(De)classifying Arunachal Pradesh Languages: Reconsidering the Evidence

ROGER BLENCH

Department of Indigenous Affairs, Government of Arunachal Pradesh

McDonald Institute for Archaeological Research, University of Cambridge

Department of History, University of Jos

Kay Williamson Educational Foundation

Table of Contents

ABSTRACT

The 'North Assam' languages of Arunachal Pradesh represent a major problem in the internal classification of Trans-Himalayan [= Sino-Tibetan] languages. A paper by Blench & Post (2014) argued that we had at that time insufficient data to assign these to the phylum unequivocally. The

last decade has seen a major expansion of documentation and the time is appropriate to reconsider the issue. This paper presents basic information about the most problematic languages, based on recent fieldwork, together with some of the hypotheses concerning their genetic affiliation. It argues that if we apply the same standards as are used in other areas of high diversity, such as Amazonia and Australia, we would certainly classify these as either isolates or small phyla. It also suggests that strategies for reconstructing Tibeto-Burman are ill-adapted to ascertaining the position of these languages.

Keywords: Arunachal Pradesh; languages; Trans-Himalayan; classification

ACRONYMS AND CONVENTIONS

#	quasi-reconstruction
*	regular reconstruction
AD	anno Domini
BC	before Christ
BP	before present
C	consonant
CTB	Common Tibeto-Burman
IPA	International Phonetic Association
kya	thousand years ago
N	nasal
TH	Trans-Himalayan
V	vowel

"The preceding remarks will have shown there is considerable difference between the North Assam dialects…The home of the North Assam tribes may be considered a kind of backwater. The eddies of the various waves of Tibeto-Burman immigration have swept over it and left their stamp on its dialects."

—Konow in Grierson (1909:572)

1. INTRODUCTION

Exactly what Sten Konow (1909) thought about the classification of the languages of 'North Assam', which largely corresponds to the modern-day state of Arunachal Pradesh, may never be clear. His account is both confused and apparently self-contradictory. However, his general conclusion was that these languages were highly diverse and showed evidence of different layers of contact with Tibeto-Burman languages spreading from the north. These perceptions have so far to make much of an impact on the world of Trans-Himalayan scholarship. The implication is that a language can effectively be classified by identifying a few words with likely Trans-Himalayan cognates. This method, while it has a certain Greenbergian charm, has problems which will be discussed at more length in §3.

[1] This is the now widely accepted replacement term for Sino-Tibetan, which was based on a cultural classification on Sinitic, rather than a linguistic analysis.

[2] This is now also a problematic term, since it evolved to group together all the Trans-Himalayan languages except Sinitic, and this is no longer considered a valid subgrouping. In this paper, the term is not used except in reference to other publications.

The purpose of this paper[3] is to take issue with this approach through a re-examination of the problematic languages of Arunachal Pradesh. It proposes we should take seriously the underlying presumption that they are isolates. Moreover, it will suggest that even where languages probably are correctly classified as Trans-Himalayan, we can in part attribute their divergent characteristics to substrates or contact with language isolates now vanished or submerged.

[3] Since the publication of Blench & Post (2014), Roger Blench has been able to travel to Arunachal Pradesh every year except during covid. The main focus of research has been Idu and Kman in the northeast of the state. Thanks to Mite Lingi, Hindu Meme and Sokhep Kri for collaboration on dictionaries and grammatical work. However, in 2023, the author was appointed Ethnographic Survey Co-ordinator for Arunachal Pradesh, which presented the opportunity to travel more widely in the state. Research is now focused on the Sajolang [= Miji] of Nafra and the Mö or Shertukpen language of Rupa. My thanks to Rijin Deru and Tshering Thongdok for their assistance in bringing teams together for elicitation work.

2. DATA SOURCES

Much of the data available for these languages does not meet modern standards of documentation. Apart from the recension of sources in Konow, van Driem (2001), Burling (2003) and Bodt (2014, 2021) review the earlier materials. While some languages, like Aka (i.e. Hruso) early drew the attention of scholars, languages like Bugun or Meyor have remained barely known. For the Tani languages, which are certainly Trans-Himalayan, Post (2011) has circulated a modern grammar and dictionary of Galo, a Tani language, and more recently Tangam (Post 2017).

Until recently, the main sources were the 'Language Guides' published by the Research Directorate of the Arunachal Pradesh government in Itanagar, included in the references. These can be supplemented by a few related publications by the Central Institute of Indian Languages, which are in the same descriptive tradition. The function of these books is rather opaque; they are part phrase books, part ethnographic

guide and part linguistic description. It is not easy to imagine why one would go to one of the most inaccessible mountainous regions of the world and offer a translation of 'the elephant is the strongest of all animals' (Simon 1976; Hill Miri).

A source for some previously unknown languages is Abraham et al. (2005, 2021) which provides the data according to a wordlist arranged for lexicostatistic coding. Fieldwork between November 2011 and April 2024 has made it possible to improve both the transcription and lexical database for some languages in Arunachal Pradesh as well as critically remapping the area where they are spoken. Despite the critical tone here, the wordlists in most sources are quite substantial and it is usually possible to isolate key morphemes and determine basic sentence structure from the grammar sketch. As a consequence, it is reasonable to say that we should have enough information to classify these languages, or possibly declassify them in the sense of excluding them provisionally from Trans-Himalayan.

3. EXCURSUS ON METHOD

Trans-Himalayan has a curious status as a phylum: long identified by a small set of widespread common lexical items, it has rarely been subject to attempted proof of its genetic unity (Matisoff 2003). In the languages further west such as Kiranti, many exhibit complex verbal morphology, suggesting the possibility that this was a feature of the proto-language. However, this model depends heavily on the internal structure attributed to the phylum. If the ancestors of Trans-Himalayan moved eastward, they would have gradually reduced this morphology, resulting in the monomorphemic structures in many branches. Indeed, this lack of morphology in many branches is problematic, since the similarity of some lexemes to those in other phyla, notably Daic [Tai-Kadai] and Hmong-Mien, has been responsible for a long history of discarded macrophyla proposals (for discussion of these, see van Driem 2008). Leaving aside constructs such as Sino-Austronesian and Sino-Caucasian, the membership is assumed to be broadly as characterised in Bradley (2002). Recent years have seen the publication of low-level reconstructions (e.g. Sun 1993; Mortensen 2003; VanBik 2007; Wood 2008; Button 2011; Pelkey 2011; Mortensen 2023) which is useful, but a long way from the goal of demonstrating the unity of the phylum.

The classification of Arunachal Pradesh languages in the literature is essentially an extension of this model—if a number of lexemes resemble reconstructed Trans-Himalayan, they are assigned to a particular branch. For example, the Ethnologue entry for Puroik says; "Lexical similarity: 22% with Eastern Miji, 20% with Western Miji, 20% with Bugun, 17% with Sartang." This statement is apparently based on Abraham et al. (2005), although despite close inspection it is not visible in their tables. Abraham used a wordlist of 210 items, so this represents some 42 shared items, although the extent to which any of them are actually cognate is never considered.

Not unreasonably, such a low figure has made scholars reluctant to accept their genetic affiliation. One strategy is to attribute the low figure to retention of putative substrate lexicon, sometimes euphemised as 'archaic'. Thus, van Driem (2001:530) says; "It would appear that these [= Kho-Bwa] languages have preserved archaic Trans-Himalayan vocabulary that has otherwise been lost everywhere else or innovated very differently and unrecognisably." The supposition is that these forms were once part of Common Trans-Himalayan and have been lost in other branches through lexical replacement. This is certainly a

possible scenario, but it makes it difficult to understand why this type of explanation would not be applicable to many cases of apparent divergence, and would undermine what is generally presumed about the unity of Trans-Himalayan.

In recent years, an alternative model has become popular, seeing divergent lexicons as resulting from unknown substrates. Thus, Burling (2003: 181) says of Bugun; "If it has a Tibeto-Burman component at all, it seems to be very deeply submerged. It is an open question whether Bugun has a larger Tibeto-Burman component which is simply obscured by a large substrate from another language, or whether it represents some fundamentally different linguistic tradition that has borrowed a bit of Tibeto-Burman vocabulary." Even where membership of Trans-Himalayan is credible there can still be evidence for substrates of an unknown affiliation. For example, the Tani languages are usually considered to pass the test of Trans-Himalayan membership in terms of numbers of cognates and at least some regularity of correspondences. Nonetheless, they incorporate significant amounts of divergent vocabulary whose source is unknown.[5] Indeed, in the Milang language, which is usually considered Tani on the basis of a large number of cognates, a high percentage of cognates seemingly has a substrate of a quite different character on which a Tibeto-Burman structure has been superimposed (Post and Modi 2011; also see §8).

[5] Sun (1993:173) wrote that "beyond the most fundamental core vocabulary, the peculiarity of the Tani lexicon becomes painfully apparent, making it extremely difficult to track down reliable extra-Tani cognates of the PT roots proposed [here]. This means that exhaustively tracing the PT initial and rhyme distinctions back to plausible PTB sources is presently quite impossible."

The core data for this paper is the comparative wordlist given in the appendix table. It tabulates the lexemes for a variety of basic terms in Arunachali languages (excluding the regions bordering Myanmar) and aligns them with the most Common Tibeto-Burman (CTB) starred forms quoted from Matisoff (2003). Apparent cognates are coded in yellow, while other more local cognate sets are assigned other colours. This provides a convenient rapid visual impression of both the correspondences with commonly accepted Trans-Himalayan and the relationships between individual languages.

4. THE MÖ (= MEY, SHERTUKPEN) CLUSTER

Mö or Shertukpen constitutes a small family of languages spoken in the valley of the Tengapani River south of Bomdila in West Kameng district. The name Shertukpen is a construct, from the settlement of Shergaon and 'Tukpen', a Monpa name for the people of Rupa town. The correct name for the Rupa is Mö and their language is Mö nyuk. A related set of lects are spoken by the Sartang, a people also called 'But Monpa' (Dondrup 2010; Bodt 2021). Two other related languages which were formerly classified under the general rubric of 'Monpa' are Duhumbi [= Chug] and Khispi [= Lish], spoken in isolated villages north of Dirang, within the Central Monpa area. Dutta (2007) includes a brief comparative wordlist of Lish and Tawang Monpa in his monograph on Central Monpa, notes its differentiation from both, but makes no comment on its possible affiliation. The Khispi live in the village of that name and in Gompatse. Fresh field data collected in November 2011 make it clear that Chug, Lish and Gompatse all are essentially the same language. The people of these settlements deny any connection with the Mey of Rupa and Shergaon. Overall, Rupa consists of three major subsets, as shown in Table 1.

Table 1. Rupa subgroups

Subgroup	Lect
Rupa	Shergaon [sdp]
	Rupa [sdp]
Sartang	Rahung [onp]
	Jergaon [no code]
Northern	Khispi = Lish [lsh]
	Gompatse [lsh]
	Duhumbi = Chug [cvg]

The ISO codes are rather unsatisfactory. Shergaon and Rupa are sufficiently distinct as to warrant their own codes. Sartang is divided into five lects, although they are all mutually intelligible (Bodt 2021). As to Khispi and Duhumbi, they are extremely close and would elsewhere be regarded as dialects of one another.

Linguistic literature on the form of Mö spoken in Rupa town is sparse. The short description by Dondrup (1988) is based on the Shergaon dialect, while Grewal (1997) includes some sentences in the dialect of Rupa. The main source is Jacquesson (2015), which has considerable issues, due to its idiosyncratic transcription. Bodt (2014) is a literature review, and Boro (2024) has published a preliminary phonology.

The text of Abraham & Kara (2021) treats Sartang, Duhumbi (their Chug) and Khispi (their Lish) as separate languages. This is not supported by the comparative wordlist in Table 2, which shows that, allowing for variations in transcription, Khispi and Duhumbi are hardly even dialects of one another. Bodt (2020) is a grammar of Duhumbi, written according to modern linguistic norms. Surprisingly, Rupa is quite distinct from the language of Shergaon. The Sartang forms given below are based on newly transcribed field data.[6] Where the Mö cluster item resembles reconstructed CTB, the line is shaded.

[6] Roger Blench would like to thank the Gaonbura of Rahung for recording a wordlist of Sartang on January 18th, 2011.

Table 2. Comparison of Mö cluster languages with CTB

Gloss	CTB	Duhumbi	Khispi	Sartang	Rupa	Shergaon
One	*g-t(y)ik	hin	hin	han	han ~ ãi	han
Two	*g-ni-s	niʃ	ɲes	niʃ	ɲik	ɲit
Three	*g-sum	om	ʔum	um	uŋ	uŋ
Four	*b-ləy	psi	pʰəhi	pʃi	bsi	phsi

Gloss	CTB	Duhumbi	Khispi	Sartang	Rupa	Shergaon
Five	*b-ŋa	kʰa	kʰa	kʰu	kʰu	kʰu
Six	*d-ruk	ʧyk	ʧʰuʔ	ʧy	kit	ʧuk
Seven	*s-nis	his	ʃis	siʔ	sit	sit
Eight	*b-r-gyat	saɾgeʔ	saɾgeʔ	sardʑe	sardʑat	sargyat
Nine	*d-gəw	t̪ʰikʰu	t̪ʰikʰu	tʰekʰe	dʰikʰi	tʰikʰi
Ten	*gip	ʃan	ʃan	sou	sõ	sõ
Head	*d-bu-s	kʰloʔ	kʰoloʔ	kʰruʔ	kʰruk	kʰruk
Nose	*na, *naar	heŋpʰoŋ	hempoŋ	apʰuŋ	nəfuŋ	nupʰuŋ
Eye	*mik	kʰum	kʰumu	kʰaʔby	kivi	khibi
Mouth	*mka	kʰoʧu	hoʧok	ʧʰo	nəʧaw	niʧaw
Ear	*r-na	kʰutʰuŋ	kʰutʰuŋ	kʰətʰyŋ	gtʰiŋ	kʰutʰuŋ
Tongue	*s-l(y)a	loi	loi	le	lapon	laphõ
Tooth	*swa	hintuŋ	ʃiŋtuŋ	nitʰiŋ	tokʧe	nuthuŋ
Arm	*g-lak	hut	hu	ik	ik	ik
Leg	*kaŋ	lai	lei	le	la	la
Stomach	*grwat	hiliŋ	hiɲiŋ	fəriŋ	sliŋ	siriŋ
Bone	*rus	ʃukuʃ	ʃukuʃ	skiʔ	skik	skit
Blood	*s-hywey	hoi	hoi	he	ha	ha
Sun	*nəy	nami	nami	nimiʔ	nini	nini
Moon	*s-la	atnamba	namba	namluʔ	namblu	namblu
Star	*s-kar	karma	karma	ʧydʑy	zik	ʧuzuk
Man	*r-min	pədəŋ	bŭɖŭn	dʑiriŋ	ʃirin	dʑuhu
Woman	*mow	dʰudma	esma	dʑymy	kʰre dʑimi	dʑimi
Dog	*kʷəy	watʰi	watʰi	petʰe	btʰa	pʰitʰa
Pig	*pwak	ʃiabaʔ	ʃaba	swaʔ	swok	swag
Tiger	*k-la	lapʧa	pʰuyam	pʰuŋ	pʰuŋ	phõ
Water	*ti(y)	kʰu	kʰau	kʰow	kʰo	kʰo
Fire	*mey	bei	bei	be	ba	ba
Tree	*siŋ, *sik	ʃiŋ	hiŋ	hiŋ	siŋtiŋ	hiŋ tʰuŋ

Gloss	CTB	Duhumbi	Khispi	Sartang	Rupa	Shergaon
Leaf	*r-pak	ulaʔ	ulap	arap	alap	alap
Name	*miŋ	biŋ	biŋ		adʒen	
Eat	*dzya	ʧʰa	ʧa	he	ʧuva, kuva	ʧuwa, kuwa

Table 2 shows that Shergaon, Rupa, Sartang, Khispi and Duhumbi form a dialect complex and that resemblances to reconstructed CTB are sporadic.

5. BUGUN AND THE PROPOSED KHO-BWA CLUSTER

The Bugun language [bgg; glottocode bugu1246] is spoken in West Kameng district of Arunachal Pradesh. The Bugun, also known as Khowa, live in some ten villages, were estimated to number 800 in 1981, but current estimates put them at around 1700 speakers.[7] The Bugun language has been barely documented. The only published source is the orthographic Dondrup (1990) which should be used with care; some phonetically transcribed data appears in the Appendix to Abraham et al. (2005) and Madhumita Barbora of Tezpur University has recorded a wordlist and sample sentences as part of an unpublished study of the phonology. Lander-Portnoy (2013) is a thesis based on recorded material. Data for this paper was recorded from Martin Glo, secretary of the Bugun Welfare Society, who is a native of Chittu village, in Tenga in January 2011.

[7] Glottolog lists no less than six lects of Bugun, which seems unlikely. These are village names.

Pandey (1996) is part descriptive ethnography, part hagiography, and again should be used with care. Despite being a small ethnolinguistic group, the Bugun are quite active in promoting their culture with an active Bugun Youth Association. Bugun may be the only language in this region to have contributed a loanword into English. The Bugun liocichla (*Liocichla bugunorum*) is an endemic bird species first described in 2006. Vanessa Cholez (pers. comm.) has completed a dissertation (2024) on the sociology of the Bugun, but it is currently not in circulation.

Inasmuch as Bugun is mentioned at all, it is assumed to be Trans-Himalayan (e.g. Ethnologue 2024). van Driem (2001:473) originally referred to unpublished and unavailable work by Roland Ruttger relating Bugun to the Shertukpen cluster suggesting this and names the resultant grouping 'Kho-Bwa'. This has been enthusiastically promoted by Bodt (2019, 2020, 2022) and even used in experiments to predict relatedness using algorithms (Bodt & List 2019; Wu et al. 2020). However, the existence of this construct is far from proven, and I suggest here it may be a chimaera. Bugun people are often able to speak Mö as a language of intercommunication and the similarities between the two may simply be borrowings. Table 3 suggests neither a regular relationship between Bugun and Mö, nor a strong resemblance to reconstructed CTB. Some words show relations of near-identity, for example 'head', 'water' and 'leaf'. /ʧ/ is conserved in 'eat' and 'liver' but /ʃ/ corresponds to /ʧ/ in 'mouth'.

On the broader question of whether Bugun and the Mey cluster are Trans-Himalayan, neither language

shows many cognates with CTB and some of those are doubtful or possible loans, such as 'pig' and 'iron'. The low number of Tibeto-Burman cognates could just as easily be explained by borrowings as by genetic affiliation. Table 3 shows some of the typical resemblances and a more in-depth search would be likely to uncover great numbers. I have marked the CTB form in the second column; it shows that only a very few forms are shared with the Bugun-Mey pair.

Table 3. Bugun-Mö cluster resemblances

Gloss	CTB	Bugun	Lish	Sartang	Rupa	Shergaon	Comment
Two	*g-ni-s	ɲeŋ	ɲes	niʃ	ɲik	ɲit	
Three	*g-sum	ɨm	ʔum	um	uŋ	uŋ	
Five	*b-ŋa	kua	kʰa	kʰu	kʰu	kʰu	
Nine	*d-gəw	dige	t̪ikʰu	tʰekʰe	dʰikʰi	tʰikʰi	
Ten	*gip	suŋwa	ʃan	sou	sõ	sõ	
Head	*d-bu-s	kʰruk	kʰoloʔ	kʰruʔ	kʰruk	kʰruk	
Nose	*na, *naar	əfuŋ	hempoɲ	mapŭn	nəfuŋ	nupʰuŋ	
Mouth	*mka	ʃyam	hotʃok	tʃʰo	nətʃaw	nitʃaw	
Ear	*r-na	kʰɔɔ̃́	kʰutʰuŋ	kʰətʰyŋ	gtʰiŋ	kʰutʰuŋ	
Leg	*kaŋ	loy	lei	le	la	la	
Penis	*m.ley	lo		lok			cf. Tani *mrak
Blood	*s.hywey	əfoy		hoi		ha	
Liver	*m.sin	ətʃiŋ		atʃĭ		atʃẽ̃	
Fat	*tsil	eyòó		ayùú		oyo	
Heart	*s.niŋ	ɛliŋ		zli			cf. Miji luŋ
Saliva	*m/s.tswa	tɕan		tɛɛ		taa	
Child	*za/*tsa	ani		nini		nunu	
Pig	*pwak	wak	ʃaba	swaʔ	swok	swag	
Water	*ti(y)	kʰo	kʰau	kʰow	kʰo	kʰo	
Fire	*mey	boe	bei	be	ba	ba	
Tree	*siŋ, *sik	hiŋmua	hiŋ	hiŋ	siŋtiŋ	hiŋ tʰuŋ	
Leaf	*r-pak	arap	ulap	arap	alap	alap	
Eat	*dzya	tʃʰa	tʃa	he	tʃuva, kuva	tʃuwa, kuwa	

6. PUROIK [= SULUNG]

The Puroik language [suv: glottocodes sulu1241 and west2872] is spoken by a few thousand people in East Kameng and Lower Subansiri districts in Arunachal Pradesh and adjacent parts of Tibet. Previously known as 'Sulung', this name has now been rejected by the community as pejorative. The ethnography of the Puroik is described in Stonor (1952) and Deuri (1982). Their basic subsistence system is hunting and gathering with a significant dependence on the sago palm, *Metroxylon*, rather as in Melanesia. All forms of agriculture appear to be recent innovations. The Puroik were formerly in a serf-like relationship with the Tani-speaking Nyishi, for whom they collect cane and labour on farms. Puroik were still being officially liberated as late as 2001 (see appended documents in Remsangphuia 2008:102-102).

The most well-described of the three Puroik varieties is the dialect of Chayangtajo circle, East Kameng, where Sanchu is the biggest and best accessible Puroik village. This variety is described in the major published sources, Tayeng (1990), Li (2004), Remsangphuia (2008) and Soja (2009). The Chinese work is summarised in Matisoff (2009) and reviewed in Jackson (2003). Matisoff (2009) has an appendix to a paper on the persistence of Tibeto-Burman roots, compares Puroik materials from Li (2004) with his CTB roots and claims numerous cognates. Many of these require the eye of faith but it is notable that there are more resemblances than are evident in southern forms which often have quite different lexemes. Since the Tibetan Puroik apparently also speak Tibetan as a second language, this should make us suspicious at the least. Work by Lieberherr (2017), assuming it is published, is likely to give us a much improved description of Puroik. The dialect of Kojo-Rojo is spoken in two, possibly three villages (Kojo, Rojo, Jarkam), and is different but mutually intelligible with the dialect of other villages in Lada circle. The third dialect is Bulu, west of Kojo-Rojo. Map 1 (from Lieberherr 2015) shows the locations of these dialects as well as neighbouring languages mentioned in this paper.

Although listed both as Kho-Bwa and 'possibly Austroasiatic' in earlier versions of the Ethnologue (e.g. 2013), the arguments for this are elusive. Concerning the classification of Puroik, a footnote to Sun (1993: fn. 14) says; 'Sulung is a newly discovered distinct Tibeto-Burman language showing remarkable similarities to Bugun, another obscure Tibeto-Burman language spoken to the west of the Sulung country.' This is a considerable exaggeration, and later, reviewing the Chinese source, Sun (1992) assumes that Puroik is Trans-Himalayan, he is pessimistic about finding the evidence for cognates. The most detailed examination of these similarities, as well as a discussion of the dialect situation of Puroik is Lieberherr (2015). This paper takes on the challenge presented in Blench & Post (2014) to show that the apparent cognates with Trans-Himalayan are in fact evidence of genetic affiliation and not simply borrowings. His method is slightly idiosyncratic, since he compares Puroik with reconstructed Kuki-Chin (VanBik 2009) rather than CTB. Kuki-Chin is certainly an authenticated branch of Trans-Himalayan, and Lieberherr's arguments are coherent. As a consequence, I regard the argument for a Trans-Himalayan affiliation as generally convincing and my prior scepticism as refuted.

A separate question is whether Puroik, Bugun and the Mö cluster form a linguistic group. If so, they would then all be Trans-Himalayan. Apart from the numerals and some body parts, the general form of the vocabulary is highly divergent and suggests the Kho-Bwa construct is not meaningful. Table 4 shows a preliminary table of lexical similarities, including Mö cognates (shaded):

Table 4. Puroik comparison with Bugun and Mö

Gloss	Puroik W.	Puroik E.	Bugun	Mö of Rupa
Two	niʔ	nii	ɲeŋ	ɲit
Seven	mə-lyɛɛ	lyɛɛ	milye	sit
Eight	mə-lyao	laa	mla	sargyat
Nine	duŋgii	doŋɹɛɛ	dige	tʰikʰi
Mouth	səm	sək	ʃyam	ni.ʧaw
Nose	poŋ	pok	e.pʰuŋ	a.pʰuŋ
Leg	a-lɛɛ	lae	loy	la
Stomach	a-ɬye-[buŋ]	a-ɬue [buk]	lui	siriŋ
Man	a-fuu	afuu	b.pʰua	ʤuhu
Woman	məruu	amui	bimi	ʤimi
Water	kɔɔ	kua	kʰo	kho
Fire	bɛɛ	bɛɛ	boe	ba
Dream v.	baŋ	bak	baŋ	baŋ

7. HRUSO [= AKA]

Hruso [hru: glottocode hrus1242] is also called Aka, although this name means 'painted', presumably a reference to facial tattooing which was formerly practised by these people. This name is still used in official publications such as Ethnologue, and Hruso is the self-name which has begun to come into use within the community. Aka is sometimes confused with Koro, which, although also called 'Aka' by the local Miji people, has now been assigned a distinct glottocode (koro1316). Hruso has been the subject of a considerable literature. A reference to the Hruso people dates to Robinson (1855). The first linguistic information is due to Needham (1886), reprised in Campbell (1874) and in Konow (1909). These sources contain substantial wordlists, although in some cases they are remarkably transcribed, while in Konow, the transcription of other languages in his survey is of a high standard. The first substantial monograph is due to Schubert (1964) and a second version of this is reprinted in Shafer (1966/1967:245-277). Modern sources include the orthographic grammar by Simon (1993). More recently, D'Souza (2015, 2018, 2021) is a detailed phonology of Hruso. The author has recorded Hruso from Sera Saring in Serba on several occasions since 2011.[8]

[8] Roger Blench would like to thank Serwa for being an enthusiastic and patient informant for Hruso.

Shafer (1947) is mainly devoted to a discussion of the classification of Hruso. He considered Hruso was Trans-Himalayan on no very good grounds, both because of the poor quality of trancriptions and because

his discussion conflates Hruso proper with Miji. Cognates with Trans-Himalayan languages are very few and involve sometimes highly ad hoc rules. Koro Aka is quite unrelated to either of these, as the appendix table shows; its affiliation is dealt with in §8. Of all the languages considered here, Hruso has the fewest roots that can plausibly be related to Trans-Himalayan. Shafer's arguments are fairly weak, and it is more credible to treat these as regional borrowing than evidence for genetic affiliation.

8. KORO AND THE POSSIBILITY OF A SIANGIC PHYLUM

The references to two Aka languages has been a source of confusion since the earliest period and indeed there are two languages under this label, Hruso proper and Koro. The Koro [jkr: glottocode koro1316] people live in East Kameng district, in Bana and surrounding hamlets, east of Thrizhino on the Seppa road. Although claims were made for its 'discovery' in 2010 (Anderson & Murmu 2010), presumably as a consequence of financing by National Geographic, a grammar sketch of this language appears in Grewal (1997) and lexical data can be extracted from Abraham et al. (2005/2021). Further lexical data was collected by the author in December 2011.[9]

[9] Roger Blench would like to thank the headman, Somo Yamde, for taking time to record a sample of Koro in Yangse village in November 2011.

A brief comparison with Hruso quickly shows that the two have virtually nothing in common, as was also stated by Anderson & Murmu (2010). However, strikingly, Koro does share a number of lexemes with Milang, a language far to the east in Siang district usually identified as Tani (Sun 1993: §3). Milang is spoken in three dispersed villages in East Siang district (Modi 2008). With the exception of Tayeng (1976) almost nothing has appeared in print on this language. Sun noted its divergent character, but treated it as an early branching of Tani. Milang is characterised by both divergent lexicon and highly irregular correspondences with the rest of Tani (Modi 2008, Post and Modi 2011). The hypothesis here is that Milang was a non-Tani language that came under heavy and repeated Tani influence. Milang is spoken a considerable distance from Koro, so shared lexicon is unlikely to be the result of contact. The proposal, set out in more detail in Post & Blench (2011), suggests there was once a chain of languages, tentatively named Siangic, stretching between West Kameng and the Siang river, whose presence can be detected both in Koro, in the substrate lexicon of Milang and in irregularities in other Tani languages now spoken in the intervening area (Table 5).

Table 5. Comparative Siangic

Semantics	Gloss	*PS	Koro	Milang	Proto-Tani	Adi	Other TB
food	cultivated field	*pu	pu	a-pu	*rɨk	a-rɨk	n/a
crops	rice paddy	*k(h)ɨ	kiraka	du-kɨ	*ma ~ mo ~ pɨm ~ am	amo, ambɨn, apin	*ma ~ *mey
	bamboo	*fu	fu	a-hu	*ɦɨə	eŋ	N/A
animals	chicken	*cjo	co-le	a-cu	*rok	pə-rok	N/A

Semantics	Gloss	*PS	Koro	Milang	Proto-Tani	Adi	Other TB
	egg	*(cjo)-ci	cu-ci	ci-ci	*pɨ	(rok-)pɨ	*ʔu, *t(w)i(y) (< water?)
	mithun	*su	sù	a-su	*a-so	ə-so	N/A
	bird	*pju	po-le	ta-pju	*pa-taŋ	pə-ttaŋ	WT/PLB *bya, Jinuo pyɔ
nature	sun	*mə(y)	me-ne	məə-ruŋ	*doŋ-ɲi	doo-ɲi	PLB *məw (Lahu mû)
	day	*nə(y)	me-ne	a-nə	*loŋ	loŋ-ə	PTB *nə(y) (Tib., Bur.)
	yesterday	*ba-nə(y)	ba-n(e)	ba-nə	*mə-lo	mə-lo	N/A
	fire	*mi	mi-la	a-mi	*a-mə	ə-mə	PTB *mey
	stone	*bu	u-bu	da-bu	*liŋ	ə-liŋ	PTB *luŋ
numerals	two	*nə(y)	(ki-)ne	nə	*ɲi	a-ɲi	PTB *ni
	seven	*rVŋ(?)	rõ	ra-ŋal	*kVnV(t)	kənɨt	PTB *ni
	eight	*ra(N)	rã-la	ra-jəŋ	*pri-ɲi	piiɲɨ	PTB *b-g-r̥yat (unlikely cognate)
body	ear	*raŋ(u)	rã	ra-ɲu	*ɲa(-ruŋ)	ɲo-ruŋ	PTB *na (widely attested). Some Tani (e.g. Bokar) has narun, apparently metathesis)
	vagina	*ce(k)	cek	a-cci	*ti(ɨ)	ɨttə (Galo)	PTB *s-tu (Lai Chin chu) ?
	neck	*laŋ	lã	a-laŋ	*a-liŋ	a-liŋ	Although given as PTB *liŋ, evidence very weak
	beard	*kjaŋ-mV	caa-mi	kjaŋ-ma	*napmɨt	nam-mɨt	initial N/A, final common *mil/mul/myal
	foot/leg	*bja	ni-bi	a-bja	*bjaŋ 'thigh'	ar-baa (Galo)	N/A
	boy	*ma	ma-le	jaa-ma	*meŋ	jaa-meŋ	N/A?
colour	green/blue	*ja-caŋ	jã-ca	jə-caŋ	n/a	ja-zee (Galo)	OC sĕŋ 'fresh', Jingpho tsīŋ 'grass', Garo thaŋ 'alive/green/raw' ?
	red	*laŋ	lã	jə-laŋ	*ja-lɨŋ	ja-lɨŋ	N/A
	arrow	*pa	pa	a-ppa	*a-puk	ə-puk	N/A
	ladder	*b(r)ja	i-bi	da-bja	*lə-braŋ	lə-bjaŋ	N/A
functors	negative verb suffix	*-ŋa	-ŋa	-ŋə	*maŋ	-maŋ	PTB *ma

Semantics	Gloss	*PS	Koro	Milang	Proto-Tani	Adi	Other TB
	locative	*la	la	l(a)	*lo	lo	PTB *la
	desiderative[10]	*-mi	-mi	-mi	*-liŋ	-liŋ	?
verbs	cut	*pi	pi	pi	*pa	pa	No cognate forms
	have (be there)	*kjo	ko	cu	*ka	ka-	N/A
	give	*ram	rã	ram	*bi	bi	PTB *bəy
	know	*fu	fu	hu	*ken	ken	PTB *kyən (WT mkhen)
	eat	*tju	to	tu	*do	do	PTB *dzya (unlikely cognate)
	imbibe (drink/ smoke)	*caŋ	ca	caŋ	*tɯŋ	tɨŋ	no cognates

[10] Seemingly only when negated in Koro.
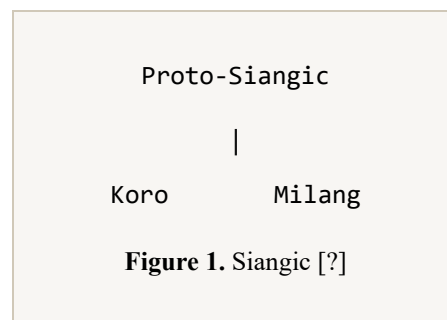

Figure 1. Siangic [?]

Figure 1 represents the configuration of the proposed Siangic phylum. If this argument is correct, then Siangic is a small phylum which is distinct from Trans-Himalayan. Milang underwent heavy cultural influence from Tani (Adi and Padam in particular), making it appear a divergent Tani language, but underlyingly it is unrelated.

9. MIJIIC

The Miji language (also Sajolang[11], Dimai [= Dhimmai]) should not be confused with Dhimal in Nepal. Simon (n.d.) reports that it is spoken in some thirteen villages around Nafra in West Kameng district and that the population was 3,549 in the 1971 census. Fresh fieldwork was undertaken among the Miji of Nafra in February and March 2024[12] which revealed serious errors of elicitation in previous work, which mixed forms from two lects. This is now corrected in more recently circulated documents.
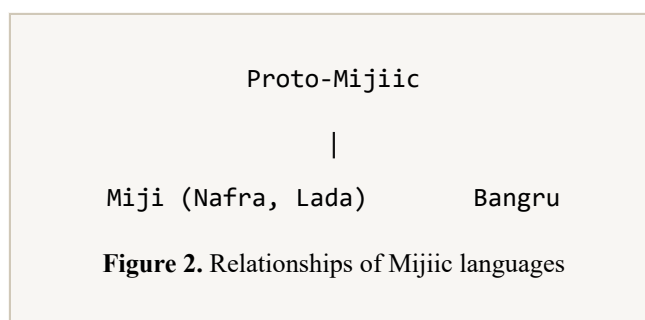
[11] The name Sajolang has been widely adopted in Nafra in reference to Western Miji, but is not in use among the Eastern Miji in Lada circle.

Miji is divided into two lects, representing Nafra and Lada circles, effectively west and east. Weedall (2021) has a more complex division of western Miji into subdialects. Whether these would be naturally mutually comprehensible is hard to determine, since almost all speakers seem to know the equivalent words in the other dialect.

Until recently, the existence of a language known as Bangru [no ISO glottocode, bang1369] was only rumoured. However an ethnographic thesis and subsequent report confirms that this is a genuine ethnic group, but also that it is a language related to Miji (Ramya 2011, 2012). Ramya's transcriptions are orthographic, but the underlying forms are easily seen when compared to Miji. Blench (2015) circulated comparative data on Miji and Bangru in support of this assumption. Bodt & Lieberherr (2015) have published a wordlist of Bangru based on the CALMSEA list, with analysis and comparisons with Miji and Hruso. While this represents a major advance on Ramya, much more remains to be done on Bangru. Separately, I compiled a wordlist with the aid of Mrs. Chera Mekia Mili and family, now resident in Itanagar. Mrs. Mili grew up in a monoglot household, but later married a Nyishi husband and came to speak the language as the main language of her home.

Andrew Hsiu kindly forwarded to me Li (2003) which is a phonology and wordlist of the Bangru spoken in Tibet which was unknown to all previous researchers. Li includes phonemes that seem to have disappeared from the Bangru of India, and also envisages a more elaborate tonal system. This may represent a more conservative form of the language and Bangru in India has undergone heavy influence from the dominant Nyishi language.

```
              Proto-Mijiic

                   |

Miji (Nafra, Lada)        Bangru
```
Figure 2. Relationships of Mijiic languages

Bangru is undoubtedly related to Miji, as the numerals in the appendix table show, although it has quite a divergent lexicon. In some cases, this is due to borrowing from Nah and Nyishi, both Tani languages. Figure 2 shows a tree of the relationships of Mijiic languages.

There is a persistent idea in the literature that Miji is related to its eastern neighbour Hruso and that there exists a subgrouping 'Hrusish'. More surprising are the cases where Bangru shows similarities to Hruso. This idea seems to derive from Shafer (1947) but is perpetuated in Bodt (2014) and Bodt & Lieberherr

(2015). The evidence for Hruso and Miji having a distinctive relationship seems to be based on a very few similarities, for example the words for 'sun', 'eight' and 'nine', which appear to be exclusively shared. The great majority of basic vocabulary appears to be quite different. Miji, Hruso and Koro share a great deal culturally, but show virtually no linguistic resemblances, except for sporadic loanwords, as might be expected between any two neighbouring languages. This can be clearly seen in the appendix table. The conclusion must be that the purported Hrusish group is spurious, and Hruso is a language isolate.

Miji has long been classified with the 'North Assam' languages and this is generally repeated in subsequent publications (e.g. Bradley 1997). However, there is no data in Konow (1909) and Shafer (1947) may be the first citation of linguistic material. As is now a rather familiar pattern, resemblances to Trans-Himalayan are confined to a few lower numerals and some basic body parts. Otherwise, few Miji lexemes resemble reconstructed Trans-Himalayan forms. I therefore argue that the Mijiic languages constitute a small independent phylum.

10. MISHMIC (~ DIGARISH)

It has been proposed that there is a Mishmic group consisting of two related languages, Idu (Luoba in Chinese sources) and Tawrã (= Taraon, Digaru, Daruang in Chinese sources). This group has sometimes been known as Digarish, or alternatively grouped together with Kman (= Miju) as Midźuish, a denomination which may go back to Shafer (1955). It is certainly the case that the Idu and the Tawrã refer to themselves in local English as 'Mishmi'. In modern times, however, it is the Tawrã and the Kman who regard themselves as a single culture, despite the fact that their languages seem to have almost nothing in common. There is a plethora of local publications which compare phrases in both languages. Quite how this cultural convergence came about is unclear, but see speculations in Blench (2024). Kman is treated in the following section as a distinct language whose classification remains unclear.

The Idu [clk: glottocode idum1241] are also known as Chulikata [= Chulikotta, Sulikota], Midu [= Ida, Midhi], Yidu Luoba, Lhoba [Chinese terms]. It is unfortunate that the ISO code is based on the pejorative term Chulikata, now discouraged. The earliest reference to the Idu language is in Brown (1837) and some material can be found in Campbell (1874) and Konow (1909). Pulu (2002) is a brief orthographic introduction to Idu in the characteristic style of Arunachal Pradesh Government publications. Idu has also been described from the Chinese side [under the name Lhoba], notably in Sun et al. (1980), Sun (1983a,b, 1999) and Ouyang (1985).

The Tawrã [mhu; glottocode: diga1241] are also known as Darang 达让僜, Daruang, Deng, Digaro, Digaru, Mishmi, Taaon, Taraon and Taying. The name 'Digaru' (vernacular name of a major river) is often used in English conversation. Records of Tawrã go back at least to Robinson (1856). Needham (1886) gives a comparative wordlist of Tawrã, Kman and Tibetan. Recensions of existing data are given in Campbell (1874) and Konow (1909). Modern publications with a 'practical' orientation include Chakravarty (1978) and Pulu (1991). Chinese scholars have also worked on Tawrã, which they call 'Daruang'. The Tawrã language has been briefly described in Sun (1999) and Jiang et al. (2013) is an extended grammar of Tawrã in Chinese. Evans & Manyu (2021) is a phonology of Tawrã in India prepared for the purposes of Bible translation, so its reliability is hard to gauge.

Whether Idu and Tawrã actually form the genetic unit claimed in the literature is questionable. They clearly share a significant amount of lexicon in some semantic fields. For example, Table 6 shows the lower numerals, which suggest a strong relationship.

Table 6. Idu and Tawrã lower numerals

Gloss	Idu	Tawrã
One	khə̀gə̀	khin
Two	kà.nyì	kayiŋ
Three	kà.sɔ̃̀	kasaŋ
Four	kà.prì	kaprayk
Five	màŋá	maŋa
Six	tāhrō	tahro
Seven	íṵ̀	wẽ
Eight	ìɭú	lɨm
Nine	khrìnī	kɨɲaŋ
Ten	hṵ̃̄ṵ̃̄	hálaŋ

By way of contrast, Table 7 compares Idu and Tawrã terms for body parts, few of which have anything in common.

Table 7. Idu and Tawrã body parts

Gloss	Idu	Tawrã
back	ìpìndò	phlíŋ
body	jóntà	kyàŋ
breast	nōbrā	ɲèè
eye	ēlōbrā	b.lm
hand	ākhó	hàprə̀
leg	āŋgēsà	gròn
lip	īnūbrū	thánù
mouth	ēkóbə̀	phùùkɛ̀
nail	āhṵ̃kò	áphlìŋ
neck	sēmbrá	pà hŋ

Gloss	Idu	Tawrã
nose	ēnāmbó	àɲàdùn
palm	lāpū	àtyòpà
skin	kòprà	pô
thigh	hàpū	sàhà
toe	ātāmbó	gròn bràn
tongue	īlìná	hèlèŋnà
tooth	tāmbrō	là

Given this divergence by semantic field, making any definitive statement about the relatedness of Idu and Tawrã is problematic. Blench (2024) argues that the two languages are ultimately unrelated, and that there has been intense bilingualism at an unknown period in the past, which resulted in the convergence of the numerals. The grammar of Tawrã is poorly known, so it is difficult to compare it with Idu, which is relatively well-described. Perhaps further in-depth studies will clarify the situation.

Whether Idu and Tawrã are Trans-Himalayan remains doubtful. The appendix data table shows that there are few evident cognates with reconstructed CTB. The pattern is much the same as noted for other languages, i.e. lower numerals, some basic body parts, sun, moon etc. I submit this is inadequate to accept as evidence for membership of Trans-Himalayan.

11. KMAN [= MIJU]

The Kman people [mxj: miju1243] live in villages around Tezu in the extreme northeast of Arunachal Pradesh.[13] Alternative names for the Kman include Eastern Mishmi, Geman Deng, Kaman, Miju. The first record of Kman appears to be Robinson (1856) which is quite accurate for the period, and his transcriptions are recognisable today. Orthographic publications on the Indian side are Das Gupta (1977) and Boro (1978) which are said to be 'practical' although the transcription of Kman is highly inaccurate by modern standards. Despite the small number of speakers on the Chinese side of the border, there have been several publications on 'Geman', the Chinese version of the name. These include Sun (1991, 1999) and most importantly, Jiang et al. (2013) which is a full-length description of the language. Kman has undergone an intriguing development in terms of its orthography; a local system of writing used for communication on Facebook has developed which is also used in a children's book (Dai et al. 2013). Separately, lexical guides have been published covering both Kman and the neighbouring Tawrã [Taraon] language (Krisikro 2006; Tawsik 2014) although the orthography bears no resemblance to any other publication. The Kman people have been the subject of an anthropological monograph (Dutta 2012).

[13] Fieldwork among the Kman began in 2015 and has continued until 2024 in collaboration with Sokhep Kri. A preliminary Android dictionary has been released in the community.

The evidence for the affiliation of Kman to Trans-Himalayan is more convincing than for Idu and Tawrã. The appendix table shows many more clear cognates with reconstructed CTB, suggesting a closer affiliation. Another aspect of Kman which shows analogies with more established Trans-Himalayan languages is in the morphology of the verb, verb stem alternation which has numerous parallels in regional languages. In the case of Kman, number marking is indexed to the head through nasal alternation in the stem. This process is only applied to a subset of verbs and is thus not predictable. The output is also atypical, since the result is verbs with final N+C sequences which do not occur elsewhere in the phonology. In Kman grammar, these are verbal nouns or gerunds, since they primary occur with action verbs. A sample of these gerunds is given in Table 8.

Table 8. Verb stems with nasal incorporation

Singular	Nasal	Gloss
ø→ -m		
dõ	dõm	saying
kà	kàm	be, is
tə̄rà	tə̄ràm	repairing
-l→ -m		
brə̄l	brə̄m	falling from a height
gyā	gyām	running
-y→ -m		
tə̄y	tẽm	going
ø→ -m		
lə̄p	lə̄mp	sitting
gyũp	gyũmp	cheating
shʉ̀p	shʉ̀mp	buying
ø→ -n		
chawk	chawnk	dwelling
kə̄t	kənt	doing
krĩt	krĩnt	laughing
thũt	thünt	blowing
thə̄k	thə̄nk	obeying
ʔ→ -nk		
phlôʔ	phlõnk	being late

Singular	Nasal	Gloss
phlûʔ	phlūnk	jumping

This type of alternation, which is sporadic and unpredictable, can be compared to verb stem alternation in other Trans-Himalayan languages, for example Tangsa (Morey 2018) and Kuki-Chin (Bedell et al. 2023). This type of morphosyntax is far less likely to be borrowed than the lexicon.

12. MEYOR [ZAKHRING]

The Meyor language [zkr: glottocode zakh1243], also known as Zakhring, is spoken in Lohit District, Walong and Kibithoo circles, Arunachal Pradesh. In 2001 there were some 376 speakers scattered in fifteen hamlets. On the Arunachal Pradesh side, the major published source on the language is Landi (2005) although Jacquesson (2001) includes some data on pronouns. Lǐ & Jiang Di 李大勤，江荻 (2001) is a brief overview of the 'Zha' language. Sun (1999) has comparative tables of language he calls 'Dza' which he relates to 'Geman' [Kman] and assumes it is a 'mixture' with Tibetan. Geman (together with Idu-Tawrã) is said to be Kachinic, although no evidence is presented for this. Both Ethnologue and Glottolog put Kman and Meyor together as a subgroup of Trans-Himalayan, but the evidence for this is thin.

Landi (2005: 164 ff.) notes the similarities to Kman, although he conflates genuine shared cognates due to common CTB inheritance with borrowings. Nonetheless, some useful observations can be extracted from his tables. Table 9 shows a sample of lexical items where Kman and Meyor share a common root. The Dza data shows more cognates with Kman than the material in Landi (2005).

Table 9. Kman-Meyor common roots

Gloss	Meyor	Kman
arrow	lowat	roowat
ask	want	wat
bear	ʧʃam	ʧʃim
beer	si	si
bird	awa	oowa
blood	awi	iwi
claw	ʧʃan	ʧʃan
comb	sipiŋ	sipin
granary	keetam	katam
hair	sam	syam
honey	ʃam	ʃamti

Gloss	Meyor	Kman
lock	dimik	ʤimik
melt	yulo	yu
mouse	aʃi	si
meat	ʧin	ʃin

These items are suspiciously similar, whereas Kman and Meyor otherwise have many completely divergent roots, suggesting borrowing. Considerably more Meyor basic lexicon is related to the Brokpa languages, such as Memba and the Senge cluster (represented in the appendix table by Tawang and Memba dialect). These languages have a relatively high proportion of Tibeto-Burman roots, preserved in a constellation very close to the hypothetical proto-form.

Landi also compares Meyor to Turung (Singpho), a Jingpho language spoken in this region, but his comparisons are all either only doubtfully cognate or are CTB and thus not relevant as evidence. Scott Delancey (p.c.) has also presented evidence for some striking grammatical similarities with Kuki-Chin type languages, in the area of pronominal indexing. Yet an examination of the lexicon using the data posted in STEDT did not produce a single example of a specific link to the Naga/Kuki-Chin languages. Meyor must definitely be left unclassified at present and indeed represents a broad problem for the usual procedures of historical linguistics.

13. SYNTHESIS

The emerging synthesis is quite strikingly at variance with the received and published classifications. Arunachal Pradesh languages are extremely diverse, not only in relation to common Trans-Himalayan but also to one another. As we learn more about their grammar, it is clear that although they are synchronically isolating languages with eroded morphology, they reveal remarkable and isolated traits, quite unlike their neighbours. Idu, for example, has a complex verbal system with lengthy stacked extensions, similar to Bantu languages of Central Africa.

The usual explanation for a low incidence of CTB vocabulary is the gradual loss of lexemes over time. What is striking here is that the presumably innovative forms—the diachronically secondary forms, according to the received view—are both (a) far greater in number than the attested CTB forms and (b) not (or not obviously) relatable to any other known language. The implications of this linguistic model for proto-historical reconstruction are extreme, and should be made plain: we are asked to believe that individual Tibeto-Burman language groups repeatedly encountered populations which so overwhelmed them that they adopted forms from these mystery languages on such a scale that the overwhelming majority of their lexicons were wholly replaced, and that these mystery languages subsequently died out, leaving only the previously marginal genetically Tibeto-Burman languages to reflect their past existence in the form of an overwhelmingly massive substrate. Why precisely this model is more persuasive than one in which it is rather a suite of non-Tibeto-Burman languages which, coming into contact with different Tibeto-Burman languages at various points in their history, adopted a handful of Tibeto-Burman forms, remains

to be demonstrated. The model adopted here is more in tune with modern contact linguistics, assuming borrowing unless inheritance is demonstrated.

In the light of this, Table 10 (next page) synthesises the new proposals presented here, omitting a detailed listing of Tibetic, Tani and Naga languages.

This represents a fairly radical departure from the conventional view of these languages. In another way, this is far from surprising. Arunachal Pradesh is highly dissected, remote and inaccessible and was by-passed by major East-West trade routes. That language isolates should have persisted here long after they were assimilated elsewhere in SE Asia is quite credible. The challenge for the future will either be to build on these hypotheses or disprove them on the basis of improved evidence.

Table 10. Proposed classification of Arunachal languages

Phylum	Branch	Language	ISO	Also
Trans-Himalayan	Jingpho	Turung	sgp	Singpho
Trans-Himalayan	Tibetic	Memba	mmc	But see text notes
	Tibetic	Brokpa	sgt	
	E. Bodish	Monpa of Tawang	dka	Dakpa, including Senge, Jang
	E. Bodish	Monpa of Zemithang	dzl	? Dzala (van Driem 2007)
	Tshangla	Monpa of Dirang, Murshing and Kalaktang	tsj	Sharchop, Tshangla
Isolate	Unclassified	Meyor	zkr	Zakhring
Trans-Himalayan	Tani	Numerous		Adi, Galo etc.
Trans-Himalayan	Tangsa Naga	Numerous		Lunchang, Jugli, Moklum, Changlang, Wancho, Nocte
Siangic [?]	Milang-Koro	Milang	none	Mala, Holon, Dalbəŋ [village names]
Isolate		Koro	jkr	Koro Aka
Isolate		Bugun	bgg	Khowa
Mö		Mö of Shergaon	sdp	Shergaon
		Mö of Rupa	sdp	Shertukpen
		Sartang	onp	But Monpa
		Khispi [= Lish]	bqh	dialect cluster with Duhumbi
		Duhumbi [= Chug]	cvg	forms a close dialect cluster with Khispi

Phylum	Branch	Language	ISO	Also
Isolate		Idu	clk	Idu Mishmi
Isolate		Tawrã	mhu	Digaru, Taraon
Mijiic		Miji	sjl	Sajalong, Dhimmai? northern dialect
		Bangru	none	
Trans-Himalayan		Puroik	suv	Sulung (pejor.)
Trans-Himalayan		Kman	mxj	Kaman, Geman, Miju
Isolate		Hruso	hru	Aka

14. CONCLUSIONS

The impetus behind this paper is the re-examination of the evidence for a Trans-Himalayan affiliation proposed for the languages of Arunachal Pradesh, in the light of the practice of repeating the work of previous scholars without an evaluation of the actual data. The conclusion is that a number of languages or clusters could well be isolates, and that the Trans-Himalayan roots they do evince may well be borrowings. Obviously, each topic requires a full-length paper, and these will be undertaken in due course, especially as better-transcribed data becomes available. Meanwhile, provisionally we may well consider Konow's summary quoted in the epigraph to be a useful image. Arunachal Pradesh consists of a chain of isolated languages, which have been on the southern edge of the core Trans-Himalayan area. A plethora of different contact situations have allowed both lexical borrowing and sometimes striking grammatical and phonological restructuring. But perhaps it would be useful to begin considering this region as more similar to the Amazon or Northeast Asia than Tibet.

In view of this, the languages of Arunachal Pradesh should be treated as a major priority on a global scale. Languages such as Basque and Burushaski have attracted high levels of scholarly interest over many decades precisely because of their status as language isolates. Those in Arunachal Pradesh have been completely bypassed. Moreover, although these languages are presently still spoken, their populations are small and pressure to switch to Hindi, promoted in both the media and via the school system, is growing. Probably by no coincidence, Arunachal Pradesh is also a major centre for biodiversity, something which attracts worldwide attention and resources. It is suggested that the little-known languages of Arunachal Pradesh should be given similar priority due to their uniqueness and endangered status.

REFERENCES

Abraham, Binny & Kara Sako. 2021. Tagin and Puroik Tribes of Arunachal Pradesh: A Sociolinguistic Survey. *Journal of Language Survey Reports* 2021 (063): 1-83. [This is an electronic publication of a report originally circulated as Abraham 2005.]

Aikhenvald, Alexandra Y. 2012. *Languages of the Amazon*. Oxford: Oxford University Press.

Anderson, Gregory D.S. 2014. On the classification of the Hruso (Aka) language. Handout, HLS, Singapore.

Anderson, Gregory D.S. and Ganesh Murmu. 2010. Preliminary notes on Koro, a "hidden" language of Arunachal Pradesh. *Journal of Indian Linguistics*, 71: 1–32.

Anderson, James D. 1896. A short vocabulary of the Aka language. Shillong: Assam Secretariat.

Bedell, George et al. 2023. The morphosyntax of verb stem alternation. *Himalayan Linguistics*, 22(1): 202-242.

Blench, Roger M. 2008. Re-evaluating the linguistic prehistory of South Asia. In: Toshiki Osada and Akinori Uesugi (eds.), *Occasional Paper 3: Linguistics, Archaeology and the Human Past*: 159-178. Kyoto: Indus Project, Research Institute for Humanity and Nature.

———. 2014. The contribution of linguistics to understanding the foraging/farming transition in NE India. In: Tia Toshi Jamir & Manjil Hazarika (eds.), *51 Years after Daojali Hading: Emerging perspectives in the Archaeology of Northeast India. Essays in Honour of T. C. Sharma*: 99-109. New Delhi: Research India Press.

———. 2020. The world turned upside down: sago-palm processors in Northeast India and the origins of Chinese civilisation. In Helen Lewis (ed.), *EurASEAA 14, Dublin 2012, proceedings*. Volume I: 190-202. Oxford: Archaeopress.

———. 2024. The 'Mishmi' languages, Idu, Tawrã and Kman: a mismatch between cultural and linguistic relations. In, Nishant Choksi (ed.), *Movements through Time and Space: Ecology and Lingua-Cultural Change in South and Southeast Asia*. Guwahati: Pragjyotish Centre for Cultural Research.

Blench, Roger M. and Mark W. Post. 2014. Rethinking Sino-Tibetan phylogeny from the perspective of Northeast Indian languages. In Nathan Hill & Tom Owen-Smith (eds.), *Selected papers from the 16th Himalayan Languages Symposium, September 2010*: 71-104. New York: Mouton de Gruyter.

Blust, Robert A. 2013. *The Austronesian languages*. Canberra: Pacific Linguistics. [Revised online version of hardcopy originally published 2009].

Bodt, Timotheus A. 2014. Ethnolinguistic survey of westernmost Arunachal Pradesh: A fieldworker's impressions. *Linguistics of the Tibeto-Burman Area*, 37(2): 198-239.

———. 2019. The Duhumbi perspective on Proto-Western Kho-Bwa rhymes. *Die Sprache*, 52(2): 141-176.

———. 2020. *Grammar of Duhumbi (Chugpa)*. Leiden/Boston: Brill.

———. 2021. Sartang (West Kameng district, Arunachal Pradesh, India) – Language contexts. *Language Documentation and Description*, 20: 162-188. doi: https://doi.org/10.25894/ldd49

———. 2022. Negation in Kho-Bwa: A typological comparison. In: Hayashi, Norihiko, and Ikeda, Takumi (eds.), *Diversity of Negation. Grammatical Phenomena of Sino-Tibetan Languages*, 5: 203-237. Kyoto, Japan: Institute for Research in Humanities, Kyoto University. http://hdl.han-

dle.net/2433/275714

Bodt, Timotheus A. and Johann-Mattis List. 2019. Testing the predictive strength of the comparative method: an ongoing experiment on unattested words in Western Kho-Bwa languages. *Papers in Historical Phonology*, 4: 22-44.

Bodt, Timotheus A. and Ismael Lieberherr. 2015. First notes on the phonology and classification of the Bangru language of India. *Linguistics of the Tibeto-Burman Area*, 38(1): 66-123. doi: 10.1075/ltba.38.1.03bod

Boro, A. 1978. *Miju dictionary*. Shillong: Director of Research, Arunachal Pradesh Government.

Boro, Hirak. 2024. A preliminary study of Sherdukpen phonology. *Himalayan Linguistics*, 23-2: 35-52. doi 10.5070/H923258183

Bradley, David 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), *Tibeto-Burman Languages of the Himalayas*: 1-72. Pacific Linguistics A-86. Canberra: Australian National University.

Brown, Nathan. 1837. Comparison of Indo-Chinese Languages. *Journal of the Asiatic Society of Bengal*, 6: 1023-1038.

Burling, Robbins. 2003. The Tibeto-Burman languages of Northeastern India. In Thurgood, Graham, and Randy J. LaPolla (eds.). *The Sino-Tibetan languages*: 169-191. London and New York: Routledge.

Button, Christopher T. J. 2011. *A Reconstruction of Proto Northern Chin in Old Burmese and Old Chinese Perspective*. 2 vols. STEDT Monograph No. 10. Berkeley.

Campbell, George. 1874. *Specimens of Languages of India: Including Those of the Aboriginal Tribes of Bengal, the Central Provinces, and the Eastern Frontier*. Calcutta: Printed at the Bengal Secretariat Press.

Campbell, Lyle. 1991. On so-called Pan-Americanisms. *International Journal of American Linguistics*, 57(3): 394-399.

Chakravarty, L.N. 1978. *A dictionary of the Taraon language*. Shillong: NEFA.

Dai, Mamang, et al. *Hambreelmai Sai*. Chennai: Tulika. 2013.

Das Gupta, Kamalesh. 1977. *A Phrase Book in Miju*. Shillong: Research Secretariat.

Deuri, Ram Kumar. 1982. *The Sulungs*. Shillong: Government of Arunachal Pradesh.

Dey, Lucky. 2015. Kinship terms in Hruso. *North East Indian Linguistics*, 7: 147-158.

Dondrup, Rinchin. 1988. *A Handbook on Sherdukpen Language*. Itanagar: Director of Research, Arunachal Pradesh Government.

———. 1990. *A Handbook on Bugun Language*. Itanagar: Director of Research, Arunachal Pradesh Government.

———. 2010. *An Introduction to Boot Monpa language*. Itanagar: Director of Research, Arunachal

Pradesh Government. [NB. The title page gives the date of publication as 2004, but the preface is dated 2009 and it was actually published in 2010.]

D'Souza, Vijay A. 2015. Towards a Phonology of Hrusso Aka. MA thesis: University of Oxford.

———. 2018. High vowel devoicing in Hrusso Aka. In Linda Konnerth, Stephen Morey and Amos Teo (eds.), *North East Indian Linguistics*, 8: 33-46. Canberra: Asia-Pacific Linguistics.

———. 2021. *Aspects of Hrusso Aka Phonology and Morphology*. PhD diss.: University of Oxford.

Dutta, Dipak Kumar. 1999. *The Monpas of Kalaktang: Arunachal Pradesh*. Itanagar: Director of Research, Arunachal Pradesh Government.

———. 2012. *The Miju-Mishmis of Arunachal Pradesh*. Itanagar: Arunachal Pradesh Research Directorate. [NB: Despite the publication date, this did not become available until 2016.]

Ethnologue 2024. *The Ethnologue: Languages of the world*. Dallas: Summer Institute of Linguistics.

Evans, Jonathan P., and Johakso Manyu. 2021. The sounds of Tawrã (Digaru-Mishmi), a Tibeto-Burman language. *Linguistics of the Tibeto-Burman Area*, 44(1): 1-26.

Grewal, Dalvindar Singh. 1997. *The Tribes of Arunachal Pradesh*. 2 vols. Delhi: South Asia Publications.

Grierson, George A. (ed.). 1909. *Linguistic Survey of India. Volume III, 1, Tibeto-Burman family*. Calcutta: Superintendent of Government Printing.

Heath, Jeffrey. 1978. *Linguistic Diffusion in Arnhem land*. Australian Institute of Aboriginal Studies.

———. 1981. A case of intensive lexical diffusion: Arnhem Land, Australia. *Language*, 57(2): 335-367.

Hesselmeyer, C.H. 1868. The hill-tribes of the Northern Frontier of Assam. *JASB*, XXXVII: 192-208.

Jacquesson, François. 2001. Person-marking in TB languages, Northeastern India. *Linguistics of the Tibeto-Burman Area*, 24(1):113-144.

———. 2015. *An Introduction to Rupa Language*. Bochum: Universitätsverlag Dr. N. Brockmeyer.

Jacquesson, François and Pascale Dollfus. 2013. *Khiksaba, a Festival in Rupa Country*. Guwahati: Spectrum Publications.

Jiang, Huo 江荻. 2023. Linguistic diversity and classification in Tibet (西藏的语言多样性及其分类). Chinese Tibetology Centre (中国藏学研究中心).

Jiang, Huo, Daqin Li & Hongkai Sun. 2013. *Darangyu yanjiu* [A study of Taraon]. Beijing: Ethnic Publishing House.

Koch, Harold, and Nordlinger, Rachel. 2014. *The Languages and Linguistics of Australia*. Berlin: De Gruyter Mouton.

Konow, Sten. 1902. Note on the languages spoken between the Assam Valley and Tibet. *The Journal of the Royal Asiatic Society of Great Britain and Ireland*, 1902: 127-137.

———. 1909. The Tibeto-Burman family. In: George A. Grierson (ed.), *Linguistic survey of India.*

Volume III, 1, Tibeto-Burman Family. Calcutta: Office of the Superintendent of Government Printing.

Krisikro, Sutik. 2006. *Tawra: tako zo bra/ käman, pït sai brat*. Tezu: Graphic Systems.

Lander-Portnoy, Maury. 2013. Let Buguns be Buguns: A Preliminary Phonetics, Phonology, and Morphology of the Bugun Language. Swarthmore College: MA thesis.

Landi, Victor. 2005. *The Meyors and Their Language*. Itanagar: Director of Research, Arunachal Pradesh Government.

Lǐ, Dàqín, & Jiang Di 李大勤，江荻 2001. 扎语概况 [Brief introduction of Zha speech]. 民族语文 *Minority Languages of China* 2001(6): 61-75 [in Chinese].

Lǐ, Dàqín. 2003. Bēngrú-yǔ gaìkuàng [Bangru language overview]. *Minzu Yuwen*, (5):64-80.

———. 2004. *Sulong yu yan jiu*. Beijing: National Minorities Publisher.

Lieberherr, Ismael. 2015. A progress report on the historical phonology and affiliation of Puroik. *North East Indian Linguistics (NEIL)*, 7. Canberra, Australian National University: Asia-Pacific Linguistics Open Access.

———. 2017. *A Grammar of Bulu Puroik*. Universität Bern dissertation. [not circulated]

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley, Los Angeles, London: University of California Press.

———. 2009. Stable Roots in Sino-Tibetan/Tibeto-Burman. In Y. Nagano (ed.), *Senri Ethnological Studies*, 75: 291-318.

Modi, Milorai. 2008. *The Millangs*. Itanagar/New Delhi: Himalayan Publishers.

Morey, Stephen. 2018. Verb stem alternation in Pangwa Tangsa. *North East Indian Linguistics*, 8: 86-109.

Mortensen, David R. 2003. *Comparative Tangkhul*. University of California, Berkeley, ms.

———. 2023. Kuki-Chin phonology: An overview. *Himalayan Linguistics*, 22(1): 5-35.

Needham, Jack F. 1886. *A few Dîgârô (Târoan), (Mîjû) M'jû, and Thibetian Words*. Shillong.

Ouyang, Jueya. 1985. *A Brief Introduction to the Luoba Language* [in Chinese], Beijing: Mínzú Chūbǎnshè.

Pandey, B. B. 1996. *The Buguns: A Tribe in Transition*. Delhi: Himalayan Publishers.

Pelkey, Jamin R. 2011. *Dialectology as Dialectic: Interpreting Phula Variation*. Berlin: Mouton de Gruyter.

Post, Mark W. 2011. *The Galo Language*. ms.

———. 2017. *The Tangam Language*. Leiden: Brill.

———. 2022. Upland languages. In Jelle J. P. Wouters & Tanka B. Subba (eds.), *The Routledge Companion to Northeast India*: 469-474. London: Routledge.

Post, Mark W. & Robbins Burling. 2017. The Tibeto-Burman languages of Northeastern India. In: Graham Thurgood and Randy J. LaPolla (eds.), *The Sino-Tibetan Languages* (2nd ed.): 213-242. London & New York: Routledge.

Post, Mark W. & Roger M. Blench 2011. Siangic: A new language phylum in northeast India. Handout circulated at the 6th NEILS Conference, Tezpur.

Post, Mark W. & Yankee Modi. 2011. Language contact and the genetic position of Milang (Eastern Himalaya). *Anthropological Linguistics*, 53(3): 215-258.

Pulu, Jatan. 1991. *A Phrase Book on Taraon Language*. Itanagar: Directorate of Research.

Pulu, Shri Jimi. 2002. *A Handbook on Idu Mishmi Language*. Itanagar: Arunachal Pradesh State Government Directorate of Research.

Ramya, T. 2011. *An Ethnographic Study of Bangrus of Kurung Kumey District of Arunachal Pradesh*. MA. Tribal Studies. Doimukh: Rajiv Gandhi University.

———. 2012. Bangrus of Arunachal Pradesh: An Ethnographic Profile. *International Journal of Social Science Tomorrow*, 1(3):1-12.

Reesink, Ger P. & Dunn, Michael. 2018. Contact phenomena in Austronesian and Papuan languages. In B. Palmer (ed.). *The Languages and Linguistics of the New Guinea Area: A comprehensive guide*: 939-985. Berlin/Boston: De Gruyter Mouton.

Remsangphuia 2008. *Puroik Phonology*. Shillong: Don Bosco Technical School.

Robinson, William. 1856. Notes on the languages spoken by the Mi-Shmis. *Journal of the Asiatic Society of Bengal*, 24: 307-324.

Schubert, Johannes. 1964. Hrusso-Vokabular. *Mitteilungen des Instituts für Orientforschung*, X: 295-350.

Shafer, Robert. 1947. Hruso. *Bulletin of the School of Oriental and African Studies*, 12(1): 184-196.

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

———. 1966/1967. *Introduction to Sino-Tibetan* (Part I: 1966, Part II: 1967). Wiesbaden: Otto Harrassowitz.

Simon, Ivan M. n.d. *Miji Language Guide*. Shillong: Government of Arunachal Pradesh.

———. 1993. *Aka Language Guide*. Itanagar: Government of Arunachal Pradesh.

Sinha, Nupur, & Barbora, Madhumita. 2021. Language endangerment amongst Hruso-Aka and Koro. *SEL Journal* 6 (II).

Soja, Rai. 2009. *English-Puroik Dictionary*. Shillong: Living Word Communicators.

Stonor, Charles R. 1952. The Sulung tribe of the Assam Himalayas. *Anthropos*, 47: 947-962.

Sun, Hongkai. 1983a. A Brief Introduction to Idu (Luoba) Language. [In Chinese], *Mínzú Yǔwén* 6. Beijing: Mínzú Chūbǎnshè.

———. 1983b. The Languages of the Peoples of the Six River Valley Region and Their Genetic Classification. [In Chinese]. *Mínzú Xuébào*. Kunming: Yunnan People's Publishing Company.

———. 1991. [Title in Chinese]. Beijing: Social Sciences Press.

———. 1999. On the Tibeto-Burman languages of the Eastern Himalayan area in China. *Linguistics of the Tibeto-Burman Area*, 22(2): 61-72.

Sun, Hongkai, Lu Shaozun, Zhang Jichuan & Ouyang Jueya (eds.). 1980. *Menba, Luoba, Dengren de yuyan* (The languages of the Menba, Luoba and Deng peoples). Beijing: Social Sciences Press.

Sun, Tianshin Jackson. 1993. *A Historical-Comparative Study of the Tani Branch of Tibeto-Burman*. PhD diss. Department of Linguistics. Berkeley: University of California.

Tawsik, Somodang. 2014. *Kaman-Tawrah*. Tinsukia: Chitra Printers.

Tayeng, Aduk. 1976. *Milang Phrase Book*. Shillong: Government of Arunachal Pradesh.

Thurgood, Graham, & Randy J. LaPolla (eds.). 2003. *The Sino-Tibetan Languages. Routledge Language Family Series*. London & New York: Routledge.

VanBik, Kenneth. 2007. *Proto-Kuki-Chin*. PhD diss. University of California, Berkeley.

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 vols. Leiden: Brill.

———. 2007. Dzala and Dakpa form a coherent subgroup within East Bodish, and some related thoughts. In Roland Bielmeier & Felix Haller (eds.), *Linguistics of the Himalayas and Beyond*: 71-84. Berlin/New York: Mouton de Gruyter.

———. 2008. To which language family does Chinese belong, or what's in a name. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros & Marie Lin (eds.), *Past Human Migrations in East Asia: Matching Archaeology, Linguistics and Genetics*: 219-253. London: Routledge.

Weedall, P. 2021. *The Western Miji Dictionary*. ms.

Wood, Daniel S. 2008. *An Phonological, Tonological, and Morphological Study of Wa (Paraok)*. PhD diss. University of California, Berkeley.

Comparative Wordlist for Arunachal Pradesh Languages

The table below presents lexical data across 16 Arunachal Pradesh language varieties: Koro, Mö, Bugun, Puroik, Hruso, Miji, Milang, Tawrã, Idu, Kman, and several Tibetic varieties (Dirang, Tawang, Memba, Meyor, Nah). Color coding indicates proposed cognate sets. CTB = Comparative Tibeto-Burman reconstruction.

Gloss	CTB	Koro	Mö	Bugun	Puroik	Hruso	Miji	Milang	Tawrã	Idu	Kman	Dirang	Tawang	Memba	Meyor	Nah	PT
One	*g-t(y)ik	e-ce	han, ãi	dʒiɔ	hwi	ă	uŋ	a-kan	kʰín	khàgõ	kume	tʰur	tʰi	gik	guk	akin	*kon² ~ təl
Two	*g-ni-s	ki-ne	ɲik	neŋ	ɲi	ksi	gni	nə	kayiŋ	kà.nyi	kinin	nitsiŋ	neˡ	ɲi	ni	aɲi	ɲi
Three	*g-sum	kala	uŋ	im	heik	ðʃ	gtʰən	ham	kasaŋ	kã.sõ	ksam	sam	sum	sum	som	aum	*hium¹
Four	*b-ləy	ko-	bisi	vi	rei, wai	pî	bli	pə	kaprayk	kà.pri	kambran	bɟi	bli	gi	dʒee	aɲo	*ŋo¹
Five	*b-ŋa	plẽ	kʰu	kua	u	pom	buɲu	pa-ŋu	maŋa	màɲá	klin	ŋa	leŋa	ŋe	ŋa	aɲo	*ŋo¹
Six	*d-ruk	su-fi	kʰit	rab	reik	ʒẽ	rɛ	sa-ap	tahro	tãhrõ	katam	kʰuŋ	gro	du	trok	akce	*kra(ŋ)¹
Seven	*s-nis	rõ	sit	milye	lye	mrə	mya	ra-ŋal	wẽ	íŭ	nin	zum	ŋis	dun	dun	kani	*ka-na¹
Eight	*b-r-	rà-lá	sargyat	mla	la	s.kzə	sige	ra-jəŋ	lĩm	ìʎŭ	grin	yen	get	gey	zat	pini	*pri¹-ɲi-
Nine	*d-gəw	gèyé	dikʰi	dige	donge	s.tʰə	stʰən	ka-pəm	kiŋaŋ	khrĩnĩ	natmo	gu	dûgu	gu	gu	kyowa	*kV-naŋ
Ten	*gip	fã-la	sô	suŋwa	guat	ʁə	lin	haŋ-tak	hâlaŋ	hũũ	kyapmo	se	ɟiʰ	ɟu thum	ɟu	ariŋ	*cam¹ ~ jiŋ¹
Head	*d-bu-s	dʒù-	khuruk	kʰruk	*a-koŋ	ekʰyẽ	wuu	dum-	kru pom	ikrũyã	ku	ɹaraŋ	got	go	aku		*kon² ~ təl
Nose	*na,	ke-pe	nupʰuŋ	epʰuŋ	pok	usˀɔ̃	nyubyuŋ	ɲokuŋ	haŋagam	ẽnã(mbõ)	mnuŋ	na uŋ	naʰ	noguŋ	na?	napʰiŋ	*ɲi²
Eye	*mik	ɲi-	khibi	meyak	kak	əni	mre?	a-mik	blom	ẽlõ	mik	rniŋ	meloŋ	mi	mik	nik	*mik
Mouth	*mka	sapu	niɟaw	ɟyam	səek	unzũ	mugo		threndom	ẽkóbà	ɟu	nowaŋ	kʰa	namdʒo	gam	*nap, *gam	
Ear	*r-na	rã	kʰutʰuŋ	ekʰõ	*a-kuñ	ufũ	m.ʒo?	ra-ŋu	kruna	ãkónã	iŋ	ney	nelãp	m			
Tongue	*s-l(y)a	sãlé²	laphõ	rhi	ruyi	əzlbra	dʒaksi	si-dal	theleŋna	ilĩná	b.lay	le	leʰ	gʃoli			
Tooth	*swa	fĩ	nutʰuŋ	siŋen	kotuwaŋ	utu	tʰu	sip-pa	lyã	tãmbrõ	sey	ʃa	waʰ	sow	ɟu	hikjuŋ	*fii
Arm	*g-lak	là	ik	wat	gəit	opəə	pʰaŋ, (m)gĩ	a-lak	apri	ãkhõ, ãtò	rawk	garaŋ	laʰ	lak	arak	lak	*lak
Leg	*kaŋ	ni-bi	la	loe	lae	əɟĩ	lay? E.,	a-byaa	grõ	ãnggãsà	pla	bi	lemi	kʰaŋ	tepro	lapa	*lo
Stomach	*grwat	gay	siriŋ	lui	loye buk	əvəkʰũ	luŋ		yaŋ kawẽ	khàpó	dak	pʰoloŋ	kepa	dogo	pʰuko	kipo	*kri
Bone	*rus	ɲirá	skit	ezeŋ	adʒay	əkəbe	məlyaŋ		reb buŋ	rũmbõ	rak	kʰaŋ	roɹba	rugo	gʃereek	alo	*loŋ
Blood	*s-	evi	ha	afoe	huɹ	cə	ʒay		harweyg	iyũ	iwii	ʒi	kʰra	tha	awi	oyik	*vii
Sun	*nəy	me-	nini	hanayaŋ	kiri	drũ	dʒo?	məə-	riŋ	ĩnyĩ	amik	ŋam	plaŋ	ɲum	mik	dom	*ɲi
Moon	*s-la	a-la	namblu	habia	ambu	hubye	łuu, tu	poo-luʰ	hallo	ẽlã	lay	laɲi	lei	dager	lo dowa	polu	*polo
Star	*s-kar	dogre	ɟuzuk	satyoŋ	pədʒeik	litsi	dutsuŋ	ta-kar	kadiŋ	ãndĩkrũ	ŋalci	karma	karma	karem	karma	taker	*kar
Man	*r-min	mur	dʒuhu	b.phua	apʰu	nəna	ɲi, nuvu	ma-lu	me	mẽyá	coŋ	soŋa	miʰ	soŋa	miŋ		?
Woman	*mow	msn	dʒimi	bimi	amwi	mĩm	nəmre?	ma-mi	miyã	yãkũ	kamay	ŋiza	ãmah	nedʒa	mainag	ɲima	?
Child	*za/*tsa	ŋwà	nunu	ani	adʒuaŋ	sa	amay		a	ã	?	za		tukto		nene	hemi
Old man	*bəw	dʒiman	frioŋ	amayin	mukʰrõ	vu kʰraŋ		mowaa	mĩɟiprã	kanaŋ	kui	kʰu		ata	im, seŋ	gidʒoŋ	ɲilo
Dog	*kʷəy	ềklề	pʰitʰa	gʰey	kayu	ɹ.ʎu	ɹadʒi?	kuak	ikũ	kui	kʰu	kʰi	kʰi	kwi	aki		*kii
Pig	*pwak	lele	swag	wak	mədow	vo	ʒo?	ayek	beleyg	ili	lii	pʰakpa	pʰa	pʰa	lìk	arik	*ryk
Tiger	*k-la	ɹaru	phõ	muɹua	neray	ɹ.dʒi	tʰuŋgraŋ	paa-tiʰ	tamya	ãmrã	topəw	goŋtak	g̃ẽn	ta?	ziktetha	abiŋ	*mro, *paŋ-
Water	*ti(y)	si	kʰo	kʰo	kua	kʰu	vu?	a-si	macey	màɟĩ	tii	ri	ɟĩ	gʰu	ati	iɟi	*si
Fire	*mey	mi-la	ba	boe	bawe	mi	may?	a-mi	naamiŋ	ãmrũ	mai	mi	meʰ	me?	mi	əmə	*mə
Stone	*r-luŋ	vuvu	liŋ	l.baw	*ka-luŋ	kun	guluŋ?		phlã	ãlãphrã	laʷŋ	luŋ	gor				*luŋ
Tree	*siŋ,	lã	hiŋ tʰuŋ	hiŋmua	ɹə(mua)	ɹõ	wuu?	haŋ-sa	masaŋ	àsimbõ	saŋ	ɟiŋ	ɹyaŋ	ɟiŋ			?
Leaf	*r-pak	nino	alap	arap	məɹay	ɹere	le?		na	ná	lap	ɹawa	palãp	lemah	əlaŋ	dunpu	seɲna
Name	*miŋ	niraŋ	azeŋ	ebeŋ	*a-bycn	aɲiɲi	m.rin		amaŋ	ãmũ	amaŋ	meŋ			alap	nane	*nə
Eat	*dzya	to	guwa,	gʰa	ɟina	tsa	tsu?	tu	tha	hà	ɹa	za	sasuʰ	sale	gʰoem	da	*do
Culture																	
Mithun	none	sù	smu	syá	ła	fu	ɹu?	a-sù	aɹya	sã	cal	meɲa		bamin	piiyee	se	
Iron	*syal		sẽ	yuŋ		si	sã	arəm	say	ɟĩ	taŋgli	perr	lʰe	ga?	gak	tagi	*ryok
Dao	*sta	kasa	handu	mudua	ɹe	vetsi	vay tsən	ayok	tara	ẽẽcẽ	sut	ɹowaŋ	kʰyop	paɹɹa	kunak	oriyuk	(a)-ryok
Banana	none	gerdʒi	musuŋ	tsyum	kapak	ruloŋ	r.laŋ		paydʒ	àdʒi brũ	hambyooŋ	leysi	lam rep		saŋjuŋ	kupak	*kopak
Arum, taro	*grwa	làm	dʒu ɹhak	dʒawk	ɹuwa	tʰrə	tca?	aaŋ	sam	sónã	gal	bozoŋ	blu	solum		aɲi	
Millet	none		gicam	ɹo	tamayi	kʰsə	dʒ.roo		haabra	yàmbã	muuŋ	koŋpu	kowp	temi	turo	tami	
Paddy	*ma(y)	kĩ	nise	nisi	amaŋ	olgi	ã?		ke	kò	ha, maŋ	ra		deyso	sipu	am	
Rice	*ma(y)	ki	nudob	nyiŋ	ambiŋ	ŏ	ã?		ke	kò	haku	kʰu	dep	dey	andek	am	*pim
Cooked rice			mãm	ha?		amaŋ	zara	tsavo mɹi			kòri	syat	togʰan		mam		

Legend: Yellow = Mö cluster cognates; Orange = Bugun-Puroik connections; Blue = Tawrã-Idu-Kman; Cyan = Milang connections; Pink = Tibetic loanwords