# The 'Mishmi' languages, Idu, Tawra and Kman: a mismatch between cultural and linguistic relations



[DRAFT CIRCULATED FOR ICEHEP 2017]

Melbourne 8-10<sup>th</sup> February, 2017

Roger Blench
McDonald Institute for Archaeological Research
University of Cambridge
Department of History, University of Jos
Correspondence to:
8, Guest Road
Cambridge CB1 2AL
United Kingdom
Voice/ Ans (00-44)-(0)1223-560687
Mobile worldwide (00-44)-(0)7847-495590
E-mail rogerblench@yahoo.co.uk
http://www.rogerblench.info/RBOP.htm

This printout: Roing, March 19, 2017

# TABLE OF CONTENTS

1. Introduction: the term 'Mishmi' and the Idu-Tawra	1
2. Existing resources on Idu, Tawra and Kman	2
2.1 The term Mishmi	2 2 2 3
2.2 Idu [clk]	2
2.3 Tawra [mhu]	2
2.4 Kman [mxj]	3
2.5 Tibeto-Burman affiliation	3
3. Cultural relations between the Mishmi peoples	3
4. Common linguistic features of Mishmi languages	4
5. Comparison of morphosyntactic markers	5
5.1 Number marking	5
5.2 Gender marking	6
5.3 Numeral classifiers	7
6. Lexical datasets	9
6.1 Comparisons across Mishmi	9
6.2 Idu-Tawra comparisons which argue for a close relationship	9
6.3 Idu-Tawra comparisons showing marked lexical divergence	10
7. Tawra and Kman cultural convergence	10
8. Exploring the cultural mismatch	12
9. West of Tani – a parallel case of convergence?	13
10. Conclusions	13
Bibliography	14
8 1 1	
TABLES	
Table 1. Common sultimal features of the three Michael macules	2
Table 1. Common cultural features of the three Mishmi peoples	
Table 2. Common typological features of Mishmi languages	
Table 4. Idu number marking on nouns	
Table 5. Tawra number marking on nouns	
Table 6. Kman number marking	6
Table 7. Gender marking on Idu animal names	
Table 8. Tawra gender marking for animals	
Table 9. Kman gender marking for animals	
Table 10. Idu numeral classifiers	
Table 11. Tawra numeral classifiers	
Table 12. Kman numeral classifiers and their categories	8
Table 13. Lexical resemblances across Mishmi	9
Table 14. Mishmi lower numerals	
Table 15. Mishmi colour terms	
Table 16. Idu and Tawra body parts	10
риотое	
PHOTOS	
Photo 1. MFPS service house, Tezu	
Photo 2. Death house, De'ing, 1916	
Photo 3. Kman death house, Yatong, 2015	
Photo 4. Hruso/Koro music video poster	14

### **ABSTRACT**

The Mishmi people consist of three separate ethnic groups speaking distinct languages, the Idu, Tawra [=Digaru] and the Kman [=Miju] living in the far northeast of Arunachal Pradesh, NE India. A few villages of each group are also found over the border in Tibet and the Tawra also have villages in Myanmar. A fourth group, the Meyor [also Zha], living in Walong and Kibithoo, are sometimes included among the Mishmi, and certainly their language shows connections with Kman. All four languages are highly endangered, with a few thousand speakers at best and Meyor a few hundred. The main predator language is Hindi, which is increasingly used in schools and media.

The three Mishmi groups share many common cultural features, such as the use of longhouses, parallel languages for shamans and hunters, a system of spiritual entities which reflect ecological zones, and much else. They all range over huge expanses of territory, stretching from the Tibetan Plateau to Assam Plains, and their cultures reflect this variability. Hunting and gathering of wild plants continues to play a major role in daily life, although the rice culture of the plains is gradually transforming this.

However, their linguistic relationships are perplexing at best. Idu and Tawra are usually linked together in a subgroup called Digarish in old sources, while Kman is distinct. On a basic wordlist, Kman shares little or nothing with neighbouring languages, except Meyor. The strong similarities between Meyor and Kman in the lexicon argue for borrowing, rather than a genetic connection. They are usually treated as Tibeto-Burman, despite a lack of evidence to support this affiliation. The relationship between the Tawra and Idu languages is complex. In some semantic areas, such as lower numerals or colours, their lexemes are almost entirely cognate. However, elsewhere in the lexicon, such as body parts, cognacy levels are extremely low. A small set of basic verbs are cognate but otherwise they are very divergent. The same is true of morphosyntax; some elements are similar, others completely different. This suggests that even the putative relationship between Tawra and Idu may be only borrowing.

At the cultural level, relationships are quite different. Despite their linguistic differences, Tawra and Kman are held to be two aspects of the same culture and unrelated to Idu. Since the 2000s, there has been a slew of local publications which include comparative wordlists, dialogues and culture guides in the two languages. Among these are reading and writing books which attempt to force them into a common orthography, despite the obvious differences in their phonology. Kman shamans chant largely in Tawra, and that innovative church-like organisations such as the Mishmi Faith Promotion Society (MFPS) hold services in a mixed Kman/Tawra shamanic register.

The Mishmi languages present a striking mismatch between local perceptions of cultural relationships and the linguistic facts, as far as they are known. Historically, all three languages are most likely of distinct origin, but living in a common environment, in close physical contact and in constant interchange set in motion powerful cultural levelling. There must have been a period when all three groups shared a common culture, in order to account for the basic similarities still observable. At some point, Tawra and Idu became intertwined, hence the intensive borrowing in restricted areas of the lexicon. Much more recently, the Tawra/Idu bond was broken, and the construction of a relationship with the Kman began. The puzzle of why there has been so little lexical interchange between the three groups and why, when it occurs, it is so selective, remains to be solved. It is suggested that in global terms this situation is highly unusual and challenges many common generalisations in sociolinguistics. There is an almost exact parallel west of the Tani languages with the Miji, Hruso and Koro peoples. Despite a striking absence of common lexicon and morphosyntax (as far as this is known) their cultures map strongly against one another. It is suggested that in global terms this situation is highly unusual and challenges many common generalisations in sociolinguistics.

Keywords: Idu; Tawra; Kman; historical linguistics; anthropology; Arunachal Pradesh

### 1. Introduction: the term 'Mishmi' and the Idu-Tawra

The Mishmi people consist of three separate ethnic groups speaking distinct languages, the Idu, Tawra [=Digaru] and the Kman [=Miju] living in the far northeast of Arunachal Pradesh, NE India (Map 1). A few villages of each group are also found over the border in Tibet and the Tawra also have villages in Myanmar. A fourth group, the Meyor [also Zha], living in Walong and Kibithoo, are sometimes included among the Mishmi, and certainly their language shows connections with Kman. All four languages are highly endangered, with a few thousand speakers at best. The main predator language is Hindi, which is increasingly used in schools and media.

The three Mishmi groups share many common cultural features, such as the use of longhouses, parallel languages for shamans and hunters, a system of spiritual entities which reflect ecological zones, and much else. They all range over huge expanses of territory, stretching from the Tibetan Plateau to the Assam Plains, and their cultures reflect this variation. Hunting and gathering of wild plants continues to play a major role in daily life, although the rice culture of the plains is gradually transforming this.

However, their linguistic relationships are perplexing at best. Indeed accounts of these languages have been blighted by amateur historical linguistics and careless use of existing data. Early sources, such as those analysed by Shafer (1955) were extremely fragmentary. All three languages are classified as the 'North Assam' group of Tibeto-Burman, although the evidence for this appears to be restricted to about ten common words, 'eye', 'sun', lower numerals etc. However, additional confusion

Map 1. The general location of the Mishmi peoples



was introduced by the publications of Sastry (1984, 1991) who conflated data from all three languages in a single monograph. Even by the standards of CIIL, this was a nadir of scholarship. Idu and Tawra are usually linked together in a subgroup erroneously called Digarish in the literature, while Kman is distinct. On a basic wordlist, Kman shares little or nothing with neighbouring languages, with the exception of Meyor. The strong lexical similarities between Meyor and Kman argue for borrowing rather than genetic affiliation.

The relationship between the Tawra and Idu languages is complex. In some semantic areas, such as lower numerals or colours, their lexemes are almost entirely cognate. However, elsewhere in the lexicon, such as body parts, cognacy levels are extremely low. A small set of basic verbs are cognate but otherwise they are very divergent. The same is true of morphosyntax; some elements are similar, others completely different. This suggests that even the putative relationship between Tawra and Idu may be only borrowing.

At the cultural level, relationships are quite different. Despite their overt linguistic differences, Tawra and Kman are held to be two aspects of the same culture and unrelated to Idu. Since 2000, a slew of local publications have appeared which include comparative wordlists, dialogues and culture guides in the two languages. Among these are reading and writing books which attempt to force them into a common orthography, despite the obvious differences in their phonology. Kman shamans chant largely in Tawra, and that innovative church-like organisations such as the Mishmi Faith Promotion Society (MFPS) hold services in a mixed Kman/Tawra shamanic register.

This paper<sup>1</sup> is a preliminary account of the linguistic differentiation and cultural convergence of the three Mishmi peoples. It describes the existing linguistic and anthropological resources for each group and compares cultural features, morphosyntactic markers and lexical cognates. It concludes with a tentative model of the history of cultural interchange and suggests a parallel with the Hruso, Koro and Miji, further went in Arunachal Pradesh. The data is derived from field trips to Arunachal Pradesh in 2105, 2016 and 2017.

# 2. Existing resources on Idu, Tawra and Kman

### 2.1 The term Mishmi

The term 'Mishmi' is used in the travel literature as far back as the early nineteenth century (Brown 1837; Rowlatt 1845) to refer to three distinct peoples, the Idu, Tawra [=Taraon and the Kman [=Geman, Miju].The numerous variant names and spellings are detailed below. All of these peoples live in the northeast of Arunachal Pradesh, bordering Tibet and Myanmar, and all have villages across the frontier (Map 1). A relationship between Idu and Tawra is recognised in Konow (1909). Robinson (1855) compared Tawra [=Taying] with Kman [=Mijhu] but did not include Idu for lack of firsthand data. All current reference sources such as Ethnologue and Glottolog treat Idu and Tawra as subgroups of Tibeto-Burman, a group sometimes known as 'Digarish' which originates with Shafer (1955). Although classified as the 'North Assam' subgroup of Tibeto-Burman by Konow (1902) there has been remarkably little published in the way of evidence for this affiliation. Indeed, Blench & Post (2013) question whether they should not be treated as isolates, in the absence of any positive argument.

# 2.2 Idu [clk]

The Idu 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). Some material can be found in Campbell (1874) and Konow (1909). The only significant publications on Idu from the Indian side are the pre-linguistic Talukdar (1962), Pulu (1978, 2002a,b). The main value of Pulu (2002) is as an elicitation guide, although the centralised Hindu-mainstream thinking that dominates its semantics means it needs to be used with caution. Idu has also been described from the Chinese side [under the name Lhoba], notably in Sun et al. (1980, 1991), Sun (1983a,b, 1999) and Ouyang (1985).

The earliest discussion of Idu social life is the brief section in Dalton (1872) which covers their social organisation, religion and warlike propensities. There are two short monographs on the social and material life of the Idu, Baruah (1960) and Bhattarcharjee (1983) both of which reflect long residence in the Dibang area. Baruah is of particular interest, since, although it was researched in the period after the earthquake of 1950, it reflects a period when the Idu still had very little interaction with the outside world.

### 2.3 Tawra [mhu]

The Tawra are also known as Darang 达让僜, Daruang, Deng, Digaro, Digaru, Mishmi, Taaon, Taraon and Taying. The name 'Digaru' (the name of a major river) often used in English conversation. Records of Tawra go back at least to Robinson (1856). Needham (1886) gives a comparative wordlist of Tawra, Kman and Tibetan. Recensions of existing data are given in Campbell (1874) and Konow (1902, 1909). Luce (1944) is an unpublished lexical list cited in bibliographies, but not in circulation in scanned form. Modern publications with a 'practical' orientation include Chakravarty (1978) and Pulu (1991). Kumar (1988) is a Tawra Hindi dictionary which probably borrows heavily from Chakravarty (1978). Chinese scholars have also worked on Tawra, which they call 'Daruang'. The Tawra language has been briefly described in Sun (1983, 1991, 1999) the last of which is in English. A summary of their findings is given in Sun (1999) which

<sup>&</sup>lt;sup>1</sup> I would like to thank the many people in all three communities who have given up their time to help me in the collection and preparation of the data that has gone into this paper, in particular Dr. Mite Lingi and the ILDC members in Roing, Sokhep Kri Barrister Manu in Tezu. Background on the phonology and orthography of all three languages can be found in a series of online resources posted on academia.edu. Note that transcriptions for Tawra are more tentative than for Idu and Kman, since phonological analysis is at a preliminary stage.

includes a brief description of the hunters' speech register. Jiang et al. (2013) is an extended grammar of Tawra in Chinese.

# 2.4 Kman [mxj]

Alternative names for the Kman include Eastern Mishmi, Geman Deng, Kaman, Miju. Ethnologue (2017) includes Miji as a variant, but this is to confuse it with the Miji people, a distinct group living in western Arunachal Pradesh. The first record of Kman appears to be Robinson (1856) which is quite accurate for the period, and his transcriptions are recognisable today. The only modern publications on the Indian side are Das Gupta (1977) and Boro (1978). These are said to be 'practical' guides and 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, Li (2003) 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, a lexical guide has been published covering both Kman and the neighbouring Tawra [Taraon] language (Tawsik 2014) although the orthography bears no resemblance to any other publication. The Kman people have been the subject of a brief anthropological monograph (Dutta 2012).

### 2.5 Tibeto-Burman affiliation

Despite being treated as Tibeto-Burman in standard sources, evidence or argument for this hypothesis is sparse in the extreme. All three languages are classified as the 'North Assam' group of Tibeto-Burman, although the evidence for this appears to be restricted to about ten common words, 'eye', 'sun', lower numerals etc. All three languages show evidence for pronominal cross-referencing, although segmental cognates with one another with existing described systems are difficult to discern. For practical purposes, these languages are best treated as isolates.

## 3. Cultural relations between the Mishmi peoples

The perception that the three Mishmi groups share aspects of a common culture is based on some genuine similarities which mark them off from their neighbours, the Tani to the west, the Tibetans to the north and the Naga and Khamti to the south and east. Table 1 synthesises some of these shared cultural features.

Table 1. Common cultural features of the three Mishmi peoples

Feature	Commentary		
Longhouses	All groups build extensive longhouses, traditionally housing multiple wives each with individual hearths. The internal architecture of these houses is very		
	similar, with a long internal corridor lined with the skulls of hunted or sacrificed animals.		
Language registers	All groups have a complex system of multiple language registers		
Shaman	The shamans chant in a language generally inaccessible to ordinary speakers, which contains innovative lexical items, periphrases and idiosyncratic grammar.		
	Common through much of Arunachal Pradesh		
Hunting	Hunters use a lexical substitution language, whereby ordinary lexical items and animal names are replaced by unrelated lexemes, some of which can be etymologised, others of which are of unknown origin. Also in use by the Meyor.		
Others	Additional registers include mediation, babytalk, cursing and poetic.		
Ecozonal deities	All three groups share a common system of ecozonal deities which are responsible for different ecological zones from the snowline to the plains. In each the rivers and lakes are looked after by a deity with a name which is a variant of Bruu. The snowline is deemed to be the most sacred area, where shamans must travel to acquire their powers		
Complex afterlife	When the soul leaves the body it becomes a ghost and must travel through a complex series of underworlds, partly reflecting the type of death an individual		

Feature	Commentary
	has undergone. To reach the final abode of ghosts, the spirit must cross a river.
	Similar elements are also recorded among some Tani
Re Festival	A festival held on or around early February. [Idu name Re, Tawra, Kman
	Tamaladu]
Ethnometrology	A complex system of weights and measures, involving finger measurements,
	bamboo tubes, baskets, distance a pig can be carried and numerous other
	common features. Some aspects also recorded among Tani
Feature	Commentary
Negative imagery of	Unusually, all three Mishmi peoples consider flowers to be ill-omened and will
flowers	not plant them around the house <sup>2</sup> . This is particularly striking on the borders of
	Tibet, where flowers in pots are considered one sign of a well-kept house
Slavery	Rather unusually, for societies with segmentary lineage structures and little
	social hierarchy, the Mishmi peoples practised slavery, most commonly of their
	own people. Slavery was legally abolished during the late 1970s, but
	discrimination against slave ancestry persists. Sex with slaves was considered
	so unacceptable that a layer of the underworld was reserved for those who
	committed this offence.

There are many more apparent common features, but weak ethnography, both of the Tani and the Tawra, makes it impossible to claim with certainty that these are exclusive to the Mishmi. These striking similarities argue that at some point the cultural lives of the three Mishmi peoples were closely intertwined, so that such distinctive common features were shared. At a later stage, the Idu and the Tawra were intertwined to such a degree that there was a significant linguistic exchange between the two (documented in §6.2). However, following that the two must have split apart, because subsequently, it was the Tawra and Kman who became culturally intertwined. Despite their linguistic differences, Tawra and Kman are now held to be two aspects of the same culture and distinct from Idu. Since the 2000s, there has been a slew of local publications which include comparative wordlists, dialogues and culture guides in the two languages. Among these are reading and writing books which attempt to force them into a common orthography, despite the obvious differences in their phonology (e.g. Tawsik 2014, Kri 2015).

# 4. Common linguistic features of Mishmi languages

All three Mishmi language show numerous common typological features, many of which are replicated in neighbouring Tibeto-Burman languages. The puzzle is that they show few segmental cognates in their morphosyntax. Table 2 summarises these features, although they do not necessarily define a Mishmi group.

<sup>&</sup>lt;sup>2</sup> The Idu have an expression, ēmò ŋī wēsà lit. 'a dream sick would be', which refers to the avoidance of flowers following a dream

Table 2. Common typological features of Mishmi languages

Feature	Commentary
Three tone-heights	Typical of Naga languages
Voiceless aspirates only	Common throughout the region
SOV syntax	Common throughout the region
TAM marking through verbal suffixes	
No marking of number or gender on nouns	Common throughout the region
Distinctive gender suffixes for domestic animals	
Question markers clause-final	
Interrogatives clause-initial or following the head noun	
9-15 numeral classifiers	Strongly focused on shapes of plants and animals
Adverbs precede verb they qualify	
Four-term comparative adjective paradigm	
Adjectives show free positioning in relation to head	
noun	

The cultural evidence suggests that at some point Idu began to develop idiosyncratic cultural features, splitting away from the pool of those represented by Kman and Tawra. Table 3 shows a number of features where Idu is (?or has become) highly divergent;

Table 3. Linguistic features where Idu is highly divergent

<b>Feature</b>	Idu	Kman
Vowel nasalisation	Extremely common	Very rare
Vowel length	Present throughout	Absent
Retracted vowels	Present	Absent
Creaky vowels	Present	Absent
Verbal extensions	Rich system of CV suffixes and allows Bantu-like stacking	Virtually none
Adjectives	Large number of CVCVCV underived adjectives	Almost all adjectives transparently derived

# 5. Comparison of morphosyntactic markers

# 5.1 Number marking

All three languages, and indeed most of the isolate branches of Arunachal Pradesh, have no bound number marking in nouns. They make do with one or several suffixed markers, distinguishing animacy. Even these plural markers are often the same as the word for 'crowd' or 'herd'. Number marking is often omitted if plurality can be inferred. Table 4 shows number marking on nouns in Idu;

Table 4. Idu number marking on nouns

Idu	Application
àlòmbrò	persons
àŗū	crowd, herd
n-do	things, inanimates

Where plurality can be inferred, Tawra does not specifically mark number. Otherwise all plurality is marked with a single suffix  $-gr\dot{\delta}$ . Animates attract an additional suffixed marker when they occur in groups, as shown in Table 5;

Table 5. Tawra number marking on nouns

Gloss	sg.	pl. suffix	group	this group, that group
			animate	animate
dog	kwág	grá	táràw	wélàŋ
dao	tárá	grá		

Similarly, number in Kman is not marked on nouns when plurality can be inferred from the context. Number is marked with the suffix  $s\bar{s}n$  for animates and  $\bar{s}ns\bar{s}n$  for inanimates (Table 6).

Table 6. Kman number marking

Gloss	sg.	pl.	Gloss	sg.	pl.
person	tsõŋ	tsòŋ sə̄n	house	bwì	bwì ənsən
pig	<u>{ī</u> j	nēs {īj	tree	sən	sèŋ ənsən
mithun	cāl	cāl sēn	dao	sùt	sùt ənsən

The number marking systems for nouns in Mishmi languages show no segmental cognacy and only broad typological similarity.

## 5.2 Gender marking

Nouns are not usually marked for gender in Tibeto-Burman languages. Clearly gender distinctions are necessary in kinship terminology and basic discussion of men and women and all three languages have such terms, although they cannot be mapped across languages. Given the importance of domestic animals, salient species are also marked for gender, a pattern found all across Arunachal Pradesh isolates. As is common in regional languages, a three-way distinction between bovids, other animals and poultry is usually made. Domestic animals in Idu are marked for gender and also reproductive status (Table 7).

Table 7. Gender marking on Idu animal names

Idu	Gloss	Example	Gloss
àŗí	male animals, general	ili àŗí	boar
àŗố	male bovids	màcū àró	bull
àlā	male (birds)	ètō àlā	cock
àpí	female animal, general	ìlì àpí	sow
èchễ	female bird	ètō èchế	hen
krú	female bovids	sà krú	mithun cow

In Tawra, gender in domestic animals is marked by suffixed nouns,  $k \dot{a} r \dot{i}$  for males and  $k r \dot{u}$  for female mammals and  $t \dot{a} l \dot{a}$  for poultry (Table 8);

Table 8. Tawra gender marking for animals

Tawra	Gloss	Tawra	Gloss
másòw kàrì	bull	másòw krù	cow
màbì kàrì	he-goat	màbì krù	she-goat
tyù tàlà	cock		

There is clearly some overlap between Idu and Tawra. The term  $k \dot{a} r \dot{i}$  for male bovids in Tawra probably corresponds to  $\dot{a} r \dot{i}$  for non-bovids in Idu and Tawra  $t \dot{a} l \dot{a}$  for poultry to Idu  $\dot{a} l \bar{a}$ . Given the strong similarities, these may well be borrowings.

Kman marks a five-way distinction, treating cattle and mithuns as separate categories, and adding distinctive terms for monkeys (Table 9).

Table 9. Kman gender marking for animals

Kman	Gloss	Kman	Gloss
māntshūw kùwwà	bull	māntshūw nàw	cow
cāl ŋālōŋ	male mithun	cāl nàw	female mithun
lī? rēhàl	boar	lī? nàw	sow
èmùk kāŋà	male monkey	èmùk nàw	female monkey
krē āpày	cock	krē nàw	hen

Kman gender markers show no obvious segmental similarities with Idu and Tawra.

### 5.3 Numeral classifiers

A major point of comparison between the three languages is the system of numeral classifiers. These are the obligatory lexemes applied when certain nouns are enumerated. Such systems are found widely in the region and in particular are common in the Tani languages where classifiers are very numerous (Post 2007). The system can be described as residual, since many nouns, including most animates, no longer have obligatory classifiers. Idu mixes bound and free classifiers; free classifiers can be added to nouns to suggest the shape of an entity which does not take an obligatory classifier. To this extent, classifiers are something like the affixes in noun-class languages, such Niger-Congo. A noun may 'normally' take a certain prefix, but this can be altered to suggest a specific feature. Thus animals in folk-tales often take a human prefix, instead of usually assigned segment.

Partial lists of Idu numeral classifiers are given in Pulu (1978) and Pulu (2002), but a more complete version is shown in Table 10;

Table 10. Idu numeral classifiers

Form	Semantic cluster	Shape or class
āļà	cloth, paper, planks	flat rectangular objects
(-m)bồ	trees, plants, main trunk of anything	trunk (trees, body)
brā	potatoes, beads, oranges, stones	small spherical objects
brū	maize, bananas, small branch of a tree, usually cylindrical	long, cylindrical objects
(-n)do	bamboo clusters, houses, hunted deer, pig carcasses	?
(-ŋ)gō	fish, rats	
(-ŋ)gò	fish, rats	
ná	specific leaves, some types of paper	
ph(r)á	small packets (such as cigarettes)	small solid rectangular objects
pò	packets, léképò necklace	medium solid rectangular objects
pồ	bundles (firewood, hay), playing cards	large solid rectangular objects
prā	leaves, paper, flat things	flat objects
рū	elephants, chickens, wild birds, cucumber <sup>3</sup> ,	large and medium round and oval
	papayas, pumpkins	things
-to ∼ tõ	single bamboo plant, reeds with cylindrical stem, sugar-cane	giant grasses

The bound classifiers –(n)do and –to are underspecified for tone, and vary in accordance with the stem tone of the noun they qualify.

Tawra has a small number of numeral classifiers, which refer largely to shape, although including some natural semantic classes, such as buildings and doorways (Table 11).

\_

<sup>&</sup>lt;sup>3</sup> Local cucumbers are short and ovoidal, unlike the European cucumber, hence they fall into this class

Table 11. Tawra numeral classifiers

Tawra	Semantic cluster	Shape or class
brá	pebbles, potatoes	small spherical objects
br <del>í</del> l	thin trees, posts snakes	long cylindrical objects
bru	tree, banana, pen, spear	long thin objects
dố	buildings	large man-made structures
hàd <sup>(?)</sup> ùm	large trees, bushes, clusters of bamboo	large, ? spherical plants
ná	cloth, paper, leaf	thin flat objects
plá	banknotes, planks, knives	circular flat objects, long sharp objects,
pùm	boxes, elephant, birds, pumpkins	large round/square? hollow objects
tí	openings, doorways	
táŋ	animals	except elephants, fish, insects or birds

There are clearly some correspondences between Tawra and Idu. The bra class for small spherical objects, the bru class for long, cylindrical objects and na class for leaves and pieces of paper match exactly. The pu(m) class, which brings together elephants and pumpkins, includes boxes in Idu, but is otherwise similar. Tawra is distinctive in having semantic classes such as buildings, while Idu focuses almost entirely on shape.

Kman numeral classifiers refer almost entirely to shape, although they include two human classes and longhouses. They are obligatory when the head noun is not single. An exception to this is  $n\dot{a}w$ , which applies to individualised living things and which can take  $m\dot{o}$ , 'one', after the classifier. The classifiers are placed after the noun but before the numeral. Table 12 shows a provisional list of classifiers;

Table 12. Kman numeral classifiers and their categories

Kman	Semantic cluster	Shape or class Comment
bõŋ	bamboo, generic, wood	
bràt	oranges, grains, pebbles	spherical things
brül	small trees, twigs, bamboo splinters	long thin wooden objects
byôŋ	long pieces of wood, bamboo, cane	long broad wooden objects
dùm	short pieces of wood, bamboo, cane	short broad wooden objects
gàw	bamboo split lengthways <sup>4</sup>	half tubular shapes
glà	human beings	
klōŋ	leaves	flat round things
kũw	human beings (used in the context of	
	headhunting and criminality)	
krôŋ	longhouses	
nāw	individualised living animals	not plants
phàl	paper, banknotes, planks	flat long rectangular things
phòŋ	bamboo, grass, leaves	anything growing in clusters
tāŋ	cloth, mats, bags	? fibre household objects
thūŋ	pieces of wood, bamboo, cane	medium size wooden objects
tūl	standing trees, bushes	

Kman contrasts with Idu and Tawra in having classifiers for humans. Of the classifiers, only two show possible correspondences with Idu and Tawra, the brat class for spherical things and the brat class for long, thin, wooden objects. The close segmental similarities between Idu and Tawra suggest a pattern of borrowing rather than retention from a proto-language.

\_

<sup>&</sup>lt;sup>4</sup> But doesn't apply to a gutter, for example

### 6. Lexical datasets

### 6.1 Comparisons across Mishmi

Lexical comparisons across Mishmi produce a very low yield of even probable resemblances. Some examples are given in Table 13;

Table 13. Lexical resemblances across Mishmi

Gloss	Idu	Tawra	Kman
sword	shàbrē	shabrē	shàbrē
beer	уū	iyu	yūī (get drunk)
fish	àŋā	tã	ອ <del>ົ</del> ງຄີ
deafness	kàpà	kàpà	kāwà
road	ālố	alyɨm	b.lòŋ
stone	ālāphrấ	phlã	phlan (lower grindstone)
bedbug	àbā	àbà	māklàp
ginger	ànjítà	àdzìŋ	də́?ìŋ
granary	àkā	aka	kētèm
name	āmū	amaŋ	<u>ām</u> àŋ

The lack of regular correspondences again suggests borrowing at some earlier period of history.

# 6.2 Idu-Tawra comparisons which argue for a close relationship

A comparison between Idu and Tawra which above all has been responsible for the hypothesis of a 'Digarish' group is the similarities between lower numerals. Numerals 'two' and 'three' also resemble reconstructed Tibeto-Burman forms, which in the magical world of regional historical linguistics is enough for them to qualify for membership. Table 14 shows a comparison of lower numerals in the three Mishmi languages.

Table 14. Mishmi lower numerals

Gloss	Idu	Tawra	Kman
One	khègè	khin	kə̄mù
Two	kà.nyì	kayiŋ	kənin, kəyin
Three	kà.sỗ	kasaŋ	kāsèm
Four	kà.prì	kaprayk	kāmbrŵn
Five	màŋá	maŋa	kəlèn
Six	tāhrō	tahro	kətàm
Seven	íǜ	wẽ	$n \tilde{M} n$
Eight	ìĮú	l <del>i</del> m	grŵn
Nine	khrìnī	kɨŋaŋ	nētmù
Ten	hữữ	hálaŋ	kyēpmù

All the Idu and Tawra forms are cognate with the possible exception of 'nine' and many are extremely close. Except for 'two' and 'three', Kman numerals are quite different.

Another area where there are striking similarities is in basic colours. Table 15 shows the common colour terms recognised in Idu and Tawra, together with the abbreviated paradigm in Kman;

Table 15. Mishmi colour terms

Gloss	Idu	Tawra	Kman
black	mà, tĩ	ma	kā?yǜm
red	shù	shi?	kā?sàl
yellow	mì	miŋ	
green/blue	prù	prue	_
white	lo	lyo	kāmphlūŋ

Again, Idu and Tawra are extremely similar.

# 6.3 Idu-Tawra comparisons showing marked lexical divergence

However, in other semantic fields, the difference between Idu and Tawra is very marked. Basic body parts are often quite stable in Tibeto-Burman languages, but as Table 16 shows, Idu and Tawra share almost no common roots. Kman is included to illustrate that it shares virtually no potential cognates with Idu and Tawra.

Table 16. Idu and Tawra body parts

Gloss	Idu	Tawra	Kman
back	ìpìndò	phlíŋ	glàwk
body	jóntà	kyàŋ	shə̀y
breast	nōbrā	ŋèè	cīn
eye	ēlōbrā	blm	mīk
hand	ākhó	hàprè	râwk
leg	āŋgēsà	gròn	plà
lip	īnūbrū	thánù	chūw dàl
mouth	ēkób <u>à</u>	phùùkḕ̃	chûw
nail	āhữkò	áphlìŋ	zůk
neck	sēmbrá	pà hŋ	hūŋ
nose	ēnāmbó	ànàdùn	mī?nyùŋ
palm	lāpū	àtyòpà	rāwk tə̀pà
skin	kòprà	pô	ີນໆ
thigh	hàpū	sàhà	kātsàwk
toe	ātāmbó	gròn bràn	plā bàn
tongue	īlìná	hèlèŋnà	blây
tooth	tāmbrō	là	síí

The same divergence can be illustrated with other common lexical fields. A similar pattern emerges as in the case of morphosyntax, namely that where Idu and Tawra resemble one another the segmental cognacy is extremely close. Otherwise, they show little or no relationship. This points strongly to borrowing rather than genetic affiliation.

### 7. Tawra and Kman cultural convergence

As has been suggested above, the Kman and Tawra are undergoing strong cultural convergence, despite their highly divergent languages. The lack of linguistic interchange, in sharp contrast to the common lexicon of Meyor and Kman, suggests this process is relatively recent, despite the fervour with which it is promoted. A striking feature of this convergence is illustrated by the practice of shamanism and associated death rituals. Both the Kman and the Tawra have closely related shamanic chanting practices and indeed Kman shamans chant largely in Tawra. In the last few years this has found expression in innovative church-like organisations such as the Mishmi Faith Promotion Society (MFPS) which holds services in a mixed Kman/Tawra shamanic register (Photo 1).



Source: Author

Historical accounts mean that we can begin to put dates on the cultural convergence of Tawra and Kman. An striking aspect of this is the creation of a highly distinctive memorial for the dead. This is essentially a small model house where the possessions of the dead are placed, potentially for their use in the afterlife. Among the Kman this is known as a *khrām*, and we are fortunate to have from Dalton (1872: 16) a description from the 1850s. He says;

Monsr. Krick describes one that he saw at a funeral ceremony. This took place over the remains of the wife of a chief who had been dead and buried three months. The tomb was near the house covered with a root, under which were suspended the deceased's clothes. For several days previous to the arrival of the priest, an attendant was employed singing a mournful devotional chant to the accompaniment of a small bell. There was also a preliminary sacrifice of a red clock and hen, the blood of which was received in a vessel containing some other fluid, and the mixture carefully examined, as it is supposed to indicate if the result will be fortunate or otherwise. At last the priest arrived, dressed like an ordinary chief, but he wore a rosary of shell and, attached to the front of his head-dress, two appendages like horns.

The khram must be constructed within five days, but there is then an interval of up to three months before the talu ceremony, which depends on the wealth of the family, who must gather sacrificial animals. If the talu is pending for over three months, the  $k\bar{a}mam$  of the dead person may become angry and express its malevolence. Photo 2 shows a Tawra death house, recorded in a photo album from  $1916^5$ , incidental to its main concern, which is road and bridge construction.

-

<sup>&</sup>lt;sup>5</sup> Thanks to Sokhep Kri for the scans of this album

Photo 2. Death house, De'ing, 1916



Source: Anonymous, Lohit Valley Album

The *khrām* are still constructed today according to the same pattern, and their rich symbolism remains important to Kman people. Photo 3 shows a Kman death house photographed at Yatong, in 2015, with an almost exactly similar character to the one in Photo 2, a century earlier.

Photo 3. Kman death house, Yatong, 2015



Source: Author

Many other examples of this cultural convergence can be cited; the mystery of why and when it began is much harder to resolve, but the lack of linguistic interchange underlines strongly its recent chatacter.

### 8. Exploring the cultural mismatch

As many of the languages of the world are known only through lexical data, lexical cognacy is often the only method of determining language relationships. However, as far back as Sir William Jones and before it was argued that morphosyntactic paradigms constituted more solid evidence as they were less likely to be borrowed. In this view, the common lexical item good/*Gut* in English and German might be borrowed, but the paradigm good/better/best far less so. Whether this is actually so is a matter for debate. Satal, a Munda language spoken in Nepal, has developed the pronominal cross-referencing characteristic of its Tibeto-Burman neighbours and quite unlike the remainder of Austroasiatic (Tolsma 2002). Presumably, even the presence of this striking grammatical paradigm would not constitute evidence for Tibeto-Burman affiliation.

Related to this is the question of whether it is the case that there is some section of 'core' vocabulary which is more resistant to borrowing. Although this widely assumed to be the case, there seem to be a significant number of counter-examples. At the interface of Austronesian and Papuan, debates about the affiliation of languages such as the Reefs/Santa Cruz group show this is not easily applied (Ross & Naess 2009). The well-known example of Thai, borrowing Chinese numbers suggests that even basic numerals are not necessarily protected. Berbice Dutch, which borrows with equal facility from Dutch, English and Arawak in all areas of the lexicon suggests this principle cannot be taken for granted (Kouwenberg 1994). It may be that these are rare cases and that conservatism is dominant. But it warns us not regard lexical similarity as a gold standard.

Historical linguists may not regard these observations as rocket science, but are an important prelude to the case of Idu and Tawra. These languages have been classified together because a rapid examination of the numerals suggests they are not only related, but very close to one another. The relationship between Idu and Tawra is less clear than a cursory examination of the numerals would suggest. The closeness of the numerals and the remoteness of many other areas of the lexicon, suggests intensive but highly selective borrowing or else a mysterious process of vocabulary replacement along the same lines. In other words, either Idu and Tawra *are* closely related but some now non-operative process has led to replacement, say of body parts, or else they are *not* related but some vanished interaction allowed intensive borrowing in the apparently related areas. Either case would surely be typologically very unusual. On balance, the fact that most morphosyntactic markers we understand are different, although typologically similar suggests that Idu and Tawra are unrelated. We can say that it is unlikely more data will resolve this issue; the existing vocabulary is large enough that it is unlikely unsuspected sound correspondences will appear in more obscure items.

# 9. West of Tani – a parallel case of convergence?

There is an almost exact parallel west of the Tani languages among the Miji, Hruso and Koro peoples. Despite a striking absence of common lexicon and morphosyntax (as far as this is known) their cultures map strongly against one another. This relationship has been obscured by some extremely foolish publicity concerning the 'hidden' Koro language, fuelled by the American television and media enterprise, National Geographic<sup>6</sup>. Koro has hardly been hidden; indeed, as Photo 4 shows, the Koro people have done their best to publicise their culture and language through music videos. The Koro, however, have a curious relationship with the neighbouring Hruso. Both Hruso and Koro are locally classified using the term 'Aka' which is commonly found in earlier literature. The Koro appear in earlier publications as the 'Miri-Aka' which may have been responsible for some of the confusion or misleading statements in media reports. This has led to the mistaken label 'Hrusish' being perpetuated by scholars who should know better; there is no evidence for a group of languages which include Hruso. The two cultures are a mirror of one another, despite the fact that the languages have virtually nothing in common. Almost all lexicographic items can be mapped to one another, a point is made clearly in the comparative study of Grewal (1997). The same is true for the Miji language, west of the Hruso, although the Miji are more culturally distinct. The pattern linking Idu and Tawra is replicated in this group, with occasional lexical similarities more suggestive of borrowing than genetic affiliation.

### 10. Conclusions

The Mishmi languages present a striking mismatch between local perceptions of cultural relationships and the linguistic facts, as far as they are known. Historically, all three languages are most likely of distinct origin, but living in a common environment, in close physical contact and in constant interchange set in motion powerful cultural levelling. There must have been a period when all three groups shared a common culture, in order to account for the basic similarities still observable.

<sup>&</sup>lt;sup>6</sup> See http://news.nationalgeographic.com.au/news/2010/10/101005-lost-language-india-science/

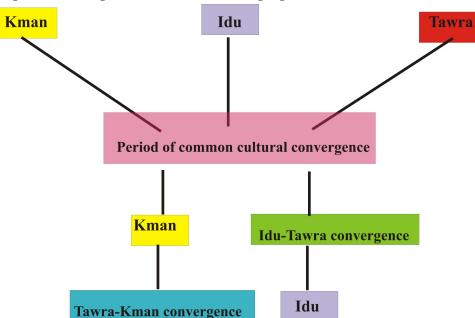
Photo 4. Hruso/Koro music video poster



Source: Author photo

At some point, Tawra and Idu became intertwined, hence the intensive borrowing in restricted areas of the lexicon. Much more recently, the Tawra/Idu bond was broken, and the construction of a relationship with the Kman people began. Figure 1 is a schematic model of the historical relationship between the three Mishmi languages. It has no dates attached and is far from explaining their unusual interchanges. Nonetheless, it represents of model for guiding future research into this striking nexus.

Figure 1. Convergence model of Mishmi languages



The puzzle of why there has been so little lexical interchange between the three groups and why, when it occurs, it is so selective, remains to be solved. It is suggested that in global terms this situation is highly unusual and challenges many common generalisations in sociolinguistics.

# **Bibliography**

Aisher, Alexander 2007. Voices of Uncertainty: Spirits, Humans and Forests in Upland Arunachal Pradesh, India. *South Asia: Journal of South Asian Studies*, n.s., Vol.XXX, no.3

Baruah, Tapan Kumar 1988. *The Idu Mishmis*. Itanagar: Government of Arunachal Pradesh. [reprint of a 1960 publication]

Bhattarcharjee, Tarun 1983. *The Idus of Mathun and Dri Valley*. Shillong: Government of Arunachal Pradesh. [original publication 1960]

Boro, A. 1978. Miju dictionary. Shillong: Research Secretariat.

Campbell, G. 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.

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

Cooper, T.T. 1873. The Mishmee Hills: an account of a journey made in an attempt to penetrate Thibet from Assam to open new routes for commerce. London: Henrhy S. King & Co.

Dalton, E.J.T. 1872. *Descriptive Ethnology of Bengal*. Calcutta: Office of the Superintendent of government printing.

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

Griffith, W. 1847. Journals of Travels in Assam, Burma, etc. Calcutta.

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.

Kouwenberg, Silvia 1994. A grammar of Berbice Dutch Creole. Berlin & New York: Mouton de Gruyter.

Kri, Doirangsi 2011. A conversation book of Kaman and Taroan Mishmi. Delhi & Itanagar: Himalayan Publishers.

Li, Daqin 2002. *Geman yu yan jiu*. (*Zhongguo xin fa xian yu yan yan jiu cong shu*) [A study of Geman]. Beijing Shi: Min zu chu ban she. [Beijing: The Nationalities Press.]

Macgregor, C.R. 1887. Journal of the Expedition under Col. Woodthorpe, R.E., from Upper Assam to the Irrawadi and return over the Patkoi Range. *Proc. Royal Geographical Society*, 9: 19-42.

Mills, J.P. 1952. The Mishmis of the Lohit Valley, Assam. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 82(1):1-12.

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.

Pulu, Jatan 1978. Idu phrase-book. Shillong: Arunachal Pradesh Directorate of Research.

Pulu, Jatan 1991. A phrase book on Taraon language. Itanagar: Directorate of Research.

Pulu, Jatan 2002a. Idu Mishmi proverbs and sayings. Itanagar: Arunachal Pradesh Directorate of Research.

Pulu, Jatan 2002b. A handbook on Idu Mishmi language. Itanagar: Arunachal Pradesh Directorate of Research.

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

Rowlatt, E.A. 1845. Report of an Expedition into the Mishmee Hills to the N.E. of Sudyah. *Journal of the Asiatic Society of Bengal*, 14, 2: 477-493.

Sastry, G.D.P. 1984. Mishmi grammar. Mysore: Central Institute of Indian Languages.

Sastry, G.D.P. 1991. Mishmi-English-Hindi dictionary. Mysore: Central Institute of Indian Languages.

Sun, Hongkai 1983. 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.

Sun, Hongkai 1999. On the Himalayan languages of the eastern Himalayan area in China. *Linguistics of the Tibeto-Burman Area*, 22: 61-72.

Sun, Hongkai et al. 孙宏开 孙宏开 1980. 们巴,珞巴,僜人的语言 [Menba, Luoba, and Deng peoples' languages], 174--231. 北京 Beijing: 中国社会科学出版社 China Social Sciences Press.

Sun, Hongkai 孙宏开 (ed.) 1991. 藏缅语语音和词汇 Zang Mianyu yuyin he cihui. [Tibeto--Burman phonetics and vocabulary]. Vocabulary 49 北京中国社会科学出版社 Beijing: Zhōngguó Shèhuì Kēxué Chūbǎnshè [Beijing: China Social Sciences Press].

Wilcox, R. 1832. Memoir of a Survey of Assam and the neighbouring countries executed in 1825-6-7-8. *Asiatick Researches*, 17: 314-469.