemann, Raphaela Stahl, Chelsea Chiovelli, Florian Knolle, Sodnom Ulziibayar, Dorjpurev Khatanbaatar, Diimaajav Erdenebaatar, Ulambayar Erdenebat, Ayudai Ochir, Ganbold Ankhsanaa, Chuluunkhuu Vanchigdash, Battu-

- ga Ochir, Chuluunbat Munkhbayar, Dashzeveg Tumen, Alexey Kovalev, Nikolay Kradin, Bilikto A. Bazarov, Denis A. Miyagashev, Prokopy B. Kononov, Elena Zhambaltarova, Alicia Ventresca Miller, Wolfgang Haak, Stephan Schiffels, Johannes Krause, Nicole Boivin, Myagmar Erdene, Jessica Hendy and Christina Warinner. 2020. 'A dynamic 6,000-year genetic history of Eurasia's eastern steppe', *Cell*, 183: 890–904.
- Kapstein, Matthew Tom. 2006. *The Tibetans*. Oxford: Blackwell.
- Kelley, Liam C. 2020. 'The centrality of "fringe history": Diaspora, the internet and a new version of Vietnamese prehistory', *International Journal of Asia Pacific Studies*, 16 (1): 71–104.
- von Klapproth, Julius. 1823. *Asia Polyglotta* [the first volume includes the separate instalment « Leben des Budd'a nach mongolischen Nachrichten » on pp. 121–144 following p. 384, and the rarer second volume contains an atlas and diagrams]. Paris: J. M. Eberhart.
- Komissaruk, Ekaterina L'vovna. 2021. *Социолингвистическая ситуация в Ладакхе (проблемы становления Ладакхи как письменного языка)* ['Sociolinguistic situation in Ladakh (Development of the written Ladakhi language)']. Moscow: National Research University, Higher School of Economics, Doctoral dissertation, 12 November 2021.
- Krejnovič, Eruxim Abramovič. 1968. *Glagol ketskogo jazyka*. Leningrad: Izdatel'stvo 'Nauka'.
- Kriz, Karel, and Guntram Hazod. 2020. *The Burial Mound Sites of Imperial Central Tibet*. Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Lee, Juhyeon, Bryan K. Miller, Jamsranjav Bayarsaikhan, Erik Johannesson, Alicia Ventresca Miller, Christina Warinner and Choongwon Jeong. 2023. 'Genetic population structure of the Xiongnu Empire at imperial and local scales', *Science Advances*, 9, article № eadf3904.
- Mainwaring, Colonel George Byres. 1898. *Dictionary of the Lepcha Language* (revised and completed by Albert Grünwedel). Calcutta: Baptist Mission Press.
- Martin, Dan. 2010. 'Zhangzhung dictionary', *Revue d'Études Tibétaines*, 19: 3–253.
- Martin, Dan. 2013. 'Knowing Zhang-zhung: The very idea', *Journal of the International Association for Bon Research*, 1: 175–197.
- Martin, Dan. 2016. *mDzod-phugs* (version of 1 January 2016). Manuscript in progress.
- Martinez, Philippe Antoine. 2021. *A Corpus-Based Account of Morphosyntactic Evidence in Discourse in Chhitkul-Rākchham*. London: Ph.D. dissertation, School of Oriental and African Studies of the University of London.
- McGranahan, Carole. 2019. 'Chinese settler colonialism: Empire and life in the Tibetan borderlands', pp. 517–540 in Stéphane Gros, ed., *Frontier Tibet: Patterns of Change in the Sino-Tibetan Borderlands*. Amsterdam: Amsterdam University Press.
- Nado, Lopon (*sLob-dpon gNag-mdog*) 'Löbö 'Nâdo. 1982. 'The development of language in Bhutan', *Journal of the International Association of Buddhist Studies*, 5 (2): 95–100.
- Namgyal, Kunzang, and George van Driem. 2022. 'Drenjongke phonology, Phonological Drenjongke and Roman Drenjongke', *Bulletin of Tibetology*, 53 (1): 9–30.
- Office of International Religious Freedom. 2022. *2021 Report on International Religious Freedom: Pakistan*. Washington: Department of State.
- Office of the High Commissioner for Human Rights. 2023. 'China: UN experts alarmed by separation of 1 million Tibetan children from families and forced assimilation at residential schools', press release of 6 February 2023 by the United Nations Office of the High Commissioner for Human Rights, Geneva.

- Oppenheimer, Stephen. 2012. 'Out-of-Africa, the peopling of continents and islands: Tracing uniparental gene trees across the map', *Philosophical Transactions of the Royal Society*, B 367: 770–784.
- Orlandi, Georg. 2021. 'Once again on the history and validity of the Sino-Tibetan bifurcate model', *Journal of Language Relationship — Вопросы языкового родства*, 19 (4): 263–292.
- Peiros, Iliia. 1998. *Comparative Linguistics in Southeast Asia*. Canberra: Pacific Linguistics.
- Peiros, Iliia. 2004. 'Linguistic prehistory of Southeast Asia: what linguists know and what they believe they know', paper presented at the conference 'Human migrations in continental East Asia and Taiwan: Genetic, linguistic and archaeological evidence', Université de Genève, 10 June 2004.
- Pompeo, Michael Richard. 2021. *Determination of the Secretary of State on Atrocities in Xinjiang: Press Statement 19 January 2021*. Washington: Department of State.
- Pulleyblank, Edwin George. 1962. 'The consonantal system of Old Chinese', *Asia Major*, IX (1): 58–144, (2): 206–265.
- Pulleyblank, Edwin George. 1966. 'Chinese and Indo-Europeans', *Journal of the Royal Asiatic Society*, 98 (2): 9–39.
- Pulleyblank, Edwin George. 1983. 'The Chinese and their neighbors in prehistoric and early historic times', pp. 411–466 in David N. Keightley, ed., *The Origins of Chinese Civilization*. Berkeley: University of California Press.
- Pulleyblank, Edwin George. 2000b. 'The Hsiung-nu', pp. 52–75 in Hans Robert Roemer and Wolfgang-Ekkehard Scharlipp, eds., *History of the Turkic Peoples in the pre-Islamic Period*. Berlin: Klaus Schwarz Verlag.
- Reddy, B. Mohan, B.T. Langstieh, Vikrant Kumar, T. Nagaraja, A.N.S. Reddy, M. Aruna, Kumarasamy Thangaraj, A.G. Reddy and Lalji Singh. 2007. 'Austro-Asiatic tribes of Northeast India provide hitherto missing genetic link between South and Southeast Asia', *Public Library of Science PLoS One*, 2 (11): e1141.
- Repucci, Sarah, and Amy Slipowitz. 2022. *Freedom in the World 2022: The Global Expansion of Authoritarian Rule*. Washington: Freedom House.
- Richardson, Hugh Edward. 1962. *Tibet and Its History*. London: Oxford University Press.
- Raza Ghalib, Muhammad. 2015. 'Revival of *yige* in Baltistan', *Stawa*, 2 (12): 21.
- Richardson, Hugh Edward. 1963. 'Early burial grounds in Tibet and Tibetan decorative art of the VIIIth and IXth centuries', *Central Asiatic Journal*, VIII: 73–92.
- Robertson, Matthew P., and Jacob Lavee. 2022. 'Execution by organ procurement: Breaching the dead donor rule in China', *American Journal of Transplantation*, 2 (7): 1804–1812.
- Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill and Johann-Mattis List. 2019. 'Dated language phylogenies shed light on the ancestry of Sino-Tibetan', *Proceedings of the National Academy of Sciences*, 116 (21): 10317–10322.
- Sahoo, Sanghamitra, Anamika Singh, G. Himabindu, Jheeman Banerjee, T. Sitalaximi, Sonali Gaikwad, R. Trivedi, Phillip Endicott, Toomas Kivisild, Mait Metspalu, Richard Villems and V. K. Kashyap. 2006. 'A prehistory of Indian Y chromosomes: Evaluating demic diffusion scenarios', *Proceedings of the National Academy of Sciences*, 103 (4): 843–848.
- van Schaik, Sam Julius. 2011. *Tibet: A History*. New Haven: Yale University Press.
- Sclater, Philip Lutley. 1864. 'The mammals of Madagascar', *Quarterly Journal of Science*, 1 (2): 213–219.

- Shafer, Robert. 1941. 'The linguistic relationship of Mru: Traces of a lost Tibeto-Burmese language?', *Journal of the Burma Research Society*, XXXI (II): 58-79.
- Shafer, Robert. 1944. 'Problems in Sino-Tibetan phonetics', *Journal of the American Oriental Society*, 64: 137-143.
- Shafer, Robert. 1950. 'Studies in the morphology of Bodic verbs', *Bulletin of the School of Oriental and African Studies*, XIII (3): 702-724.
- Shafer, Robert. 1951. 'Studies in the morphology of Bodic verbs', *Bulletin of the School of Oriental and African Studies*, XIII (4): 1017-1031.
- Shafer, Robert. 1953a. 'Classification of the northern-most Naga languages', *Journal of the Bihar Research Society*, XXXIX (3): 225-264.
- Shafer, Robert. 1953b. 'East Himalayish', *Bulletin of the School of Oriental and African Studies*, XV (2): 356-374.
- Shafer, Robert. 1954. 'The linguistic position of Dwags', *Oriens, Zeitschrift der Internationalen Gesellschaft für Orientforschung*, 7: 348-356.
- Shafer, Robert. 1955. 'Classification of the Sino-Tibetan languages', *Word, Journal of the Linguistic Circle of New York*, 11: 94-111.
- Shafer, Robert. 1963. 'Shafer, Robert', *Orbis, Bulletin International de Documentation Linguistique*, XII (1): 350-353.
- Shafer, Robert. 1966. *Introduction to Sino-Tibetan, Part I*. Wiesbaden: Otto Harrassowitz.
- Shafer, Robert. 1967. *Introduction to Sino-Tibetan, Part II*. Wiesbaden: Otto Harrassowitz.
- Shafer, Robert. 1968. *Introduction to Sino-Tibetan, Part III*. Wiesbaden: Otto Harrassowitz.
- Shafer, Robert. 1974. *Introduction to Sino-Tibetan, Part IV*. Wiesbaden: Otto Harrassowitz.
- Siiger, Halfdan, and Jørgen Rischel. 1967. *The Lepchas: Culture and Religion of a Himalayan People* (Parts I and II). Copenhagen: National Museum of Denmark.
- Snowden, Frank Martin, Jr. 1970. *Blacks in Antiquity: Ethiopians in Graeco-Roman Experience*. Cambridge, Massachusetts: Harvard University Press.
- Snowden, Frank Martin, Jr. 1989. 'Bernal's "blacks", Herodotus and other Classical evidence', *Arethusa*, 22 (3): 83-95.
- Sørensen, Per Kjeld, and Guntram Hazod. 2007. *Rulers of the Celestial Plain: Ecclesiastic and Secular Hegemony in Mediaeval Tibet: A Study of Tshal Gung-thang* (two volumes). Vienna: Verlag der Österreichischen Akademie der Wissenschaften.
- Sprigg, Richard Keith. 1982. 'The Lepcha language and three hundred years of Tibetan influence in Sikkim', *Journal of the Asiatic Society*, XXIV: 16-31.
- Starostin, Sergej Anatol'evič. 1989. *Rekonstrukcija drevnekitajskoj fonologičeskoj sistemy*. Moscow: Izdatel'stvo 'Nauka'.
- Starostin, Sergej Anatol'evič. 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, Sèvres, 14 octobre 1994.
- Stein, Rolf Alfred. 1959. *Les tribus anciennes des marches sino-tibétaines: Légendes, classifications et histoire*. Paris: Imprimerie Nationale.
- Stein, Rolf Alfred. 1962. *La civilisation tibétaine*. Paris: Dunod.
- Takeuchi Tsugohito and Nishida Ai. 2009. 'The present state of deciphering Old Zhangzhung', *Senri Ethnological Studies*, 75: 151-165.
- Tamsang, Kharpoo (i.e. Kharpú). 1994. *The Lepcha-English Encyclopaedic Dictionary*. Kalimpong: Mrs. Mayel Clymit Tamsang.
- Tournadre Nicolas. 2014a. *Le prisme des langues*. Paris: L'Asiathèque.

- Tournadre Nicolas. 2014b. 'The Tibetic languages and their classification', pp. 105-130 in Nathan Hill and Thomas Owen-Smith, eds., *Trans-Himalayan Linguistics*. Berlin: Mouton de Gruyter.
- Tshering, Karma, and George van Driem. 2019. *The Grammar of Dzongkha, Revised and Expanded, with a Guide to Roman Dzongkha and to Phonological Dzongkha*. Santa Barbara, California: Himalayan Linguistics.
- Tsomo, Tsering, et al. 2022. *Sucked Our Marrow: Tibetan Language and Education Rights under Xi Jinping (Thematic Report 2022)*. Dharamsala: Tibetan Centre for Human Rights and Democracy.
- Tucci, Giuseppe. 1950. *Tombs of the Tibetan Kings* (Serie Orientale Roma, 1). Rome: Istituto Italiano per il Medio ed Estremo Oriente.
- Uray, Géza. 1960. 'The four horns of Tibet according to the royal annals', *Acta Orientalia Academiae Scientiarum Hungaricae*, 10 (1): 31-57.
- van Walt van Praag, Michael Christian. 1987. *The Status of Tibet: History, Rights and Prospects in International Law*. Boulder, Colorado: Westview Press.
- van Walt van Praag, Michael Christian, and Miek Boltjes. 2020. *Tibet Brief 20/20*. Denver, Colorado: Outskirts Press.
- Voltaire, alias François-Marie Arouet. 1803 [An XI]. 'Catalogue... des écrivains français qui ont paru dans le siècle de Louis XIV', pp. 185-324, Vol. v in *Siecles de Louis XIV, et de Louis V, par Voltaire*. Paris: De l'imprimerie et de la fonderie stéréotypes de Pierre Didot l'aîné, et de Firmin Didot.
- Vovin, Alexander. 2000. 'Did the Xiong-nu speak a Yeniseian language?', *Central Asiatic Journal*, 44 (1): 87-104.
- Vovin, Alexander, Edward Vajda and Étienne de la Vassière. 2016. 'Who where the *Kjet 羯 and what language did they speak?', *Journal Asiatique*, 304 (1): 125-144.
- Warner, Cameron David. 2022. "'Do what you think about': Fashionable responses to the end of Tibet', pp. 155-179 in Haidy Geismar, Ton Otto & Cameron David Warner, eds., *Impermanence: Exploring Continuous Change across Cultures*. London: University College London Press.
- Winter, Werner. 1986. 'Aus der Arbeit des Linguistic Survey of Nepal'. pp. 453-459 in Bernhard Kölver and Siegfried Lienhard, eds., *Formen kulturellen Wandels und andere Beiträge zur Erforschung des Himālaya: Colloquium des Schwerpunktes Nepal, Heidelberg, 1.-4. Februar 1984 (Nepalica 2)*. Sankt Augustin: Vereinigung für Geschichtswissenschaft Hochasiens Wissenschaftsverlag.
- Zhang, Menghan, Shi Yan, Wuyan Pan and Li Jin. 2019. 'Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic', *Nature*, 568: 112-115.
- Zhao, Yong-Bin, Hong-Jie Li, Sheng-Nan Li, Chang-Chun Yu, Shi-Zhu Gao, Zhi Xu, Li Jin, Hong Zhu and Hui Zhou. 2011. 'Ancient DNA evidence supports the contribution of Di-Qiang people to the Hân Chinese gene pool', *American Journal of Physical Anthropology*, 144: 258-268.
- Zhao, Yong-Bin, Ye Zhang, Hong-Jie Li, Ying-Qiu Cui, Hong Zhu and Hui Zhou. 2014. 'Ancient DNA evidence reveals that the Y-chromosomal haplogroup Q1a1 admixed into the Hân Chinese 3,000 years ago', *American Journal of Human Biology*, 26: 813-821.
- Zhao, Yong-Bin, Ye Zhang, Quan-Chao Zhang, Hong-Jie Li, Ying-Qiu Cui, Zhi Xu, Li Jin, Hui Zhou and Hong Zhu. 2015. 'Ancient DNA reveals that the genetic structure of the northern Hân Chinese was shaped prior to 3,000 years ago', *Public Library of Science PLoS One*, 10 (5): e0125676.