

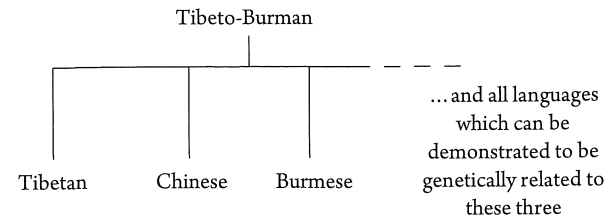
## The Naga Language Groups within the Tibeto-Burman Language Family

George van Driem

The Nagas speak languages of the Tibeto-Burman family. Yet, according to our present state of knowledge, the >Naga languages< do not constitute a single genetic subgroup within Tibeto-Burman. What defines the Nagas best is perhaps just the label *Nāgā*, which was once applied indiscriminately by Indo-Aryan colonists to all scantily clad tribes speaking Tibeto-Burman languages in the northeast of the Subcontinent. At any rate, the name *Nāgā*, ultimately derived from Sanskrit *nagná* >naked<, originated as a titular label, because the term denoted a sect of Shaivite *sādhus* whose most salient trait to the eyes of the lay observer was that they went through life unclad. The Tibeto-Burman tribes labelled *Nāgā* in the northeast, though scantily clad, were of course not Hindu at all. The diverse Naga language communities observed their own indigenous religions and represented a lineage of cultural traditions entirely distinct from the Hindu, Jain and Buddhist traditions of India proper. Other etymologies have been proposed relating the ethnonym *Nāgā* to words or expressions in one or other language of the many diverse Naga languages. One such etymology proposes that the term derives from an expression denoting piercing, a practice prevalent amongst some of the Naga tribes. Whether or not any such alternative explanations in fact represent the genuine etymology of the name, or just reflect folk etymologies, they evidently take their inspiration from popular perceptions by outsiders of salient features of Naga material culture.

Ethnically, many Tibeto-Burman tribes of the northeast have been called *Nāgā* in the past or have been labelled as >Naga< in scholarly literature who are no longer usually covered by the modern more restricted sense of the term today. Linguistically, even today's >Naga languages< do not represent a single coherent branch of the family, but constitute several distinct branches of Tibeto-Burman. This essay aims (1) to give an idea of the linguistic position of these languages within the family to which they belong, (2) to provide a relatively comprehensive list of names and localities as a directory for consultation by scholars and interested laymen who wish to make their way through the jungle of names and alternative appellations that confront any interested reader of ethnographical and linguistic literature on the Nagas and their closest Tibeto-Burman neighbours in northeastern India, and (3) to address the issue of the provenance of the various Tibeto-Burman language communities designated collectively as Naga.

The Tibeto-Burman family vies with Indo-European for the title of the language family with the most speakers in the world. Numerically, however, most speakers of Tibeto-Burman languages are represented by a single branch. The Sinitic languages Mandarin, Hakka, Cantonese and the other Chinese >dialects< account for most speakers as a historical consequence of the southward expansion of Hân language and culture from northeastern China from the 3rd century BC onwards. Yet Sinitic is just one of many branches. Most of the branches of the Tibeto-Burman language

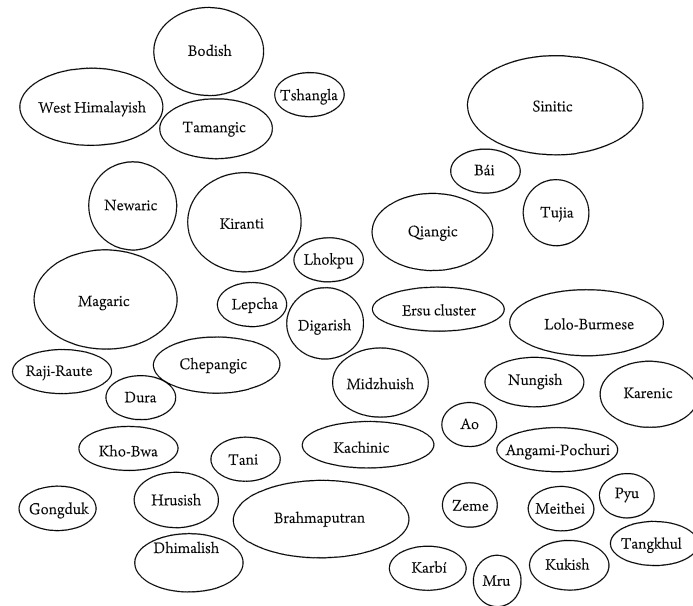


△ Fig. 1. One of the language families identified by Julius Heinrich von Klaproth in his polyphyletic view of Asian linguistic stocks (1823a, 1823b). He explicitly excluded languages today known to be Kra-Dai or Daic (e.g. Thai, Lao, Shan) and known to be Austroasiatic (e.g. Mon, Vietnamese, Nicobarese, Khmer).

family are represented solely in India. In fact, the linguistic and prehistorical centre of diversity of the language phylum lies decidedly within the Indian subcontinent.

The Tibeto-Burman family of languages was first identified in Paris by the German scholar Julius von Klaproth in his *Asia Polyglotta* in 1823. The first explicitly polyphyletic view of Asian linguistic stocks had been presented over a century earlier, in 1692, by Nicolaes Witsen, former mayor of Amsterdam. Yet by the beginning of the nineteenth century, enough language documentation had accumulated in Europe that the well-travelled and knowledgeable Klaproth was able to identify and distinguish twenty-three Asian language families based on his systematic comparison of lexical roots. Some of his families have been augmented, diminished or redefined, but today his Tibeto-Burman model remains the most well-informed model and also still represents the most agnostic and thus inherently least controversial theory of the genetic relationship between the diverse Naga languages and languages such as Burmese, Chinese, Tibetan, Sherpa, Limbu, Newari and other languages of the same linguistic stock.

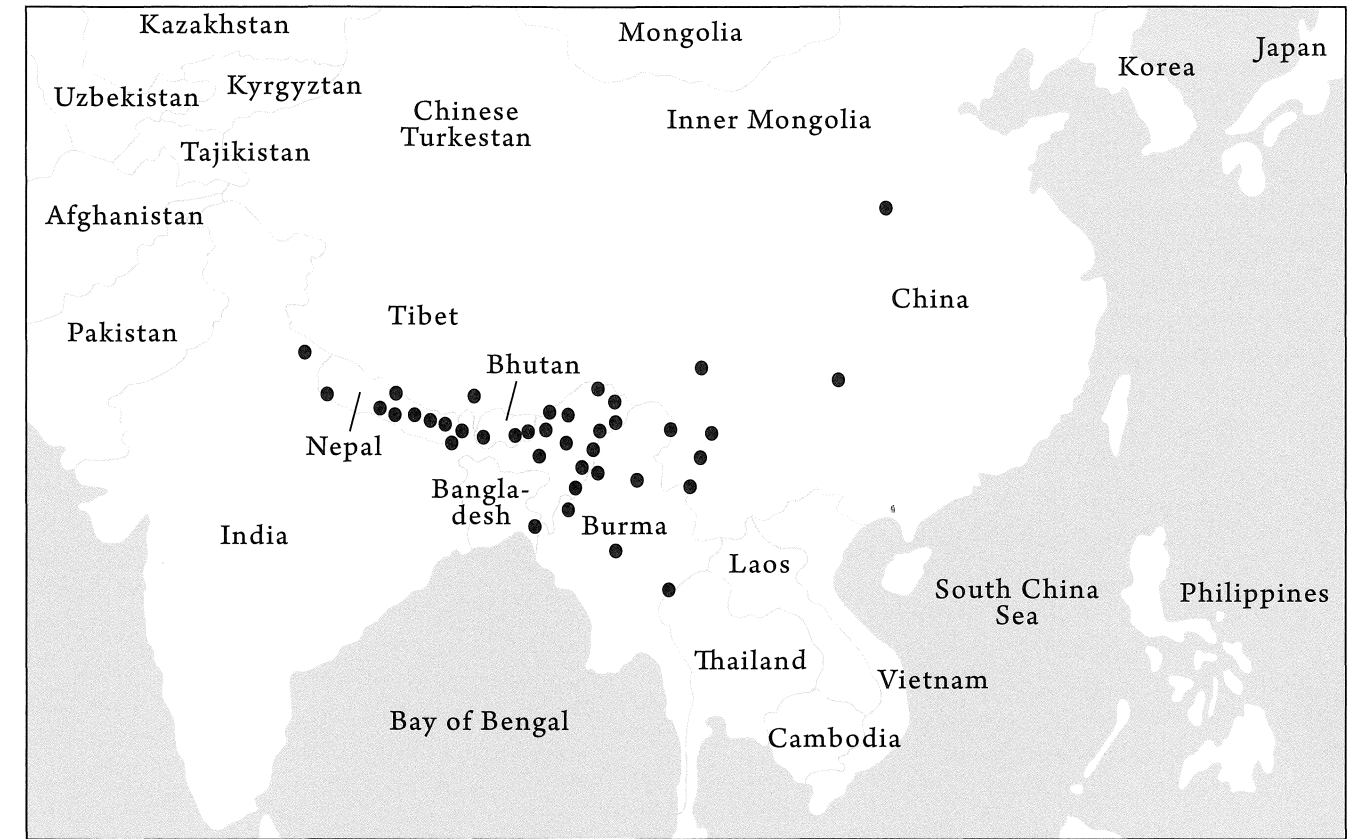
We know more now than Klaproth did. Many previously unknown Tibeto-Burman languages and subgroups have been identified since 1823. In 2001 in Cambridge, I introduced the metaphor of fallen leaves illustrated in Fig. 2 (van Driem 2001). The model attempts to identify all the constituent branches of the family and draw the focus of



△ Fig. 2. Major branches of the Tibeto-Burman language family (van Driem 2001).

attention back to the centre of Tibeto-Burman linguistic diversity, which lies in the eastern Himalayas and the Indo-Burmese borderlands. The patch of fallen leaves on the forest floor provides a more informative framework than a false tree, such as the misleading and now defunct Indo-Chinese or >Sino-Tibetan< model (van Driem 2003).

The metaphor of fallen leaves implies the existence of a tree. Yet we cannot lift our gaze from the forest floor to see the tree because we cannot look directly into the past. Instead, historical comparative work will hopefully enable us to see the shadows which the branches cast between the leaves on the forest floor. Whether a language family appears to be more rake-like or more tree-like is often a function of the state of the art in historical comparative linguistics rather than a statement about linguistic phylogeny. It is relevant to note, however, that with the inexorable progress of Indo-European linguistic studies, even the twelve branches of this most well-studied language family, once depicted in the pleasing shape of a branching oak, have gradually assumed a more rake-like appearance and so come closer to the fallen leaves model.



Each branch of the family shown in Fig. 1 contains between one and over forty completely different languages. The representation is schematic and neutral with respect to different higher-order subgrouping hypotheses about the branching of the family, e.g. Sino-Bodic, Newaric, Brahmaputran. The subgroups Sinitic, Bái, Tǔjiā, Qiāngic and the Ērsǔ cluster lie within the present-day political borders of China. The Bodish, Tshangla and Lolo-Burmese branches are represented both inside China and on the Indian subcontinent. All other groups lie within India, Nepal, Bhutan, Burma and Bangladesh.

The key to understanding the prehistory of this language family therefore lies in India, and the centre of diversity lies specifically in the northeastern corner of the Subcontinent, where we also find the Naga languages. Below are listed the groups in India which are called Naga, some neighbouring groups that have been loosely categorised as

△ Fig. 3. Geographical distribution of the major branches of Tibeto-Burman. Each dot represents a major subgroup, not a single language. (van Driem 2002).

>Nagas< in the past, and other groups of northeastern India that sometimes tend to be confused with the Nagas. The most common alternative names for the various language communities are provided in order to aid the unravelling of the complex ethnic and linguistic nomenclature used locally and in the specialist literature. The geographical locations of the language communities are also indicated.

The groups numbered here from 1 to 8 live to the west or to the north of the Nagas. Group 9 comprises the communities usually called >Northern Nagas<, even though Northern Nagas are linguistically more closely affiliated to Brahmaputran, i.e. Group 4, and perhaps also to Kachinic, i.e. Group 17. Groups 10 to 12 are the linguistic groups

narrowly defined as Nagas, comprising the languages of the Ao branch, the Angami-Pochuri branch and the Zeme branch. Groups 13 to 16 represent the immediate southern and eastern neighbours of the Nagas. Groups 18 to 38 are branches of Tibeto-Burman languages spoken outside of northeastern India. More information on all of these groups can be found in my 2001 handbook.

#### Tibeto-Burman subgroups and language communities

- 1 Lepcha – Sikkim and Darjeeling district
  - 1a Lepcha (a.k.a. Rong)
- 2 Hrusish – in Kameng district in eastern Arunachal Pradesh
  - 2a Dhímmai (a.k.a. Miji) – between Dirang, Seppa and Rieng
  - 2b Levai (a.k.a. Bangru) – northeast of Rieng, northwest of Yomtám
  - 2c Hruso (a.k.a. Aka) – southwest of Seppa, east of Shergaon
- 3 Kho-Bwa – in Kameng district in eastern Arunachal Pradesh
  - 3a Khowa (a.k.a. Bugun) – around Bomdila, esp. in Wanghoo and Singchung villages
  - 3b Sulung (a.k.a. Puroit, Puroik) – north of Rieng along the upper reaches of the Par river
  - 3c Lishpa (pass themselves off as >Monpa<) – around Dirang
  - 3d Sherdukpen – mainly in and around Shergaon, Rupa and Jigaon
- 4 Brahmaputran (a.k.a. Kachári-Koch or Bodo-Koch)
  - 4a Koch
    - 4a-1 Atong (a.k.a. At'ong, Atong >Garo<) — Somasvara and Baghmara in the lower Garo Hills
    - 4a-2 Ruga (a.k.a. Ruga >Garo<, Rugha) – downhill from the Atong
    - 4a-3 Rabha – skirting the north of the Garo Hills, on both sides of the Brahmaputra just before it bends south
    - 4a-4 Pani Koch – western fringe of the Meghalaya below the Garo Hills
    - 4a-5 Rajbangsi (no longer speak any Tibeto-Burman language but a dialect of Bengali, historically Rajbangsi are Koch assimilated linguistically in the Kamarupan period) – Siliguri, Jalpaiguri, Alipur *duar* (trade route) and other former Bhutanese *duar*
  - 4b Bodo-Garo
    - 4b-1 Mech – Jalpaiguri district, Bhutanese *duar*
    - 4b-2 Garo – Garo Hills
    - 4b-3 Bodo – Darrang

- 4b-4 Tiwa (a.k.a. Lalung) – in the Karbi Anglong or Mikir Hills around Umswai
- 4b-5 Dimasa (a.k.a. >Bodo<) – northern Cachar Hills
- 4b-6 Hojai (a.k.a. >Bodo<) – northern Cachar Hills
- 4b-7 Kachari (a.k.a. >Bodo<) – Goalpara
- 4b-8 Kokborok – Tripura
- 4c Chutiya (a.k.a. Deori Chutiya) – in Lakhimpur and Sibsagar districts
- 5 Tani (a.k.a. >Abor-Miri-Dafla<) – in Arunachal Pradesh
  - 5a Western Tani (a.k.a. >Miri-Dafla<, Nishi group) – west of the Subansiri river
    - 5a-1 Apatani – between the Panior and Kamla rivers
    - 5a-2 Bengni (a.k.a. Bangni, >Western Dafla<)
    - 5a-3 Nah (a.k.a. Chendar, Haphi, Hari and Tisi) – Upper Subansiri district in seven villages of Taksing administrative circle
    - 5a-4 Nishing (a.k.a. Nishi, >Eastern Dafla<) – Kameng valley in East Kameng district
    - 5a-5 Nyisu – possibly a sub-branch of Nishi
    - 5a-6 Yano – possibly a sub-branch of Nishi
    - 5a-7 Tagin – northeastern Subansiri district, inc. Denekoli and Taliha
    - 5a-8 Hill Miri (a.k.a. Sarak) — enclaves between Nishi and Gallong territory
    - 5a-9 Gallong (a.k.a. Galo) – southern half of West Siang district, west of the Siyom river
    - 5a-10 Pailibo (a.k.a. Libo) – north of Gallong territory along banks of the Siyom or >Yomgo<, esp. Tato and Payum circles of West Siang district
    - 5a-11 Ramo – Mechukha subdivision of West Siang district to the northwest of the Pailibo area, esp. between Machukha and Tato
    - 5a-12 Bokar – Monigong Circle of Machukha subdivision in West Siang district, esp. Gesing, Pangri, Yangrang, Taihiyong, Rote, Pidi, Ruying, Kate, Luto, Ramni, Hemi or Mote, Pote, Karle, Monigong, Ingo, Pulum, Simegong, Papigro, Tadadege, Lapugora, Yorkongdo and Namasiba

- 5b Eastern Tani (a.k.a. >Abor<, Adi, Padam-Minyong group) – east of the Subansiri
  - 5b-1 Damu – perhaps near the Subansiri, precise whereabouts unknown
  - 5b-2 Bori – along the Siyom and Sike in an area enclosed by the Luyor hills on the east, the Piri hills on the west and on the north by the closing together of these two ranges, esp. Payum, Dupu, Yiyo, Pame, Gasheng, Paying, Gatte, Gameng, Bogu and Mega villages
  - 5b-3 Mishing (a.k.a. Plains Miri) – East Siang at Oyan, Mer and Namsing
  - 5b-4 Padam (a.k.a. Bor Abor, i.e. >Great Abor<) – the Dibang, Siang and Yamne valleys in East Siang, from the Assam border in the south to the Sidip river in the north, and in western Lohit between the Siang and Sisiri or Sikang river, esp. Anpum, Bijari, Bolung, Bomjir, Dambuk, Kangkong, Kappang, Tapat and Poblung.
  - 5b-5 Minyong (a.k.a. Adi Minyong) — west bank of the lower Siang river, downstream of the Bori and Karko language communities and to the east of the Gallongs, esp. Ledum, Sido, Mirem, Yagrung, Rengin, Rotung, Kebang, Yemsing, Pangin, Yeksing, Boleng, Dosing, Pankang, Parong, Yebuk, Sitang, Riga, Riu, Komsing, Rumgong, Pessing, Molom, Moput, Lorging and Jomlo
  - 5b-6 Shimong (a.k.a. Simong) – left bank of the Siang in and around Yingkiang, esp. Ngaming, Jido, Anging, Singiang, Palin, Likor, Pugging, Gete, Shimong and Gobuk villages
  - 5b-7 Pasi — Pasighat area in the foothills, esp. Gine, Kelek, Roing, Monku, Balek, Rasam and Tigra
  - 5b-8 Panggi – lower Yamne valley above the confluence of the Yamne and the Siang, esp. Geku, Sumsing, Sibum, Jeru and Pongging
  - 5b-9 Tangam – northernmost Siang district in the villages of Kuging, Ngering and Mayum

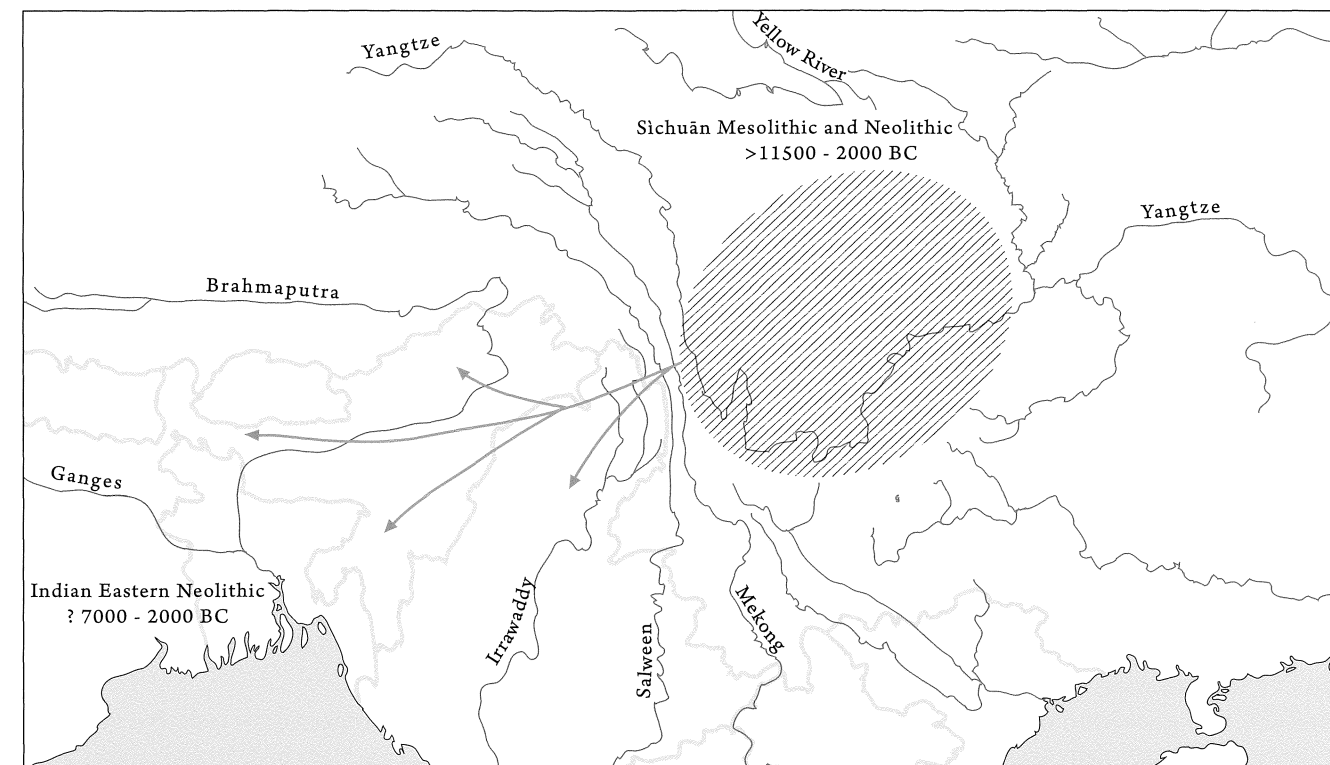
- 5b-10 Karko (a.k.a. Karka) – across the river from the Shimong to the west, mainly in Karko village and nearby Ramsing and Gosang
- 5b-11 Ashing (a.k.a. Asing) – headwaters of the Siang near the Tibetan border, from Ramsing in the south to Tuting in the north, esp. Pango, Bomdo, Ninging, Minging and Mosing.
- 5c Milang
- 5c-1 Milang – three villages of Milang, Dalbing and Pekimodi in the upper Yamne valley in Mariyang subdivision of East Siang district
- 6 Digarish (a.k.a. >Northern Mishmi<) – Lohit district of Arunachal Pradesh
- 6a Idu (a.k.a. Chulikata >cropped hair< Mishmi, >Bebejia Mishmis<) – in the Dibang, Mithun and Dri valleys
- 6b Taraon (a.k.a. Digaro, Digaru Mishmi) – in the Lohit valley, between the Delei and Lati rivers in the east, the Kharem in the south and the Digaru in the west
- 7 Midzhuish (a.k.a. >Southern Mishmi<) – Lohit district of Arunachal Pradesh
- 7a Kaman (a.k.a. Miju, Mijhu Mishmi) – upper reaches of the Lohit on both banks of the river around Parsuram Kund
- 7b Zaiwa (a.k.a. Zakhring and Meyor dialects) – less than two hundred members of the Zakhring and Meyor clans in the vicinity of Walong
- 8 Karbí (Mikir)
- 8a Karbí proper – the territory all around Hojai in Assam
- 8b Amri Karbí – lower Meghalaya hills northeast of Nongpoh
- 9 Konyak a.k.a. Northern Naga
- 9a Tangsa-Nocte cluster
- 9a-1 Tangsa – in Changlang and Miao subdivisions of Tirap district in Arunachal Pradesh around Tirap and also in neighbouring parts of Burma: including Moklum, Longcang (Lungchang), Longphi, Hawi (Have, Havi), Jogli (Jugli, Yogli), Tikhak, Longri, Rangpang, Moshang, Rangpang, etc.
- 9a-2 Nocte – in Arunachal Pradesh between Pong and Tirap
- 9b Konyak-Wancho cluster
- 9b-1 Wancho – west of Tirap district in Arunachal Pradesh around Pong
- 9b-2 Konyak – the northern tip of Nagaland
- 9b-3 Phom – northern tip of Nagaland, north of Tuensang
- 9b-4 Khiamngan – northern tip of Nagaland
- 9b-5 Chang – northern Nagaland, in and around Tuensang
- 9c Unclassified Northern Naga
- 9c-1 Kuwa – Burma east of Tirap district
- 9c-2 Haimi – Burma east of Tirap district
- 9c-3 Ponyo – Burma around Lahe, east of Tuensang in Nagaland
- 9c-4 Welam – Burma around Lahe, east of Tuensang in Nagaland
- 9c-5 Nokaw – Burma around Lahe, east of Tuensang in Nagaland
- 9c-6 Htangan – Burma around Lahe, east of Tuensang in Nagaland
- 10 Ao branch – central Nagaland, north and east of Wokha
- 10a Yacham
- 10b Tengsa
- 10c Ao Chungli
- 10d Ao Mongsen
- 10e Lotha (a.k.a. Lhota)
- 10f Sangtam (a.k.a. Thukumi)
- 10g Yimchungrü (a.k.a. Yachumi)
- 10h Nruanghmei
- 11 Angami-Pochuri branch – southern Nagaland around Kohima, neighbouring portions of Burma and the northern fringe of Manipur
- 11a Angami
- 11b Chakri (a.k.a. Chokri)
- 11c. Sema
- 11d Rengma (a.k.a. Anzang)
- 11e Pochuri (a.k.a. >Southern Sangtam<, >Eastern Rengma<)
- 11f Kezhama (a.k.a. Kezha)
- 11g Senkadong
- 11h Mao
- 11i Ntenyi
- 11j Maluri
- 12 Zeme branch – the southwestern tip of Nagaland and the northwestern corner of Manipur, starting from Tamenglong and moving to the north and north-northeast
- 12a Mzieme
- 12b Khoirao
- 12c Maram
- 12d Puiron
- 12e Zeme (a.k.a. Empeo Naga, Kacha Naga, Kochu Naga)
- 12f Nruanghmei (a.k.a. Rongmai, Kabui)
- 12g Liangmai (a.k.a. Kwoireng)
- 13 Tangkhul branch – the northeastern corner of Manipur around Ukhrul and in neighbouring portions of Burma
- 13a Tangkhul
- 13b Maring
- 14 Meithei (a.k.a. Manipuri) – in Manipur
- 15 Kukish (a.k.a. Mizo-Kukì-Chin)
- 15a Mizo (a.k.a. Lushai) – Mizoram
- 15b Hmar – northern Mizoram and Cachar
- 15c Anal (a.k.a. Lamgang Kuki) – in southeastern Manipur
- 15d Paite – in southeastern Manipur
- 15e Gangte – in southeastern Manipur
- 15f Thado (a.k.a. Thadou) – in southeastern Manipur
- 15g Kom – in southern Manipur
- 15h Purum – in southern Manipur
- 15i Chiru – in southern Manipur and an enclave south of Silchar in Assam
- 15j Biete – in an enclave northeast of Silchar in Assam
- 15k Chorei – in an enclave southwest of Silchar in Assam
- 15l Bawm – on the Tripura-Assam border
- 15m Hrangkhoh – in an enclave in northern Tripura
- 15n Lakher (a.k.a. Mära) – in southern Mizoram
- 15o Simte – northeastern Mizoram and the neighbouring portion of Burma
- 15p Zo – northeastern Mizoram and the neighbouring portion of Burma
- 15q Vaiphei – northeastern Mizoram and in Burma around Tiddim
- 15r Tiddim Chin – northeastern Mizoram and in Burma around Tiddim
- 15s Falam Chin – in southeastern Mizoram and in Burma around Falam
- 15t Haka Chin – in southeastern Mizoram and in Burma around Haka
- 16 Mru (not to be confused with Märu) – central hills of the Chittagong
- 17 Kachinic
- 17a Luish
- 17a-1 Andro [in Manipur; the Andro now speak Meithei]
- 17a-2 Sengmai [in Manipur; the Sengmai now speak Meithei]
- 17a-3 Kadu [in Upper Burma; the Kadu now speak Burmese]
- 17a-4 Chakma [in Chittagong hill tracts of Bangladesh, now speak Bengali]
- 17a-5 Sak [southern Chittagong hill tracts of Bangladesh, still speak Sak]
- 17b Jinghpaw (Kachin)
- 17b-1 Singpho – in the Arunachal hills to the east of Sadiya, north of the Tangsa area (see 5a-1)
- 17b-2 other Jinghpaw dialects in northern Burma and southwestern China
- 18 Raji-Raute – Pithauragadh district in Uttarakhand and the western Nepalese Terai
- 19 West Himalayish – in Himachal Pradesh and Garhwal
- 20 Bodish – in Tibet and neighbouring portions of Pakistan, Nepal, India and Bhutan
- 21 Newaric – in Nepal.
- 22 Kiranti – in Nepal.
- 23 Tamangic – in Nepal.
- 24 Dura – in Nepal.
- 26 Magaric – in Nepal.
- 26 Chepangic – in Nepal.
- 27 Dhimalish – in the southeastern Nepalese Terai and on the Bhutanese *duar*
- 28 Lhokpu – in Bhutan

- 29 Gongduk – in Bhutan
- 30 Tshangla – in Bhutan
- 31 Nungish – in Yúnnán province in  
China and in northern Burma
- 32 Karenic – in eastern Burma and western Thailand
- 33 Pyu [extinct] – central Burma
- 34 Sinitic – in China
- 35 Bái – in Yúnnán province in China
- 36 Tǔjiǎ – in China
- 37 Qǐāngic – in Sīchuān province in China
- 38 Ěrsū cluster – in Sīchuān and Yúnnán provinces in China

A perennial issue is the likely provenance of the Nagas, a question that we can ask about any people or language community in the world. Inevitably, an ethnolinguistically informed answer to such a question is necessarily many-faceted because the prehistory of any language community is likely to have been complex. More fundamentally, the linguistic ancestors of a community were not necessarily the same people in time and space as the biological ancestors of that community. The material culture and agricultural package of a community represents a third quantity altogether. Nonetheless, a detailed understanding of the linguistic affinities and population genetics of language communities and the archaeological record of the region which they inhabit sheds much light on the possible prehistory and provenance of peoples and on the interactions which these peoples may have had in the past with other groups.

The Neolithic cultural assemblage which is associated with northeastern India is the Eastern Indian Neolithic. This cultural assemblage is entirely distinct from other Neolithic traditions of the Indian Subcontinent. The antecedents of the Eastern Indian Neolithic lie to the northeast, i.e. possibly in what today is Sīchuān province. This is shown in Fig. 4. In several publications (van Driem 1998, 2001, 2002), I have argued for the hypothesis that the intrusion of this cultural assemblage into the northeast of the Subcontinent could represent Proto-Tibeto-Burmans originating from Sīchuān, spreading southwest onto the lower Brahmaputran plain and introducing themselves and their Eastern Indian Neolithic culture to resident Austroasiatic populations. Indeed, the two major linguistic phyla Tibeto-Burman and Austroasiatic meet in this area. More recently, I have compared and discussed alternative and competing interpretations for the spread of Tibeto-Burman as reflected in the archaeological record (van Driem 2007a).

The main uncertainty about this scenario is the dating of the relevant cultural assemblage. Archaeologists have estimated the Indian Eastern Neolithic to date from between 10,000 and 5,000 BC (Sharma 1989b; Thapar 1985). If these estimates are taken at face value, it would mean that northeastern India had shouldered adzes at least three millennia



△ Fig. 4. Brahmaputra basin and surrounding hill tracts colonised by ancient Tibeto-Burmans from Sīchuān (van Driem 2001: 415). The dating of the Eastern Indian Neolithic is currently a matter of archaeological conjecture and still not yet based on any sound stratigraphy or reliable radiocarbon or accelerator mass spectroscopy dates.

before they appeared in Southeast Asia. Therefore, most archaeologists now tend to ascribe younger dates to the Indian Eastern Neolithic. Yet such archaeological dating estimates lie in the realm of conjecture, since a solid stratigraphy and calibrated radiocarbon dates are still unavailable for this major South Asian cultural assemblage.

What arguments can be advanced to decide whether the advent of ancient Tibeto-Burman language communities to the Indo-Burmese borderlands took place in the Neolithic or much later, say, in the Bronze Age? The provenance of Tibeto-Burman groups in northeastern India and the Indo-Burmese borderlands and the timing of their arrival is

inextricably connected with the question of the homeland and provenance of the Austroasiatics, whose current geographical centre of gravity lies in this very region. In addition to the geographical distribution of modern Tibeto-Burman and Austroasiatic language communities, we can examine the findings of population genetics and attempt to trace the spread of agriculture and crops.

Linguistic palaeontology, a term introduced by Adolphe Pictet in 1859, is an attempt to understand ancient material culture on the basis of the lexical items which can be reliably reconstructed for the oldest level of the common ancestral language. Linguistic palaeontology and native lore both suggest that the ancient Tibeto-Burmans were cultivators of broomcorn millet *panicum miliaceum* and foxtail millet *setaria italica*, whereas linguistic palaeontology qualifies the ancient Austroasiatics as the most likely candidates for the first cultivators of rice.



The oldest millet cultures have been found in the ancient Yellow River basin and on the North China Plain. Currently, the earliest attested domestic millet dates from before 6000 BC at 興隆溝 Xīnglǒnggōu near 赤峰 Chìfēng, where there was a Neolithic culture without sickles (Fuller *et al.* 2007). The rice story, however, is more complex, and the plot of the story has changed more than once in recent decades. Whereas the oldest rice cultivation was once held >incontestably< to lie in the Indian subcontinent (Haudricourt & Hédin 1987), subsequent scholarship moved the homeland of rice agriculture from the Ganges to the Yangtze. For years conventional wisdom in archaeological circles dictated that rice was domesticated in the Middle Yangtze, perhaps as early as the sixth millennium BC.

Currently available evidence indicates that immature morphologically wild rice may have been used by foragers before actual domestication of the crop, e.g. at the 八十擋 Bāshídàng site (7000–6000 BC) belonging to the 彭頭山 Péngtóushān culture in the Middle Yangtze and at sites in the Yangtze delta area such as 跨湖橋 Kuàhúqiáo, 馬家浜 Mǎjiābāng 河姆渡 (5000–3000 BC) and Hémùdù (5000–4500 BC). However, only c. 5000 BC was the actual cultivation of rice probably first undertaken by people in the Lower Yangtze, who at the time relied far more heavily on the collecting of acorns and water chestnuts.

Today, our understanding of the palaeoethnobotanical picture is more complex. The two main domesticated varieties of rice, *oryza indica* and *oryza japonica*, are phylogenetically distinct and would appear to have been domesticated separately. *Oryza indica* derives from the wild progenitor *oryza nivara* and was first cultivated in South Asia or western Southeast Asia, perhaps in two separate domestication events. *Oryza japonica* derives from the wild progenitor *Oryza rufipogon* and was first cultivated to yield the early *Oryza japonica* along the Middle Yangtze. Harvey *et al.* (2006) have critically reassessed the morphometrics of rice finds associated with various Neolithic sites throughout the Yangtze basin in the light of recent genetic findings. The wild progenitor *Oryza rufipogon* was not fully domesticated in the Lower Yangtze to yield the early *Oryza japonica* until c. 4000 BC.

More recently, scholars have increasingly begun to take note of findings that would move the original homeland of rice cultivation back to the Indian subcontinent. At the Lahuradewa site (26°46' N, 82°57' E), the early farming phase, corresponding to period 1A in the site's clear-cut stratigraphy, has radiocarbon dates ranging from c. 5300 to 4300 BC. Carbonised material from period 1A was collected by the flotation method, yielding *Setaria glauca* and *Oryza rufipogon* as well as a morphologically distinct, fully domesticated form of rice >comparable to cultivated *Oryza sativa*< (Tewari *et al.* 2002). More recently, accelerator mass spectroscopy dates were obtained on the rice grains themselves, corroborating the antiquity of rice agriculture at the site. Most recently, new radiocarbon dates for rice agriculture have been coming from the Ganges basin, with the Tokuvā site near Allahabad now yielding similar dates (Vasant Shinde [Vasant Śivarām Śinde], personal communication 27 November 2007).

Turning to northeastern India and the Indo-Burmese borderlands, where we find the Naga peoples and related Tibeto-Burman language communities, we must recognise that, notwithstanding the excellent archaeological work done in the Ganges and Yangtze river basins, the archaeology of ancient rice agriculture is simply not known because no substantive archaeological work has been done on the Neolithic in the most relevant areas, e.g. Bangladesh, northeastern India and Burma. The sheer dearth of archaeological research in these areas leaves entirely open the possibility that rice cultivation may have originated in this region. Perhaps the remains of the very first rice cultivating cultural assemblages lie buried forever in the silty sediments of the lower Brahmaputran basin or were washed out long ago into the depths of the Bay of Bengal.

At least four species of wild rice are native to northeastern India, viz. *Oryza rufipogon*, *Oryza officianalis*, *Oryza perennis*, *Oryza meyeriana*, and over a thousand varieties of domesticated rice are currently in use in the region (Hazarika 2006a). The different varieties of rice in northeastern India are cultivated in three periods by distinct cultivation processes. In the process of *ahu kheti*, the rice is sown in the months of *fagun* and *sot*, i.e. mid February to early April.

The seedlings are not transplanted but ripen in just four months in fields which must be constantly weeded. In *bao kheti*, the rice seedlings are sown from mid March to mid April in ploughed wet fields and likewise do not need to be transplanted. In *sali kheti*, the rice is sown from mid May to mid June, and the seedlings are transplanted. *Sali kheti* rice varieties are held to have been derived from the wild *officianalis* rice still widely found in swampy village areas. The wild *rufipogon* rice cannot be used for human consumption because the plants shed their seeds before they ripen, so that *rufipogon* rice is used in Assam and other parts of northeastern India as cattle feed (Hazarika 2006b).

So, might ancient Tibeto-Burmans originating in what is today China indeed have spread southwest onto the Brahmaputran plain and introduced themselves and the Eastern Indian Neolithic cultural assemblage to resident Austroasiatic populations, or will emergent population

genetic findings and new archaeological research in the area compel us to embrace another interpretation of prehistory? New population genetic findings, especially on Tibeto-Burman populations in the Himalayas, e.g. Parkin *et al.* (2006a, 2006b), Kraaijenbrink *et al.* (2006, 2007), have shed light on the biological ancestry of populations in the northeastern portion of the Subcontinent. I have had occasion to discuss the possible interpretations of emergent population genetic findings with respect to the ethnolinguistic prehistory of Tibeto-Burman and Austroasiatic populations at some length (van Driem 2006, 2007b, 2008), and new results of genetic assays will soon appear in print. Ultimately, the answers to our questions about the provenance of the Nagas and other peoples of the northeast will have to come from a careful reconstruction of the disparate societal and migrational processes which have shaped the ethnolinguistic prehistory of all of northeastern India and adjacent regions.