

PEOPLING OF



THE NORTHEAST

Part 2

Tanmoy Bhattacharya

Teaches at Centre of
Advanced Studies in Linguistics
Faculty of Arts, University of Delhi
& Chief Editor of Indian Linguistics

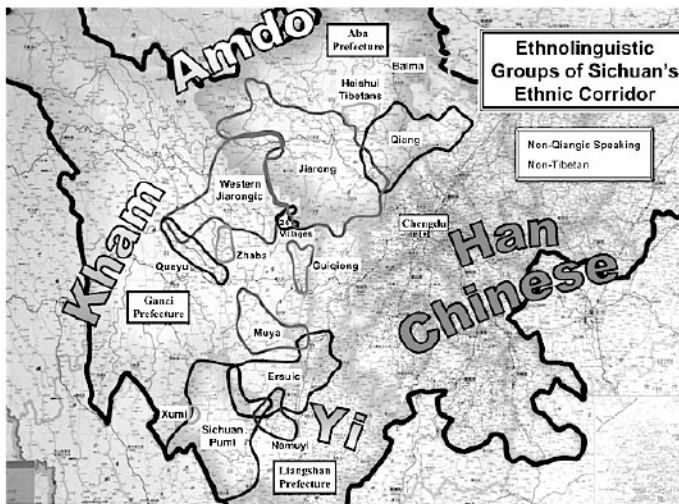
IN Part 1 (see vol. 02, issue 03, p. 66-73), we saw that based on the archaeological evidence for the proposed Northeast Indian Neolithic culture, it can be established that the “Western” Tibeto-Burman group (to which, Meeteilon belongs) split off the earliest from a common ancestor inhabiting most probably the South-western Sichuān province of China; and the earliest date for such an intrusion into the northeast of India was set at 7000 to 6000 BCE. It was also pointed out that this is by far the earliest intrusion into India of a Neolithic culture, as the aceramic culture of Burzahom in Kashmir is dated around 2800-2500 BCE.

*Tawang Gate, Arunachal Pradesh
Dhrubazaanphotography, CC BY-SA 4.0*



THE MULTI-ETHNIC CHINA

HOWEVER, this might lead us to believe that the modern Tibeto-Burman speaking population and their languages are a derivative of the broadly – and wrongly – termed “Chinese” people and language. This, however, cannot be further from truth. First, despite the obvious perception created mostly by the uninterested (and uninformed), China is a highly multicultural and multilingual entity. Its multicultural character is evidenced by the existence of the well-known “ethnic corridor” extending from southern Gansu and eastern Qinghai, including western Sichuān and south-eastern Tibet, and ending at western Yúnnán and northern Burmese and Indian borders (Sun Hongkai, 1990, “Languages of the ethnic corridor in western Sichuan”, LTBA 13:1, 1-31). There are about 56 ethnic groups in China, nearly half of which reside in parts of Sichuān and Yúnnán; these numbers refer only to “officially” recognised ethnic groups, the actual figures are higher. One reason for this high concentration of ethnic groups (and languages) is due to the fact that this area served as a corridor for four of the five language families of southeast Asia -- Sino-Tibetan (Tibeto-Burman and Sinitic), Tai-Kadai, Austro-Asiatic (Mon-Khmer), and Hmong-Mien; Austronesian being largely a matter of South or South-eastern China. A part of the ethnicity is visible from Map 1.



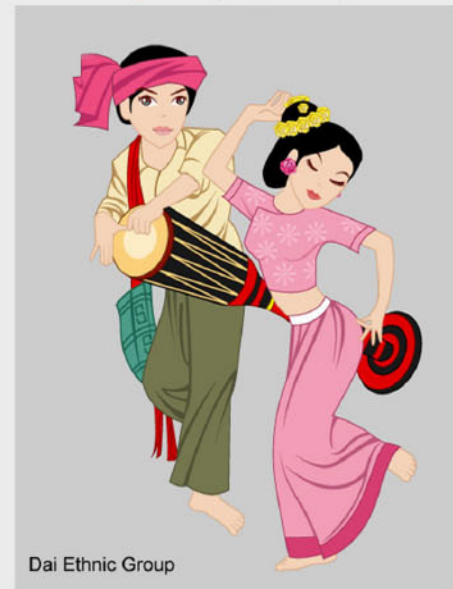
Map 1: The Sichuān ethnic corridor

These various ethnic groups have differing customs, rituals, costumes, totems, etc.; some of the examples are given in the accompanying graphics.

✓ Dai (Southern Yúnnán)



✓ Dai (Costume)



✓ De'ang Totem

✓ De'ang (Yúnnán)



✓ Yi Marriage

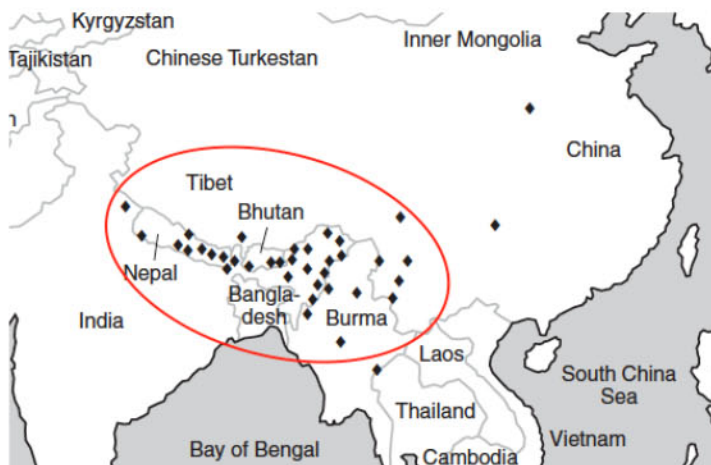
✓ Yi (mostly Sichuān)

Figure 1: Chinese Ethnicity

Multilinguality of China too derives mostly from the multi-ethnicity of this ethnic corridor. Some of the languages that are spoken in this ethnic corridor are: Tibetan, Hàn, Yi, Tai, Jǐngpō (Kachin), rGyarong, Qiāng, Primi (Pǔmi), Miáo, Palaung, Bái, Nàxī, Nù, etc. Most of the languages of this ethnic conglomeration are Tibeto-Burman, a few of them are Austro-Asiatic (Palaung), Tai-Kadai (Tai), and Hmong-Mien (Miáo).

A TAXONOMIC CONFUSION

THE concentration of Tibeto-Burman languages in this corridor matches the observation that most of the Tibeto-Burman languages of the Sino-Tibetan group lie in the cluster area defined by the foothills of the Himalayas, north-eastern India, and the stretch between south-eastern Tibet and northern Burma. In one count (van Driem, 2005, “Tibeto-Burman vs Indo-Chinese: implications for population geneticists, archaeologists and prehistorians”, in *Peopling of East Asia*, eds. L. Sagart, R. Blench, and A. Sanchez-Mazas. Routledge: London & NY), 3/4th of the 39 groups of Tibeto-Burman languages fall in this area, whereas, most number of speakers of the language family titled “Sino-Tibetan” inhabit the rest of China in the north and the east. This somewhat puzzling state of affairs (shown in Map 2, where the cluster area is encircled) is at the heart of issue in the classification of these language groups.



Map 2: Distribution of Tibeto-Burman language groups (from van Driem, 2005)

That is, whether the name of the group that includes Chinese -- in some ways the “odd” member of the group, and the Tibeto-Burman taxon, which includes the majority languages spoken in the geographical area, should bear the adjective Sino or Tibeto-Burman? In fact, the appellation ‘Burman’ should also be questioned in this light since Burma is only a minor actor in this retelling of the migration narrative, as only Lolo-Burmese and Karenic group of languages travelled deep south to Yúnnán, Burma and Thailand.

The existence of other allied linguistic groups like the Tai-Kadai (‘Daic’ as in the older name) apart from creating further (scholastic) confusion within this region had driven scholars before the WW II to also include this group within a greater Sino-Tibetan family, inspired no doubt by attempts in the latter half of the 19th century to club Chinese and Thai together. Another geographically rational phylum Hmong-Mien, is conjectured to have moved southwards – under pressure from the incoming Hàn Chinese – from central China (near Pengtoushan) to the valley between the Yangtze and the Mekong (Roger Blench, 2014, “Origins of Ethno-linguistic Identity in Southeast Asia”, in *Handbook of East and Southeast Asian Archaeology*, eds. J. Habu, P. Lape, J. Olsen, and J. Zhichun, Springer US). The fourth group in the region, namely, the Austroasiatic (Mon-Khmer) languages, which also traces its origin to about 7000 BCE (like the Tibeto-Burman/ Sino-Tibetan group), along with the Tibeto-Burman group, is one of the two main actors of the story about peopling of the Northeast of India.

Whether or not the Tibeto-Burmans encountered the Austroasiatics in their journey to the Northeast of India is a story that will require its own space and time, which shall be revealed in the third part of the series in a subsequent issue. For the time being, we can look at the cartographic representation of one of the proposals of the original homeland of the Austroasiatic group (R.A. Blust 1996, “Beyond the Austronesian homeland: the Austric hypothesis and its implications for archaeology”, in *Prehistoric Settlement of the Pacific* Vol. 86, Pt. 5, (ed.) Ward H. Goodenough 117–60, American Philosophical Society: Philadelphia). According to this view the Austroasiatics emerged from the area ‘in which the Salween, Mekong and Yangzi run parallel at their narrowest watershed’, shown in Map 3.



Map 3: The original homeland of Austroasiatic group (based on Blust 1996)

In short, the multi-ethnicity and multilinguality of China is the reason behind scholars of Sino-Tibetan linguistics/ languages going back and forth on a number of taxonomic proposals throughout almost the last two centuries. This confusion in the taxonomic nomenclature can be summarised as in Figure 2:

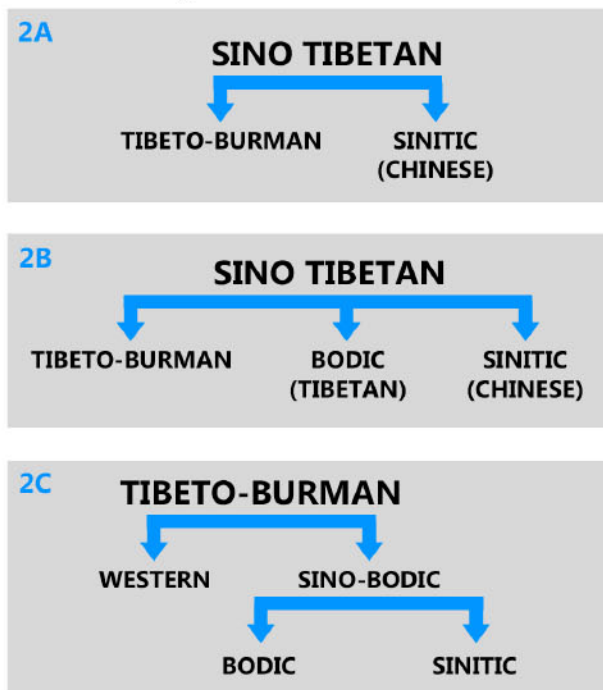


Fig 2: Taxonomic Classification of Sino-Tibetan Languages

Among these classifications, I adopt the proposal in (2c), which is roughly based on van Driem (2005); further justification for this classification is being provided in the following section, in particular, in support of the taxon “Sino-Bodic”, which lies at the heart of our story in this Part.

THE SINO-BODIC

THE most advanced Neolithic culture in China is considered to be the region around the Yellow River (Huang He) basin/ valley in the North, rather than Southwest China – the purported precursor of the North-eastern Neolithic of India. In fact, the two Neolithic civilisations that flourished in 6500-5000 BCE in Gānsù/ Shǎnxī provinces (Dàdīwān) and in Hebei/ Henan provinces Péilǐgǎng-Císhān) are located in the North, in the fertile reaches of the middle course of the Yellow River. In fact, van Driem (2005) conjectures that it is perhaps the affluence of the North that drove the early Tibeto-Burmans from the Southwest (in Sichuān) to the North. However a split between the Southern and Northern Tibeto-Burmans around 7000 BCE, prior to the dawn of the Northern Neolithic cultures, preceded this northward migration. It was the southern Tibeto-Burman group that set off in a westerly direction through the eastern Tibetan plateau looking for fertile agricultural land elsewhere, namely, in the northeast of India. These two different directions of displacements are shown in Map 4 (in red and black arrows, respectively):



Map 4: Westerly and Northern Movement from Southwest

If the southern Tibeto-Burmans moved in the westerly direction towards Northeast India, what did the northern Tibeto-Burmans do after moving northward? Mainly two important subsequent events ensued that involved this group that are important for our story. The civilisations that succeeded the Dàdīwān and the Péilǐgǎng-Císhān, led to westward and eastward movements, respectively. The westward movement was through the trade route across the Himalayas into Kashmir that gave rise to the Kashmir

Neolithic; linguistically, this was identified as the Bodic group. The Bodic made its appearance in a preliminary form a long time ago, in Julius von Klaproth's observation almost two centuries ago (1823) that Tibetan and Chinese are closer than they appeared. This closeness was captured by recognising a higher classification called 'Sino-Bodic' (van Driem, 1998, Neolithic correlates of ancient Tibeto-Burman migrations. In: Blench R, Spriggs M, (eds.) *Archaeology and language II: archaeological data and linguistic hypotheses*. New York: Routledge; 67–102); Tibetan, Sherpa, Ladhaki are prominent Bodish languages.

The eastward movement led to the evolution of the Hân Chinese stock, which populated the rest of China (92% in 2010) throughout the intervening millennia, giving rise to the impression of a homogenous culture in modern China, no doubt, aided in the process by an aggressive homogenizing political process of the 20th century.

We return next to the main theme of our present story – in adopting Figure 2c for classifying the languages of the area encompassing China, northern parts of the South East Asia landmass, and the Northeast of India, we have adopted three taxa – Western, Bodic, and the Sinitic. And for a purpose!

IS THERE A BODIC-WESTERN CONNECTION

THE objective is to raise this question, a relation that the taxonomy in (2c) forbids, shown in Figure 3 in red.

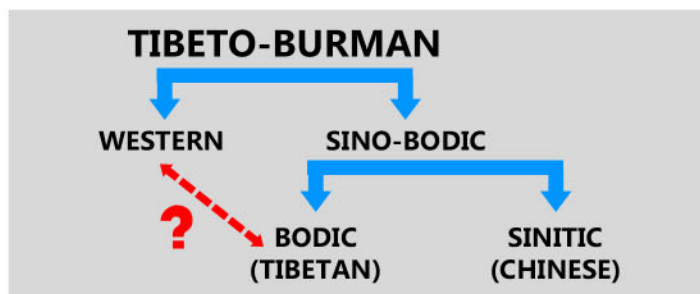


Figure 3: Connecting "Western" and Bodic

The careful reader will have appreciated already the validity of this question, it is but obvious that both these migrations led to movement into India, one through the Northeast (the "Western" group, the story of Part 1) and the other through the North (the Bodic into the Kashmir valley). However, these two intrusions have been

separated in time by at least three millennia. Whereas the "Western" group entered India around 7-6000 BCE, the Bodic group is supposed to have entered much later around 3000 BCE, coinciding with the Northern Neolithic culture of Kashmir, which predates the Aryan intrusion through the Northwest later around 2000-1500 BCE.

The fact that these two groups of languages belong to the same phylum often escapes the attention of scholars working with these languages, with the result that apart from a nomenclature identity sometimes invoked between these two groups, for example, Tibeto-Burman and Tibeto-Himalayan, there is very little work that *wants to* establish the connection shown in Figure 3. The reason for this absence is obvious, and is known to plague many disciplines of thought – 'we believe what we see', and in case of languages, 'we believe what we hear'. The Himalayan languages (for example Darma, Rangpo, etc.) do not *sound* the same as, say, Meeteilon. However, I will show next that this anti-rationalist stance, at least in this case, can be defeated in two ways; one, by playing the game by their rule and defeating them, and two, by scientifically comparing the properties of the two groups of languages. The first is the genetic story, which will be more elaborately discussed in Part 3 of this series, and the second is the linguistic story.

THE GENETIC EVIDENCE

BEFORE we consider another set of evidence – apart from the migrational and the archaeological evidence recounted in Part 1 and here so far – it must be kept in mind that speculations about the connection between the linguistic and the genetic affinity is coloured by a gap between the two; while linguistic affinity through the method of "historical reconstruction" can reconstruct proto-stages of languages to typically less than 10K y.a. (years ago), genetic antiquity of a population can be dated much farther back to typically between 25-40K y.a.

With this clarification, let me come back to the query why I identified the genetic evidence as 'playing the game by their rule' a while back. The rationalist stand to knowledge is that it need not rely on surface facts, in facts that are easily observable and experiential, but rather that knowledge lies in reasoning. The genetic studies on the other hand start with the null hypothesis that 'seeing is believing'. The fact that the people of the

Himalayas speaking Bodic languages may look similar to people of the Northeast speaking “Western” Tibeto-Burman, lies at the heart of the genetic evidence.

Before we establish the genetic connection between the Bodic and the “Western”, let us consider, genetically, that the “odd” group (though with the most number of speakers), namely, the Chinese, may not be the odd member of the Sino-Tibetan family either. The findings of the Chinese Human Genome Diversity Project (reported in Chu et al. 1998, “Genetic relationship of populations in China”, *Proceedings of the National Academy of Sciences of the United States of America* 95: 11763–8, van Driem, 2005) show that genetically East Asian populations are derivative of the Southeast Asian populations, as such therefore, the Chinese ancestral populations could have migrated to the Yellow River area from South-western China. This can be seen as the justification for re-naming the phylum as Tibeto-Burman, as in Figure 2c.

Genetic studies in this domain revealed correlations between genetic markers and geographical distribution of language communities, however, the correlations were not with the maternally inherited mitochondrial DNA (or mtDNA) but with the paternally inherited Y chromosome. This has been termed as the Father Tongue hypothesis by van Driem (2002, “Tibeto-Burman replaces Indo-Chinese in the 1990s: review of a decade of scholarship”, *Lingua* 111: 79–102.). The Tibeto-Burman speakers of India carry both the Y-chromosomal haplogroup O2a (M95), which has been identified as the

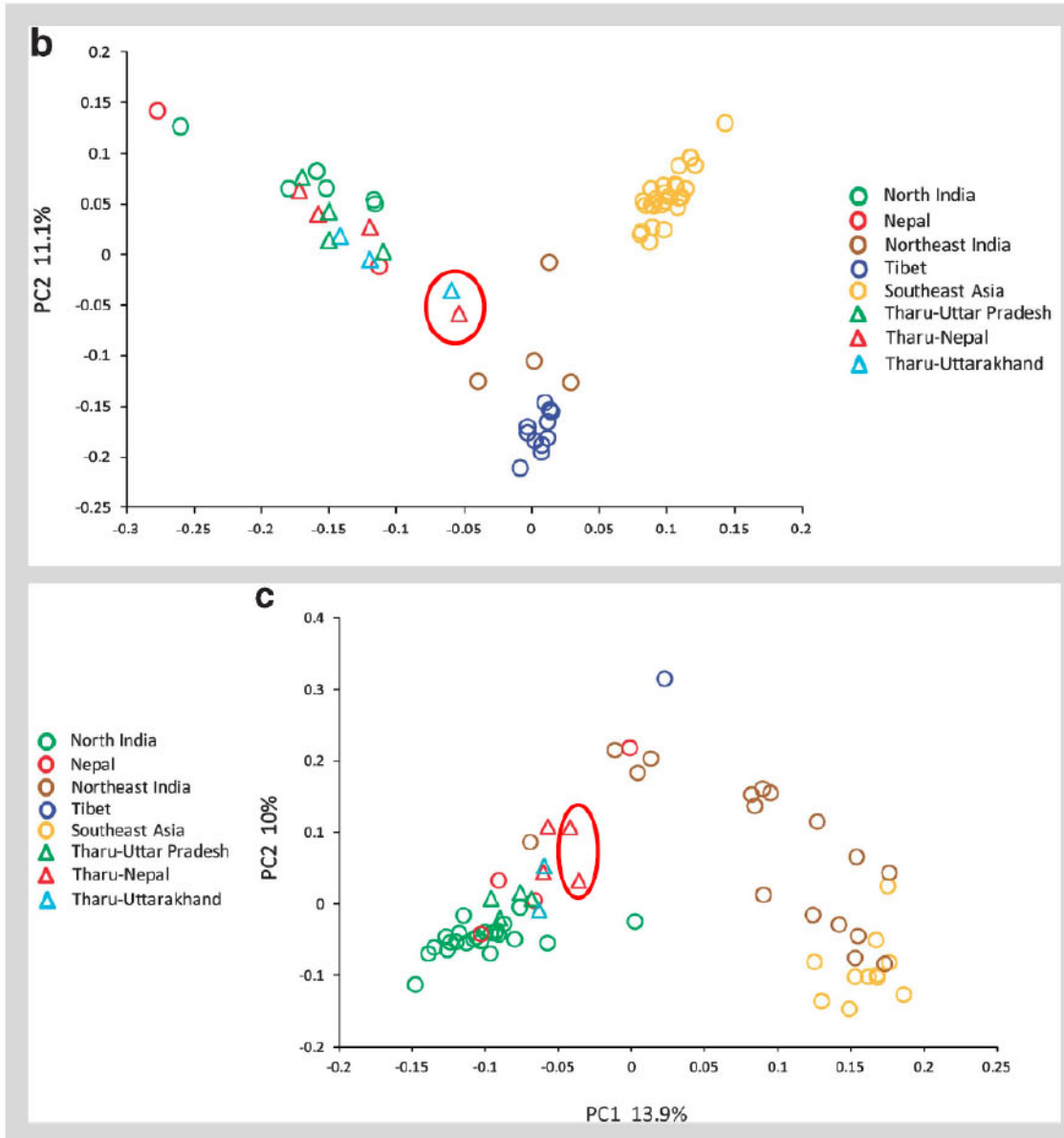
definitive SE Asian (and Indian, as in Munda) Austroasiatic marker, as well as the O3a3c (M134) haplogroup, which is not found in the Austroasiatic populations. Generally, the O3 (M122) area is considered to be where North-eastern India, South-eastern Tibet, and Northern Burma meet. According to van Driem (2005), “the bearers of Y chromosomal haplogroup O3a3c (M134) expanded eastward into Sichuān and Yúnnán, north and northwest across the Tibetan plateau as well as westward into the Himalayas and southward into the Indo-Burmese borderlands”.

There are quite a few studies, which have shown through Y-chromosome and mtDNA studies, a high degree of genetic homogeneity between Himalayan and Northeast Indian Tibeto-Burman groups. However, as I noted earlier, intrusion dates of these two groups differ considerably. As a result of this difference, for reasons of the route taken or the settlements preferred, there are also distinct differences in the compositions of the gene-profiles of these two groups. For example, one study (Gayden et al, 2009, “Genetic insights into the origins of Tibeto-Burman populations in the Himalayas”, *Journal of Human Genetics* 54, 216–223), noted that the Himalayan population of their study had both low genetic diversity compared to the North-eastern group (126 versus 151 alleles variants), and comparatively high inter-population diversity (0.012 versus 0.003) indicating multiple genetic sources, genetic drift, and/or founder effect – all of which can be attributed to the delay in intrusion.

In another significant study, Chaubey et al 2014 (“Unravelling the distinct strains of Tharu ancestry”, *European Journal of Human Genetics* 22, 1404–1412) corroborated the presence of an East Asian genetic component in the Tharu in terms of Y-Chromosomal and mtDNA markers. The Tharu were described by ethnographers in the 19th century as representing an East Asian phenotype, among others. The name Tharu is used as an umbrella term for a variety of endogamous groups that speak different dialects; “morphologically” they are described as Mongoloids. However, they themselves claim to be descendants of Rajputs, and are therefore a north Indian Hinduised ethnic group.

The study found that in the Principal Component Analysis (PCA) based on mtDNA and Y-chromosome relative frequencies of haplogroups within populations, although the Tharu groups (from Nepal, Uttarakhand, and UP) align closely to the north Indian populations, they tend to tilt towards Southeast Asian and Tibetan groups (see Graphs 1b,c). This is consistent with observations that the East Asian specific genetic component among the Nepali and the Tharu groups has entered the region from the north over the Himalayas.

It is quite clear from these graphs that the UP Tharu is closer to the Indian population than either the Uttarakhand or Nepal Tharu. The fact that the maternal and paternal lineages among the Tharu trace back to East Asian populations, is clearly seen in Table 1, which shows that 15-51% of their profiles contain the relevant haplogroups.



Graph 1b,c: Principal component analysis (PCA) based on mtDNA (a) and Y-chromosome (b) relative frequencies of haplogroups within populations (from Chaubey et al, 2014)

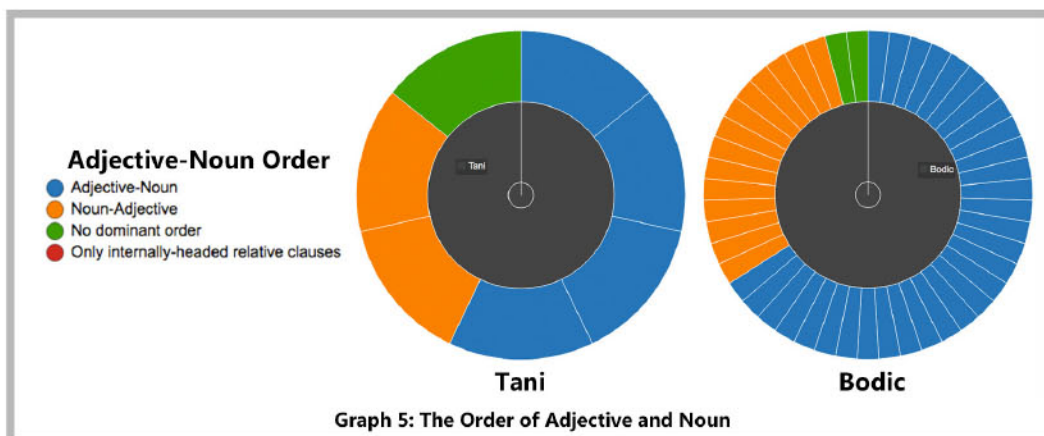
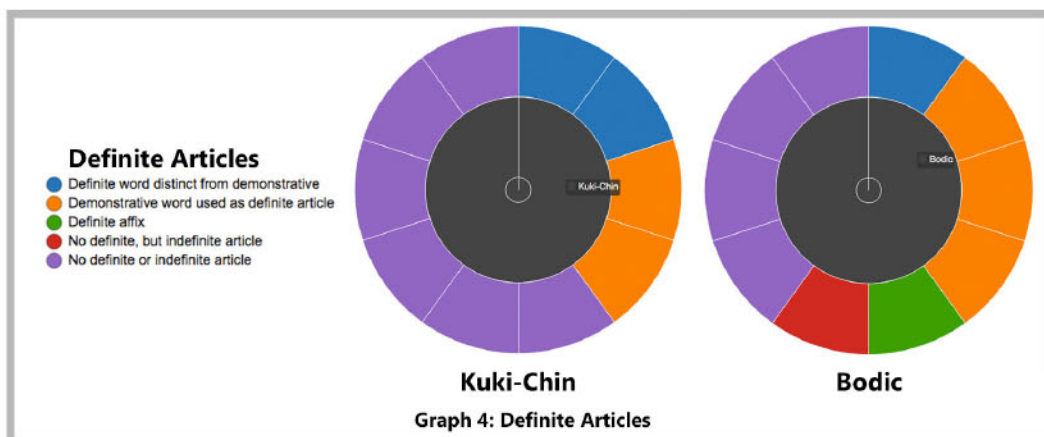
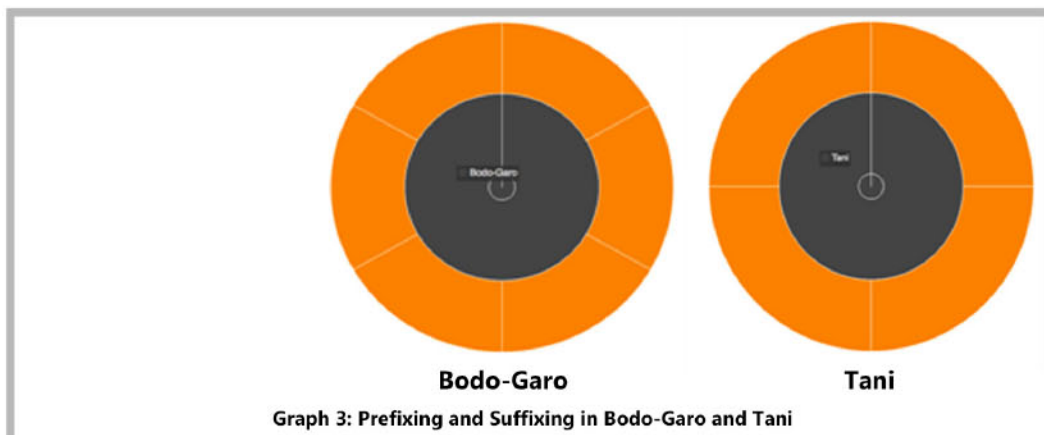
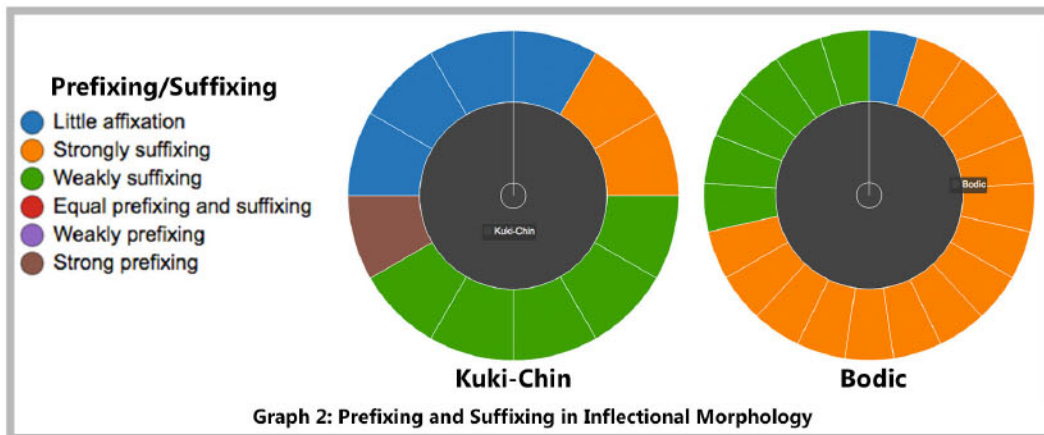
	mtDNA E/SE Asia	Y-Chromosome E/SE Asia
Tharu-Nepal	51.47	40.59
Tharu-UT	42.22	26.67
Tharu-UP	25	14.63

Table 1: mtDNA and Y-chromosomal haplogroup profiles in 3 Tharu ethnic groups

The genetic evidence therefore points towards a situation which can be most appropriately described as an inversion of the popular adage ('looks can be deceptive'), here, it seems, that looks cannot be deceptive. The Tharu, and many such populations, who may have lost their languages, ethnically still carry a substantial genetic imprint of their origin.

THE LINGUISTIC EVIDENCE

WITH regards to the linguistic evidence, it was Asko Parpola, the famous Indologist who first pointed out the relation between the Northern Neolithic and the Tibeto-Burman populations in the Himalayas (Parpola, A.1994. *Deciphering the Indus Script*. Cambridge: Cambridge University Press). He pointed out that "The language (assuming there was only one) spoken by the people of the Northern Neolithic may have died out, but not without influencing the later languages of the regions. The many phonological and syntactic peculiarities of the Indo-Aryan Kashmiri, which set it apart from the rest of the Dardic group, point to an extinct substratum language ...". And further, with regard to the Himalayan languages (as opposed to Tibetan), he stated that it "may have arrived by a different route, from the northwest, and thus the Northern Neolithic may have been Proto-Himalayan-speaking. This is suggested also



by the fact that a manuscript relating to the Bon religion of western Tibet has been discovered in Dun Huang, written in the extinct Zhang Zhung language, which appears to have been closely related to Kanauri (=Kinnauri)". As pointed out by van Driem (1998), the apparently archaic nature of Kiranti, East Bodish, and other Himalayan languages and their affinity with Old Chinese supports this.

I will now examine some of the similarities between the two language groups by way of comparing certain properties of these groups using the WALS database. The properties that I am considering are the morphological and syntactic ones like the following:

- ☎️🌀🕒 Prefixation/ Suffixation
- ☎️🕒🕒 Definite articles
- ☎️📏🕒 Adjective-Noun order

Instead of considering a lot of data, let us view the presence or absence of these properties in the two groups in graphic forms.

Of course these representations are somewhat skewed since the extent of the colour patches depends upon the number of languages the WALS database has data on, and the rough area chosen to plot the graphs. For example, the Kuki-Chin cluster in Graph 2 has only two languages (Mikir and Tangkhul) with the strongly suffixing property (in orange), whereas the Bodic languages have 15 such language samples in the database. However, what we can note is that, in spite of the

outward dissimilarity, the range of the colours is comparable. This will be true for the other two properties as well.

If we extend the area chosen to further northeast, we do get a Bodic pattern as well with more strongly suffixing language groups like the Bodo-Garo and the Tani group of languages; this is shown in Graph 3.

So whatever doubts that might have arisen after examining the Kuki-Chin pattern in Graph 2, will be allayed once we have a broader group of language samples in our frame.

In Graph 4 too, the range of colours are quite comparable, thus in the property of the presence or absence of definite articles, both groups show

similar range in the absent category (marked here in purple). The odd ones here within the Bodic group marked in red and green are Byangsi (spoken in eastern Kumaon) and Balti (spoken in Pakistan and Kargil), respectively, which have an indefinite article and a definite affix.

In Graph 5 too, we find a pattern that is comparable in range, both Tibeto-Burman (represented by Tani) and Bodic languages have green, blue, and orange colour coded segments. Note that in both groups, there are more languages that have Adjective-Noun order, although in Bodic, the number of language samples is more in the area chosen that have the opposite, i.e. Noun-Adjective order.

CONCLUSION

Thus, at the end of this excursion of comparing the two groups of people and languages in terms of their genetic profile and Morphological/ Syntactic properties, respectively, we pose the difficult question of identity by asking 'Who is the other?' Our domestic help may no longer have access to the language or the culture of a Tibeto-Burman past, but the fact that she hails from an area in UP not far from a concentration of Tharu population, the possibility of an ethnic Tibeto-Burman make-up cannot be ruled out. Every tribe, time and again, has claimed some 'other' place as their home, and in this constant vortex of human migration, who is the other? Is the other the different? I will take this question up in Part 3 of the series.

