The linguistic position of Tani (Mirish) in Tibeto-Burman: A lexical assessment*

Jackson T.-S. Sun
Institute of History & Philology
Academia Sinica

INTRODUCTION

The obscure Tani (Mirish, Mishingish) languages of southern Tibet and Arunachal Pradesh have only recently begun to receive the attention they deserve (Chhangte 1990, 1992; Sun 1993, 1994). The aim of this paper, which is part of an ongoing project to study the phonological and lexical diachrony of these languages, is to contribute toward clarifying the linguistic position of Tani languages in the Tibeto-Burman family from the vantage-point of reconstructed Proto-Tani (hereafter PT).[1]

Section 1 surveys and contrasts existing views on the affiliations of Tani in Tibeto-Burman. Section 2 inspects in detail a number of Tibeto-Burman languages which have been nominated in the literature as possible close relatives of Tani. After screening out a few unlikely contestants, a pilot lexical study is conducted in section 3 to weigh the degrees of lexical affinity between Tani and the remaining candidates as compared with three control languages, Written Tibetan, Written Burmese, and Garo. The implications of the output of this study on the phylogenetic position of Tani are then discussed. In the concluding section, we consider the nature of the relationship between Tani and Digarish (consisting of two known languages: Taraon and Idu), the language group which turns out to be most akin to Tani in basic vocabulary.

1. EXISTING VIEWS ON THE PLACE OF TANI IN TIBETO-BURMAN

The genetic affiliations of Tani with Tibeto-Burman have seldom been called into question,[2] and should now be considered *proven beyond reasonable*

[1] A preliminary phonological reconstruction of Proto-Tani is proposed in Chapter II of Sun 1993, from which the reconstructed PT roots cited herein are taken. The Proto-Tibeto-Burman (hereafter PTB) reconstructions are based mainly on Benedict 1972 (hereafter STC).

[2] The great lexical differences between Tani and other Tibeto-Burman languages (only 12.5% agreement of basic vocabulary with Tibetan and Burmese according to his calculation) has led Marrison to doubt not only the genetic affiliations of Tani with Tibeto-Burman, but also "the

doubt in view of the accountability of much of the PT phonological developments in terms of PTB (Sun 1993, chapter IV).[3] However, there is no consensus yet as to how Tani interrelates with other Tibeto-Burman languages. In fact, as shown in the following survey of the subgrouping literature, opinions diverge sharply from each other with regard to both lower-level and higher-level affiliations of Tani in Tibeto-Burman.

1.1. Konow: 'North Assam'

In the colossal Linguistic Survey of India, Tani languages, along with other little-known Tibeto-Burman languages of Arunachal Pradesh, were brought together in the so-called 'North Assam' group. This was clearly meant to be an expedient, geographical grouping, as shown in the following quote from Sten Konow, the linguist originating this term (Konow 1909:568, 569, emphasis ours):

> The North Assam group is *not a well-defined philological group* with salient grammatical features distinguishing it from other Tibeto-Burman forms of speech...In many important points, however, Mishmi[4] differs from Abor-Miri, and *the points of correspondence just referred to are not of an importance sufficient to prove a close connexion between the two forms of speech.*

As for higher-level connections, Konow made only a vague suggestion (Konow op. cit.:572):

> The North Assam forms of speech can be described as links which connect the Tibetan and Himalaya dialects with the languages of the Bodo, Naga, Kuki-Chin and Kachin groups.

---

reality of the Tibeto-Burman language family as generally accepted...The Tibeto-Burman family is an unsatisfactory construct, and this whole field of investigation should be reopened" (Marrison 1988:216). My own lexical study, however, has turned up much higher cognate figures between Tani and both Tibetan and Burmese (see 5.3. below). Even if Marrison was right about the cognacy rates, his radical view on the status of Tibeto-Burman, we believe, would be hard to accept for most Sino-Tibetanists.
[3] For instance, the regular sound correspondence between PTB *-əy and PT *-i is backed up by as many as eleven cognate sets, all belonging to basic vocabulary (Sun 1993: 4.3.1.2.).
[4] As shown by ensuing research, the Mishmi languages do not form a coherent linguistic unit either. Rather, there is a fundamental cleavage between Digaro-Chulikata-Midu (Taraon-Idu) and Miju (Kaman). Thurgood 1985:81 claims that the Mishmi languages belong with Nungish under a supergroup 'Kaman-Nung' with 'fully substantiated lower-level genetic relationships'. We believe that this claim, which remains totally unproven, underestimates the great differences between the two Mishmi groups (for a more conservative view, cf. Sun et al. 1980:299-315).

1.2. Shafer: Mishingish (Bodic/Burmic)
The distinctness of the 'North Assam' languages is further underscored in Shafer 1955:102, where no less than four separate groups are recognized: Mishingish (= Tani), Digarish (= Taraon-Idu), Midźuish (= Kaman-Meyöl), and Hrusish (= Hruso = Aka). Shafer did not attempt a further classification but suggested that all of them are 'possibly sections of Bodic, possibly of Burmic, *certainly not of Baric*' (Shafer op. cit.:102).

1.3. Benedict: Mirish (Major Tibeto-Burman Nucleus)
While positing Abor-Miri-Dafla (i.e. Mirish in the narrow sense = Tani) as one of the major nuclei of the Tibeto-Burman family, Benedict (1972:5) suggests that to this division perhaps also belong not only the three Arunachal neighbors of Tani: Taraon, Kaman, and Hruso, but also the geographically more distant Dhimal group of Sikkim and Nepal. This claim, in effect, upgrades for the first time Konow's 'North Assam' from an *areal* to a *genetic* grouping. He further speculates that this group (Mirish in the extended sense) could ultimately be linked with Kachin (Jingpo), Baric (Bodo-Garo and Konyak), Nungish, and Lolo-Burmese under the supergroup 'Burmic' (op. cit.:11). This view was soon given up. In Benedict 1976:178; fn. 14, he proposes instead that, as far as core vocabulary is concerned, Tibetan, Chepang, Tamang (i.e. Bodic), Burmese-Lolo-Nungish, Lushai (Kuki-Chin-Naga), and Miri (Tani) form one supergroup as against Kachin, Garo, Konyak languages, and Chairel (or Chakpa, a Luish language according to Bradley 1993:7).[5]

Benedict's revised view on the linguistic position of Abor-Miri-Dafla (AMD = Tani) can thus be interpreted as follows: At a lower-level, AMD is most closely related to Hruso, Taraon, Kaman, and Dhimal; these languages are allied further with Lolo-Burmese, Bodic, and Kuki-Chin-Naga, as against Kachin and Baric. It is important to note that while Benedict ventures explicit claims about possible lower-level close relatives of Tani, he agrees with Shafer that *Tani is not akin to Baric.*

1.4. Other Ideas
Egerod 1974 also contains a classification of Tibeto-Burman, founded largely on Shafer and Benedict's frameworks. According to Egerod, Mirish (=

---

5 Incidentally, Benedict's revised view on the special relationship between Jingpo, Bodo-Garo, and Northern Naga seems to be receiving growing endorsement (Burling 1971, 1983; French 1983). The most drastic move in this direction is taken by Weidert 1987: fn.22, where Jingpo is put directly under one of the three branches of Barish: Western Barish (= Bodo-Garo, or Burling's Garo branch); Eastern Barish-I or Arunachal Barish (= Tangsa, Nocte, Wancho); and East Barish-II (= Konyak, Phom, Chang, Khlamngan, and *Jingpo*). An alternative view groups Jingpo rather with Lolo-Burmese, forming a '*Jiburish*' subgroup on the strength of hundreds of cognates between Jingpo and Lolo-Burmese and some parallel phonological developments (Matisoff 1974). In Matisoff 1991:481, however, Jingpo (Kachinic) and Lolo-Burmese are treated as separate major Tibeto-Burman subgroups.

Tani) is one of the major branches of Tibetic (= Shafer's Bodic); further, all of the other sections (Dhimalish, Digarish, Midźuish, Hrusish, Newarish, and Dzorgaish) left unclassified between Bodic and Burmic by Shafer are directly assigned to 'Other Tibetic'. Although further genetic subrelations among these Tibeto-Burman groups are not explored by Egerod, it is clear that, like Shafer and Benedict, he does not consider Mirish to be closely affiliated with Baric.

Matisoff 1991, DeLancey 1991 and Bradley 1993 are among the most recent statements on the genetic relationships among the Tibeto-Burman subgroups.[6] Incorporating information on the newly described Tibeto-Burman languages as well as some recent low-level subgrouping proposals, they all depart in significant ways from their predecessors. In DeLancey 1991, an expanded notion of Baric is suggested, subsuming not only Bodo-Garo and Konyak-Naga (= French's Northern Naga), but also Kuki-Naga, Kachinic (Jingpo), and Mirish. What is more, in this classification 'Mirish' includes the three Mishmi languages in addition to Tani proper, but not Dhimal (assigned to Bodic) or Hruso (not mentioned in his framework). This extended conception of Baric may be inspired by the geographically-based *Kamarupan* (i.e. Assamese Tibeto-Burman) group first proposed in Matisoff 1985b: fn. 8, where, however, the term is explicitly stated to be 'a neutral overall designation for the TB languages of NE India and adjacent areas'. Matisoff 1991:480-1 proposes a simplified heuristic subclassification model of Tibeto-Burman with seven major Tibeto-Burman subgroups including Kamarupan (again with the disclaimer that this is a 'purely geographic rubric'), under which we find Kuki-Chin-Naga, Mikir, Meithei, Mru, Bodo-Garo, as well as Abor-Miri-Dafla. Unlike DeLancey's Baric, however, Kamarupan does not include Jingpo, which is assigned to form a subgroup (Kachinic) by itself. Bradley 1993 contains a wealth of valuable new demographic and sociolinguistic information, especially concerning the Tibeto-Burman languages of India and Burma, but is unconventional in many ways. Adopting purely geographic labels, Bradley classifies Tibeto-Burman into four major groups: Western (Bodic), Northeastern India (= Burling's Sal group plus Kuki-Chin-Naga and Luish), Southeastern (Burmese-Lolo and Karenic), and Northeastern, a tentative medley group containing not only languages which Benedict 1972 puts under Mirish (i.e. Tani, Hrusish, Dhimalish, and the Mishmi languages), but also Nungish, Qiangic,[7] and some widely divergent minor languages such as Sulung and Bugun.

---

6 Dai 1989 and Sun 1988 are not considered here because they deal only with the Tibeto-Burman languages of China.
7 Surprisingly, Bradley's Qiangic also includes such languages as Tujia, Baima, and even Bai. Chinese scholars now tend to regard Baima as a divergent dialect of Tibetan (Zhang Jichuan and Huang Bufan, p.c.)

It is evident that there is hardly any agreement among Tibeto-Burmanists today concerning the precise linguistic affiliations of Tani in Tibeto-Burman. While this indeterminacy reflects the immature state of higher-level Tibeto-Burman subclassification in general (Thurgood 1985, Sun 1988, Dai 1989, DeLancey 1991, Matisoff 1991),[8] the uncertainty surrounding the linguistic position of Tani and related languages in particular can be directly attributed to the shortage of comparative data essential for recovering the linguistic histories of these languages, which in turn makes definitive subclassification well-nigh impossible.

Yet, what is relatively uncontroversial is that languages of the Tani group (i.e. Shafer's Mishingish, Benedict's Mirish in the narrower sense) do form a compact unit, more closely related to each other than to any other Tibeto-Burman language. We think it is important for the clarification of this issue to assert with certainty that *no other Tibeto-Burman language known to us deserves a place on the same taxonomic level as the two major Tani subgroups (Eastern and Western Tani).* Hence, earlier proposals which subsume languages like Midu (Thurgood 1986:93),[9] Aka (Nishida 1979:77), or Sulung and Bangru (Sun 1983:267)[10] directly under Tani proper are untenable. This is not to deny, of course, that Tani may not be grouped further with other Tibeto-Burman languages in a co-ordinate relation under some higher Tibeto-Burman division, the topic of the next section.

2. POSSIBLE CLOSE RELATIVES OF TANI

What, then, are the *collateral relatives* of Tani proper in the Tibeto-Burman family? A number of languages have been mentioned in the literature as showing particular affinity with Tani, including Lepcha (Bodman 1988); rGyarong (Nagano 1984); Dhimal (Benedict 1972, Bradley 1993); Hruso (Benedict 1972; Nishida 1979, 1984; Bradley 1993), and the Mishmi languages (Benedict 1972; DeLancey 1991:431; Bradley 1993). These proposals will be considered below in light of our improved understanding of the Tani evidence.

---

[8] One important reason for the lack of agreement in Tibeto-Burman subgrouping may be the different criteria (often implicit) used in the various subgrouping proposals. Thus, Thurgood puts Nung in his 'Rung' supergroup apparently on morpho-syntactic evidence only (Thurgood 1985). What is not explained is the considerable amount of shared basic vocabulary between Nung and Lolo-Burmese (STC:8; Benedict 1976: fn 14).

[9] Thurgood claims that 'Even from the limited LSI sample of data, it is clear that the Chulikata Mishmi [=Midu]...must be subgrouped with these Adi languages rather than with the Miju language' (Thurgood 1986:93). Actually, Midu should be equated with Idu (autonyms: Idu, Midu, Dudu), which Thurgood in the same paper correctly assigns to the Taraon group.

[10] Sun Hongkai's tentative inclusion of Sulung and Bangru under the Nishi-Bangni subgroup of Tani (Sun 1983:267) was done apparently at a time when linguistic data on these languages was not yet available to him. His more recent view is that Sulung and Bokar (other Tani languages are not mentioned) are distinct languages belonging to the 'Jingpo' supergroup, which also contains Jingpo, Nungish, and the Mishmi languages (Sun 1988:69).

2.1. Lepcha

The phylogenetic position of Lepcha, a Tibeto-Burman language of Sikkim, has also been highly controversial. Earlier analyses have aligned Lepcha with Naga (specifically, the 'Northern Naga' branch of Shafer 1955:106),[11] Tibetan-Kanauri and Kiranti (Benedict 1972:7-8), and Mikir (Bauman 1976). In a valuable recent revisit to the issue, Bodman (1988) compares Lepcha with a number of Tibeto-Burman languages which are lexically most similar to Lepcha, including an unidentified variety of Adi extremely similar (if not identical) to Padam. The substantial evidence of the lexical affinity between Lepcha and Adi comprises a list of 130 cognate pairs, based on which some important Lepcha-Adi consonantal correspondences are worked out.

On close inspection, however, many such sets appear to be *common retentions* from the original Tibeto-Burman lexical stock, and do not demonstrate by themselves any special lexical relationship between Lepcha and Adi. They include the following: *blood, blow, bow (weapon), carry on back, crab, cry (weep), dig, dream, drink, eat, eye, fire, fish, five, flat, four, give, horn, male of animals, leech, otter, ripe, road, seed, shade, smell v., snake, son-in-law, star, stone, three, tongue, two, and wood.*

Furthermore, the cognacy of the following items seems doubtful:

• 'sew' Lepcha *hrap*, Adi om-*kap*: The true Tani root for 'sew' is the first element om- (< PT *ɦom); the second element -kap,[12] on which the comparison is based, is a verbal particle signifying 'closure'. Thus, the precise meaning of Adi om-kap is 'sew up'. This makes Adi -kap semantically less compatible with the Lepcha form.
• 'spirit' Lepcha *a-pil*, Adi a-bɯr a-jo (listed as a-bɯn a-jo in Lorrain 1907:361; a typo?). The Adi form a-bɯr a-jo can indeed mean 'spirit', but the phonology does not match (Like Lepcha, Padam preserves -1, but the form in question ends in -r).
• 'crumb' Lepcha *p'yol*, Adi pim-*pil*: The Lepcha form, which does not mean 'crumb' at all, is an adverbial which occurs in reduplicated form p'yol p'yol (e.g. *p'yol p'yol* glo nóŋ 'to fall into pieces'). The Adi word is a compound composed of the 'grain' root PT *pim plus an element pil (< PT

---

[11] This is not the same as the 'Northern Naga' (or Konyak) languages of French 1983. Rather, it refers to the group of Naga proper which Weidert terms Naga-II, comprising Ao, Lhota, Sangtam, Yimchunger, and Northern Rengma (Weidert 1981: fn. 3).
[12] From PT *kap 'cover'. Cf. rGyarong pkap; Jingpo mă³¹kap³ⁱ; Dulong ta⁵⁵kɔ̌p⁵⁵; Kaman ŋkhap⁵³; WT 'kheb~'gebə 'cover'.

*pjɯl) which refers to small rounded objects in general and appears also in such compound words as 'grain', 'coin', 'uvula', 'clitoris', and 'kidney'.

• 'dig up' Lepcha *bɔl*; *byol*, *byul*; Adi du-*bur*. The Adi compound, which has a more specific meaning of 'dig up (earth) and make it powdery', contains the morpheme du-, the real root for 'dig' (< PT *du); the -bur element, semantically incompatible with the Lepcha forms, is a (resultative) verbal particle meaning 'so as to be powdery'.

• 'beetle' Lepcha *bŭt*, Adi je-*pɯt*. The Lepcha word is glossed 'insect that eats and causes destruction' in Mainwaring-Grünwedel 1979:258, and seems to be derived from the verb bŭt meaning 'pulverize, decay (of tooth)'. The Adi form, on the other hand, refers to 'scarab, dung beetle' and is transparently composed of je 'excrement/ dung' plus pɯt 'burrow/bore v.'.

• 'steep' Lepcha dóp, Adi tap-. The Adi form seems to be a resultative verbal particle which means rather 'down, become horizontal (of something upright, e.g. a tree)'.

• 'stick, adhere' Lepcha *krap*, Adi *gap*. The Adi morpheme, which appears in the compound geŋ-*gap* 'adhere/stick to', actually means 'grasp/hold' and is here used as a resultative verbal particle after geŋ-, the true root for 'adhere, stick, heal'.

• 'close (v.i.)' Lepcha *zap*; Adi a-*dap*. The central meaning of the Lepcha root zap is 'place compactly'; zap seems to take on the meaning 'close together' only in an adverbial phrase sǎ-zǔ-sǎ-*zap*.

The following pairs seem to involve convincing cognates; however, further comments can be added to them:

• 'divide, distribute' Lepcha ór, Adi or. The two words involve different (nevertheless related) meanings in the respective languages. The Lepcha form means 'separate (people or things) that which are close together', whereas the Adi form (< PT *hor) means rather 'distribute'.

• Lepcha *rŭm* 'god', Adi u-*rom* 'ghost': Lepcha rŭm seems to refer more generally to 'benevolent spirits' and is thus semantically closer to the Adi word, which is from PT *rom 'ghost (ancestral)' (contrast PT *ju 'evil spirits').

• 'pubic hair' Lepcha *mǎt*, Adi a-*mɯt*. Actually, the semantics of the given roots in both languages goes beyond 'pubic hair'. The Adi form goes back to PT *mɯt, a general 'hair' root (for both body hair and hair of head). The Lepcha root mǎt also appears in the compound bon-*mǎt* 'beard (mouth-hair)'. Also to be noted is the shared -t final, rarely found in Tibeto-Burman words for 'hair'. The cognacy of these forms to PTB *mul is dubious, as there is otherwise little evidence for the *-l > -t shift in either language. In fact, PTB *mul is directly attested in the Lepcha doublet a-*myal* 'body hair, feathers, armor', as well as in the Mising L forms nam-*mur*; soŋ-*mur* <*nap-mul; *čok-mul 'beard' (PT *čok 'chin/jaw').

• 'taboo, omen' Lepcha nyo, Adi ño. The Padam Adi form is a verb which means 'be tabooed or quarantined for religious reasons'; the Lepcha form is

glossed as 'be ominous, have a bad effect'. The really remarkable fact, not mentioned by Bodman, is that both of these forms show the same variant form with -t (suffix?): Lepcha nyot; Padam-Mising L ñot!

What is surprising about Bodman's comparative list is that many cases of plausible lexical comparability between Lepcha and Padam Adi (Eastern Tani) coincide with the east-west lexical split among Tani languages, and the forms more common in Western Tani do not resemble the Lepcha forms at all. Consider the following examples:

• 'breeze' Lepcha fár, Adi a-sar. This is an Eastern Tani word; cf. Western Tani: *rji (< PTB *g-ləy).
• 'swell' Lepcha bróm; Adi pom (< PTB *(s-)bvam). This form appears to occur in Padam only; other Tani < PT *bruŋ (< PTB *bliŋ~pliŋ 'full').
• 'fear' Lepcha ro(-m), Adi le-ro. Milang T a-re-ma; Padam-Mising L le-ro; aŋ le-lo (aŋ ='heart'); other Tani < PT *pV-so ~ bV-so.
• 'sky, heaven' Lepcha tă-lyañ, Adi ta-leŋ ~ ta-jeŋ. This is mainly an Eastern Tani form (see section 3.2.2).
• 'return, (give) back' Lepcha lót, Adi -lat. This form, another verbal particle, is used only in Eastern Tani; contrast Western Tani -kur.
• 'girdle' Lepcha a-rek, Adi mag-rek. This form is found in Padam only.

We can also contribute a few more items to the list of Lepcha-Tani comparabilia:

• Lepcha pán 'be forgetful, absent-minded', PT *mit-pan 'forget' (PT *mit- = 'extinguished').[13]
• Lepcha pán 'break off v.i.' vs. Lepcha fán (< *ph-?) 'break off v.t.'; Padam-Mising L ben~bet 'break off v.i.'; Padam-Mising L pen~pet 'break off v.t.'. This is one of the rare instances where Tani preserves the familiar Tibeto-Burman transitivity-based voicing alternation (cf. Xiandao Achang bio '(of thread) be broken v.i.' vs. phio 'break (thread), v.t.'; Taraon bɹun⁵³ '(of ropes) be broken' vs. phɹun⁵³ 'break (ropes)' (Sun et al. 1980:205).[14]
• 'nest' Lepcha a-šap; PT *sup.
• 'revolve in mind; reason' Lepcha myón; PT *muŋ 'think'.
• 'take' Lepcha lóŋ; PT *laŋ.
• 'bowels' Lepcha tă-klí; PT *kri. Matisoff 1978a:214-5 suggests that these forms may originate from PTB *kləy 'excrement'. The root also occurs in compound words for 'belly' and 'navel' in Tani, but not in Lepcha.
• Lepcha mlo 'world, universe'; PT *mroŋ 'world/land/earth'.

---

13 Cf. Damu OY mit-pan to-mit 'forget'. Prof. Matisoff suggests that the *mit- element may reflect PTB *m-yit 'mind'. This is possible, but the normal PT 'mind/think' root is *muŋ.
14 Causativity in modern Tani is normally expressed by means of affixation (usually involving the morpheme 'do/make' mo:) rather than by stem-modification.

We have shown that although Bodman's original list of Lepcha-Adi comparisons needs revision, the rather remarkable lexical tie between these languages cannot be overlooked. In addition to a few new items added to the list (further search will doubtlessly uncover more), we have also made the discovery that despite the geographical location of the present Lepcha-speakers to the west of the Tani language area, it is in Eastern Tani (particularly Padam Adi), that the more striking similarities are found. Does this mean that Lepcha and Tani are close kin on the Tibeto-Burman genealogical tree? We will defer judgement until this issue is further explored below.

2.2. rGyarong

We now turn to rGyarong, another language supposedly showing special affinity to Tani according to Nagano 1984. One of the most noteworthy claims in this work is that rGyarong in its deepest lexical stratum is more intimately related to AMD (i.e. Abor-Miri-Dafla) than to either Tibetan (the traditional view) or Qiangic (a view espoused by leading Qiangic specialists of China; see for instance Sun 1982 and Huang 1991).[15] In order to demonstrate this new linguistic alignment, Nagano presents a comparative list of about a hundred core vocabulary items (mostly verbs) with which to establish sound correspondences between the GC (i.e. lCog-rtse) dialect of rGyarong and AMD. The AMD data is taken from Yano B unless otherwise stated (actually, forms are often cited from the distinct Tagen B variety instead), interspersed with Abor-Miri forms (hereafter AM) taken from Lorrain 1907. To one's puzzlement, Ao Naga and Mikir forms are included under the AMD heading, though these languages had never been considered to belong to the AMD group. What is also peculiar is Nagano's decision to use modern lCog-rtse forms instead of reconstructed Proto-rGyarong roots in his rGyarong-AMD comparison.[16] Rather than presenting a thorough review of the rGyarong-Tani lexical connections suggested by Nagano, the following sample set of comparisons supposedly representing rGyarong-Tani *dental-stop* correspondences (Nagano 1984:142), will be examined; the highlighted segments in the GC and AMD forms therein being the proposed equations:

• 'dig': GC **tuʋ**, Yano B **du-to**. The Yano B form goes back to PT *du which, like the rGyarong form, is a reflex of the prevalent PTB etymon *du~tu (STC #258). This is a common TB root attested in various TB branches and cannot be regarded as evidence of a special lexical link between rGyarong and Tani.
• 'hit': GC **tom**, AM **dem**. This rGyarong form is derived from PTB *dup~dip; *tup~tip 'beat' (STC #399). The nasal-final form **tom** 'I shall hit' is clearly

---

[15] Wolfenden 1936:168 also suggested that rGyarong may be a moderately close surviving relative of Xixia (Tangut), which is now generally considered to be another Qiangic language (Sun 1988:67, Matisoff 1991: 482).
[16] This is perplexing given the general principle that if two languages bear a true genetic relation, then the further back one traces their histories, the more similar they should be.

secondary (< top + ŋ), cf. the infinitive form ka-top from the same lCog-rtse dialect cited in Anonymous 1991 (hereafter ZMYYC):1081 and Qu 1984: 79. Padam-Mising L dem has a more specific meaning 'beat (with a stick, etc.)' and is clearly a separate root. The true cognate with rGyarong -top 'hit' is rather PT *tup 'strike', both being reflexes of PTB *tup.

• 'big': GC kte; Yano kte. No such Yano B form exists. The real Yano B root should be just -tè, a bound morpheme occurring with classifiers. Again, both forms may reflect a common PTB root *tay (STC #298).

• 'see': GC mto; Yano kâ-to. This is a misinterpretation. Instead of the real root kâ (< PT *kaŋ) 'look/see' which is mistaken for a 'prefix' (op. cit.:90), the Yano morpheme selected for comparison, -to, is an imperative marker which appears on all citation-form verbs in Bor's Yano-Tagen wordlist.

• 'straight': GC sto; AM adong. This Padam L form actually means 'long' (cf. PTB *duŋ, STC p.19) rather than 'line', contra op. cit.:143.

• 'cold': GC sytak (i.e. [ʃtak]); Yano po-teng-pa. This Yano B form actually means 'dry (of clothes)' (cf. Bengni S pɯ-tuŋ). We fail to see any possible connection, formal or semantic, between these GC and Yano words.

• 'go': GC thal; AM to. The AM form is unknown. As far as we know, no Tani language has this form with the given meaning.

• 'put': GC tha; AM tâk. The rGyarong form exemplifies a well-attested Tibeto-Burman root PTB *ta (STC #19), with an open rhyme. The AM form, occurring in a compound tak-po 'put (cover) on', is semantically compatible but the fact that tak- is a checked syllable makes the connection dubious.

• 'ask (enquire)': GC tho; Yano B tao-to. Tani languages, like some other Sino-Tibetan languages, use the same verb root for both 'listen/hear' and 'ask (i.e. cause to listen)'.[17] We believe that the variant forms Padam-Mising L tau, Yano B and Tagen B tao for the meaning 'ask' may reflect the same PT root *tas. The association of the Tani and rGyarong forms, though superficially plausible, is weakened by the fact that rGyarong (lCog-rtse dialect) uses a completely separate root for the meaning 'hear/listen' ka-reŋ-na (ZMYYC).

• 'give': GC dit; Yano ǰi. The palatal initial in the Yano B form is secondary. The real PT root should be *bi (< PTB *bəy, STC #427), which is cognate rather with the regular GC word for 'to give' və (< Proto-rGyarong *bi?, cf. Dashuigou rGyarong bi-).[18]

---

[17] Cf. the Chinese parallelism: wén 聞 'hear' vs. wèn 問 'ask'.

[18] Medial -w- in lCog-rtse rGyarong often comes from earlier (phonetically prenasalized) voiced stops (cf. lCog-rtse ta-wo; Dashuigou rGyarong ta-ɱbo 'deaf n.'; cf. Jingpo na³¹phaŋ⁵⁵; Mawo Qiang bu; Queyu rɲi⁵⁵pa¹³; Muya na³³ɱbɑ²⁴; Nusu boŋ⁵⁵; WB nâ-pâng; Garo beŋ-a; Tangsa ³baŋ; < PTB *baŋ; lCog-rtse tə-wɾo; Dashuigou rGyarong tə-ŋgɾo 'sinew; tendon'; cf. Mawo Qiang gəⁱ; Xide Nosu gu⁵⁵tsɛ³³; WB aʾ-krô; Nusu gɹɯ⁵⁵; Dulong dɯ³¹gɹɯ⁵³; WT rgyus. Dashuigou 大水溝 (previously known as Benzhen 本真), like the better-known lCog-rtse and Suomo varieties, belongs to the Eastern dialect of rGyarong. The Dashuigou data cited herein were collected by the author in two recent field trips to western Sichuan.

• 'arrive': GC **Ndu**; AM **tok**. The AM form **tok** actually means 'descend'. The real Padam-Mising word for 'arrive' should be puɪŋ (< PT *puɪŋ, attested mainly in Eastern Tani languages, cf. also Bokar OY puŋ).

• 'meet': GC **rdo**; Yano **che-tok**. The 'Yano' form is actually a word from Mikir, which is not even a Tani language. The real Yano B word for 'meet' is **gue-ter-ra** (i.e. go + ? + verbal particle of reciprocality, cf. Bokar gɯ-tum-raː).

In short, eight ('hit', 'see', 'straight', 'cold', 'go', 'give', 'arrive', 'meet'), or two thirds, of the twelve proposed cognate sets above are probably misidentified, while the sets for 'dig', 'give', and 'big', although legitimate for setting up rGyarong-Tani consonantal correspondences, are of limited value for proving the proposed lexical affiliation since common TB roots are involved. Therefore, although Nagano starts with the sensible idea of probing deep lexical relations by focusing on a selected area of core vocabulary, namely basic verbs,[19] the forms randomly picked from modern Tani languages, unfortunately, failed to provide him with a reliable basis for comparison.

Nagano's alignment of rGyarong with Tani may strike those who have examined the structures of both language groups as quite surprising, for the two groups diverge from each other in almost every linguistic subcomponent. Phonologically, rGyarong has a much richer system of segmental contrasts. In contradistinction to the situation in Tani, aspiration is phonemic in rGyarong stops/affricates. Moreover, while Tani has only one (palatal) series of affricates, rGyarong distinguishes as many as four (dental, retroflexed, alveopalatal, and palatal). Although consonant clusters are not unknown in Tani (especially Western Tani), they cannot begin to compare in number and variety with the impressive array of consonant clusters found in rGyarong. The differences in morphosyntax are even more fundamental. Although both languages utilize considerable affixation, rGyarong is predominantly *prefixing* while the Tani languages are mainly *suffixing*. In terms of function, rGyarong boasts highly complex derivational as well as inflectional morphology, in contrast to Tani where morphological processes are much less abundant. Furthermore, rGyarong is an ergative language[20] with many head-marking features (Nichols 1986), including a system of verb agreement which indexes not only person and number, but also direction (or person hierarchy, i.e. direct vs. inverse) of the discourse participants. All Tani languages, on the other hand, display the so-called 'anti-ergative' pattern (LaPolla 1992), where agents are generally not

---

[19] Cf. Matisoff 1976 in which body-part terminology is chosen as the target semantic area in an exploration of shared contact vocabulary between Sino-Tibetan and Austro-Tai.

[20] Patients carry no case-marking in rGyarong. In this regard rGyarong differs from languages of the 'Qiangic' group (to which rGyarong has been assigned by some Chinese scholars).

case-marked while a single 'object' case marks a number of semantic roles, including patients, recipients, beneficiaries, and even temporals.[21] The two languages also employ distinct verb-phrase structures. In Tani, various complements and modifiers of the verb, along with such other categories as tense, aspect, polarity, and modality, are generally expressed by a large set of postposed 'verbal particles'. This characteristic is so important in Tani that it may not be too wide of the mark to say that the study of the Tani verb phrase is largely the analysis of such verb particles. No comparable phenomenon obtains in rGyarong, where many of these categories are conveyed by verbal prefixes instead. This, in short, leaves the lexicon as the only likely linguistic sub-system in which possible *close* genetic ties between rGyarong and Tani can be sought.

In order to assess the assertion that rGyarong is closely affiliated with Tani in its deepest lexical core, I have examined a total of 383 basic adjectives (stative verbs) and verbs listed in ZMYYC, yielding the following comparable pairs between rGyarong (i.e. Proto-rGyarong as proposed in Nagano 1984)[22] and Proto-Tani in these two basic semantic areas (states and actions):

---

[21] For more discussion, see J. T.-S. Sun 1994:4.2.
[22] Unfortunately, only a limited number of Proto-rGyarong roots are proposed in Nagano 1984:133-9. Where Proto-rGyarong reconstructions are unavailable, modern (lCog-rtse) forms (unasterisked), are cited from ZMYYC.

Gloss	Proto-Tani	rGyarong
'big'	*tə~*ta	*k-Te
'come'	*puɯ ('arrive')	*bo
'cover'	*kap	*p-Kap
'die'	*si	*syi
'dig'	*du	*duɯ
'dream'	*maŋ	*r-mo
'eat'	*do	*za
'exist'[23]	*duŋ	ndo
'heavy'	*ji	*li
'itch'[24]	*fak	*ya
'lean (against)'	*grəŋ	kʋ-nə-ŋgrə
'melt, thaw'	*jit	kə-ndʑi
'ripe, cooked'[25]	*min	*s-min
'run'	*duk~juk	kʋ-rʑʑək[26]
'smell'	*nam	*nam[27]
'stand'	*rop[28]	*ro[29]
'sweet'	*tiː	*ci
'thin (of people)'	(Bokar OY gi)	kə-nə-khi
'vomit'	*b(r)əs	kʋ-mʋ-mphʋt
'wait'	*jaŋ	ka-na-jo
'weep'	*krap	ka-ŋa-kru

Table 1.
Comparison of Selected Basic Verbs in Tani and rGyarong

That is, out of the 383 sets compared, only twenty-one pairs (or about 5%) show enough resemblance to be considered *probable* cognates. Furthermore, rather than revealing uniquely shared rGyarong-Tani lexical relations, the

---

[23] The PT root also means 'sit/stay/dwell'. rGyarong uses a completely different form ka-ñi for 'sit/dwell'.

[24] Nagano posits an open-syllable proto-form *ya; however, a lCog-rtse form -jak, with a checked syllable, appears in ZMYYC.

[25] This PT root means only 'ripe'.

[26] Cf. WT rgyug.

[27] Nagano 1984 provides the lCog-rtse form nam-nam. Compare the different form kʋ-nʋ nsə nsət in ZMYYC.

[28] The 'stand' meaning of PT *rop is preserved in Bokar OY. Reflexes of this root occur elsewhere mainly as an adverbial verbal particle meaning 'upwards'.

[29] It is not clear why Nagano chose to reconstruct this root as an open syllable despite the lCog-rtse form ka-rjap (ZMYYC).

majority of such pairs (e.g. 'die', 'dig', 'eat', 'heavy', 'smell', 'ripe', 'stand', 'vomit', 'weep') involve roots widely attested in the Tibeto-Burman family.

To assess further the lexical relations between rGyarong and Tani *vis-à-vis* other Tibeto-Burman members, I conducted another sample comparison including Tibetan and Burmese, two other languages showing considerable affinity to rGyarong. The items utilized for ths pilot study are narrowed down to the seventeen verbs from the Swadesh 100 core vocabulary list:[30]

GLOSS	Proto-Tani	rGyarong	Written Tibetan	Written Burmese
'drink'	*tuŋ	*mot	'thung	sok
'eat'	*do	*za	za	sá
'bite'	*g(j)am	kha mtʃik khə-lət	so brgyab	kuik
'see'	*kaŋ-paŋ	mto	mthong	mrang
'hear'[31]	*tas-paŋ	*r-na	thos; rna-ba 'ear'	krâ; na
'know'[32]	*kən	*sye	shes; mkhyen [hon.]	si'
'sleep'[33]	*jup	*r-ma	nyal; gnyid	ip
'die'	*si	*syi	si; 'chi	se
'kill'	*man	*sat	gsod	phyak; sat
'swim'	*bjaŋ	*pjaw	rkyal; 'phyo	po
'fly v.'	*bjar	*N-pjam	'phur	pyaṁ
'walk'	*in	ptʒe	'gro	hlyok; hrok
'come'[34]	*(ɦ)aŋ	*bo	yong~'ong; 'byon	la; waŋ

---

[30] The main roots are italicized; cognates with PT roots are boldfaced.

[31] The rGyarong root is cognate with WT *rna*-ba 'ear' and WB nâ 'ear', na 'listen'.

[32] The predominant rGyarong words for this gloss are cognate with WT shes and WB si' < PTB śey (STC #182); cf. lCog-rste ka-ʃə (ZMYYC), Tsanla ka-nga-syis, Khamto ka-syi, Suomo ka-nə-msyi, Chos-kia ko-syu (Nagano op. cit.:109). Nagano also gives the alternative PG root *gye-s which he links with PTB *m-kyen (and which is thus supposedly cognate with PT *ken), but it is not clear what data support this reconstruction.

[33] Nagano associates this rGyarong root with WT rmi < PTB *r-mwəy 'sleep'. The equation rGyarong -a <-> PTB *-əy, however, seems restricted to this single example.

[34] WB waŋ means 'enter'. WT 'byung 'emerge, come, go' is listed in the cognate set for PG *bo in Nagano op. cit.:84; however, if this rGyarong root came from PTB *byon (STC #179) as Nagano suggests, then the true WT cognate should rather be 'byon 'go, arrive, appear'.

GLOSS	Proto-Tani	rGyarong	Written Tibetan	Written Burmese
'sit'	*duŋ	ɲi^{35}	'dug; snye(s) 'recline,lean against' (?)	thuiŋ
'stand'	*dak; *rop	*ro	lang; 'greng	rap
'give'	*bi	dit; və	sprad; sbyin	pè
'say'	*lu; *ban	ka-rjo	bshad; smra	prô

Table 2.
Comparison of Selected Tani Verb Roots With rGyarong, Tibetan, and Burmese

Table 2 yields the following pairwise cognate numbers: Tani-rGyarong 4/17, Tani-Tibetan 8/17, Tani-Burmese 7/17; rGyarong-Tibetan 8/17-10/17;[36] and rGyarong-Burmese 8/17.[37] It is important to note that rGyarong has twice as many cognates with Tibetan and Burmese than with Tani, and that the rGyarong-Tani pair shows the *lowest* cognate count among all five pairs. To the extent that cognate counts derived from such a limited sample can be suggestive of the *relative* strength of lexical ties among the languages compared, rGyarong appears to be much more closely related in basic vocabulary to Tibetan and Burmese[38] than to Tani. This fact, coupled with the striking structural differences between the two Tibeto-Burman groups, makes their intimate genetic connection highly improbable.

2.3. Dhimalish

Dhimal (in Darjeeling and the Jalpaiguri area of Sikkim and eastern Terai, Nepal), and the closely related Ṭoṭo (south of the borderline between Bhutan and West Bengal) are two small languages comprising the obscure Dhimalish section of Shafer 1955:102. The only documentation on these languages available to us are Hodgson 1847 for Dhimal and Sanyal 1955 for Ṭoṭo. The association of this group to Tani is vaguely suggested by Benedict in STC, and we quote: "Abor-Miri and Dafla make up the nucleus of the 'North-

---

35 This rGyarong root is linked with WT snye(s) 'lean against, lie down'; again, the equation between rGyarong -i and WT -e(s) is limited to this pair.
36 The following glosses are considered to involve rGyarong-WT cognates: 'eat', 'see', 'hear/ear', 'know', 'die', 'kill', 'swim', 'come'. The cognacy of the pairs PG *r-ma, WT r-mi 'sleep', and PG *nyi 'sit', WT snye(s) 'lie down' is possible but uncertain. Thus, the number of rGyarong-WT cognates in this sample ranges from eight to ten.
37 The following items are judged to involve rGyarong-WB cognates: 'eat', 'fly v.', 'hear', 'know', 'die', 'kill', 'stand', and 'swim'.
38 The strong rGyarong-Lolo-Burmese lexical ties, suspected by Benedict (p.c.), is an area awaiting further investigation.

Assam' group of Konow and the Linguistic Survey of India. Aka (or Hruso) has
the most points of contact with this nucleus, and *Dhimal (in Sikkim) the fewest"*
(p. 6). From this statement alone it is not certain whether Benedict refers to a
contact or genetic relationship. However, on the previous page (p. 7), he does
consider Dhimal to be a likely addition to the Abor-Miri-Dafla (Mirish) nucleus.

A revisit to the Dhimalish sources, however, has failed to reveal too many
significant points of agreement between Tani and Dhimalish. The following test
comparisons, utilizing again the seventeen basic verbs from the Swadesh 100-
word list, should be suggestive of the genetic distance between the two
groups:[39]

GLOSS	Proto-Tani	Dhimal	Ṭoṭo
'drink'	*tuŋ	ám	āng
'eat'	*do	chá	cā
'bite'	*g(j)am	---	cā-pir
'see'	*kaŋ-paŋ	dó; khang	kāng; ting
'hear'	*tas-paŋ	hén	hing
'know'	*ken	gé	gē
'sleep'	*jup	jim	jing-ju; jin
'die'	*si	si	shi-pu
'kill'	*man	shé	pāi
'swim'	*bjaŋ	nó-i	---
'fly v.'[40]	*bjar	*bhír*	bi -u
'walk'	*in	hi-gil	tē
'come'	*(ñ)aŋ	lé	lé
'sit'	*duŋ	*yong*	i-ung
'stand'	*dak; *rop	jáp	lŏ-o; lo -
'give'	*bi	pí	pi
'say'	*lu; *ban	dóp	jāng

Table 3.
Comparison of Selected Basic Verbs in Tani and Dhimalish

The Dhimal and Ṭoṭo words for 'eat', 'die', 'give' and 'look' are
undoubtedly cognate with the PT roots. The cognacy of the Ṭoṭo form for
'stand', and the Dhimalish words for 'fly v.' and 'sit' (italicized in the table) to
the corresponding PT roots are uncertain. Everything considered, we get at

[39] Data transcription follows the original sources. Probable cognates with the PT roots are
boldfaced; suspicious look-alikes are boldfaced and italicized.
[40] PT *bjar reflects PTB *byer. The Dhimalish forms may come rather from PTB
*pur~pir, now considered a separate root (STC fn. 249).

most 7 cognates out of 16 pairs compared, which is equivalent to the cognate figure between Tani and Burmese obtained by using the same test sample. The set for 'look/see' (PT *kaŋ, Dhimal khang, Ṭoṭo kāng) may appear to be a striking parallel between the two groups; yet, this root occurs also in many *Kiranti* languages, e.g. Bahing koŋ 'look, watch'; Chamling, Bantawa khaŋ 'look, see', Newari khan- 'see'. On the other hand, Dhimalish seems to exhibit many more lexical links with Kuki-Chin, and especially with Tibetan, as pointed out in Shafer 1950:207.

At any rate, the similarities between Tani and Dhimalish are far from numerous,[41] otherwise they would not have escaped the attention of both Konow and Shafer. It seems, therefore, futile to search for deep connections between Tani and Dhimalish, although more extensive inquiry (and with much better Dhimalish data) needs to be done to properly assess the 'points of contact' between the two groups which prompted Benedict to place them in the same subgroup.

2.4. Hrusish

The obscure Hrusish branch is named after its best-known representative, the Hruso (paleo-exonym Aka) language of West Kameng, Arunachal Pradesh. The remarkable linguistic divergence of Hruso from neighboring Tibeto-Burman languages was already noted by Konow (1909b). Shafer 1947 compares various early wordlists of 'Aka' and concludes that actually two very distinct 'dialects' of Hruso can be established: Dialect A and Dialect B. To Dialect B, or *Hruso proper*, belong most early records of 'Aka'. Shafer's Dialect A of Aka is actually a distinct language, represented only by Campbell's (1874) variety of 'Aka'. We have recently made the discovery that Shafer's 'Dialect A of Hruso' is none other than the language of the *Dhammai* (exonym: *Miji*) tribe distributed to the north of the Hruso country. For this important language, which is more conservative than Hruso proper, we are now able to consult Simon 1979, a far more ample source than any available to Shafer. There is at least one more Hrusish language in Arunachal Pradesh, namely the language of the *Bangru* tribe of North-western Upper Subansiri district.[42] Publications on the Bangru language are completely non-existent. Our limited fieldwork data on Bangru[43] reveals such striking resemblances

---

[41] This is also the impression of Dr. Sueyoshi Toba (p.c. 1993), who has been working on this Tibeto-Burman group in Nepal.

[42] The Bangru (autonym Levai /lə³¹vɐ⁵⁵/) tribe consists of about a thousand souls whose villages are distributed in the Lagong area along the Tibetan-Indian border (Anonymous 1989:248). Note the similarity between the name *Levai* and the Miji autonym *Dhammai* (/ठ̌wṃ-mai/). It is possible that the Levai represents a northeastern subbranch of the Mijis of Eastern Kameng. The name Bangru (/bwŋ-ru/) is a Bengni exonym; cf. also the Sulung exonym of Levai: Buzwa (/bu³³zwɐ⁵³/).

[43] I recorded about a thousand Bangru words from my Sulung consultant, who has a speaking knowledge of this language, during field work in Tibet in the summer of 1992.

between Bangru and Dhammai that they may even turn out to be dialects of the same language.

The lexical similarities between the Hrusish languages and Tani (especially Western Tani) are indeed notable and deserve to be carefully investigated.

2.5. Languages of the 'Mishmi' Tribes

Comparable to Hrusish languages of the west, the languages spoken by the Mishmi tribes are the most important linguistic neighbors of Tani in the east. Unlike Tani or Hrusish, however, these languages by no means form a coherent unit. Instead, they fall into two distinct groups, Taraon-Idu (Shafer's *Digarish*) and Kaman (Shafer's *Midźuish*). Sun et al. 1980: 299-315, to date the only comparative study of the Mishmi languages based on accurate first-hand data, turns up remarkable differences. Of the 2477 native lexical items compared, 2089 or 84.4% are non-cognate, including quite a few core Tibeto-Burman items such as 'man (homo)', 'snake', 'sit', 'hand', 'hair', 'weep', 'know', 'buy', 'tooth', 'hear', 'rain', and 'house'. The morpho-syntactic disparity between the two groups is also considerable. For example, Kaman has pronominal verb agreement while Taraon and Idu do not; moreover, Kaman sometimes uses prefixes (e.g. taŋ^{55}- 'nominalizer', mai^{55}-/mɯ^{31}- 'negator', ai^{53}- 'prohibitive marker') while Taraon and Idu, like Tani, always use suffixes (e.g. Taraon -ja^{31} 'nominalizer', -jim^{55} 'negator', -ja^{53} 'prohibitive marker'). These languages, therefore, do not appear to be as intimately related to each other, contrary to what Thurgood 1985 suggests. Thus, before we even begin to compare them further with Tani (or with any other language), we must bear in mind that the alleged unity of the Mishmi languages is still an unproven hypothesis.

As stated above, most Tibeto-Burman classifications place the Mishmi languages close to the Tani nucleus. Indeed, even a cursory glance at the data shows considerable parallels between Tani and these languages (in particular Taraon and Idu), calling for more detailed exploration.

In summary, after inspecting a few alleged close relatives of Tani, we have decided to screen out rGyarong and Dhimal as improbable candidates. In the following section, the remaining languages will be further assessed by means of a more detailed lexical test.

3. TANI'S NEXT OF KIN: A FURTHER SEARCH

3.1. Methodological Perspectives

Much doubt has been cast on the validity of lexicostatistics in historical linguistic research; Matisoff 1978a:1.14 outlines the hazards of a particular application of this method, namely the use of cognate counts in setting up

subgroups among related languages.[44] However, the following statement seems quite reasonable (Thomas and Headley 1970:411, emphasis ours):

> Lexicostatistics is not a precision tool. Careful phonological reconstruction is necessary if one desires detailed information about language relationships. *Lexicostatistics is useful, however, for giving a quick general picture of language groupings.*

In fact, the authors of the preceding quote claimed that the results of their lexicostatistic analysis of Mon-Khmer internal relations can be 'presented with the confidence that the general outlines will still be standing after detailed phonological reconstruction has been done' (Thomas and Headley op. cit.). The ensuing two decades have seen considerable advances in comparative Mon-Khmer and phonological reconstruction of many Mon-Khmer subgroups (Monic, Waic, Aslian, etc.); indeed, the Thomas-Headley subgrouping framework turns out to have stood the test of time, judging by a recent authoritative statement on Mon-Khmer subclassification (Diffloth and Zide 1991).[45] Consider also the small-scale lexicostatistic study presented in Benedict 1976, where Tibetan, Burmese, Kachin, Garo, Lushai, and Pwo Karen were compared with Mandarin Chinese in terms of the Swadesh 100-word list, with the primary purpose of testing the solidarity of the Tibeto-Burman grouping vis-à-vis Chinese and Karen. It is on the basis of this analysis that Benedict proposes the '*basic cleavage line*' in Tibeto-Burman between the Baric-Jingpo supergroup and practically all other TB groups. This hypothesis has been corroborated by a follow-up comparative study of Northern Naga (i.e. Benedict's Konyak group), leading the author to conclude with confidence that the validity of the Bodo-Garo-Northern Naga-Jingpo supergroup 'should no longer be in doubt' (French 1983:727). A key factor behind these two useful (in the sense of producing new and viable ideas, inspiring further research, and contributing eventually to growing consensus) applications of lexicostatistics is that the investigators are all specialists in the respective language families, which means that the risk of cognate misidentification was minimized, and sensible adjustments in the Swadesh wordlist could be made to fit the

---

[44] The two most serious problems pointed out by Professor Matisoff being (a) How can one ensure that one's cognate identification is reliable, when detailed knowledge about the sound laws in the languages compared may be lacking? (b) How can an all-or-none (i.e. cognate vs. non-cognate) scoring method reflect the gradient nature of phonological-semantic relationships in the lexical data?

[45] Both scholars are among the world's leading Austro-Asiaticists. They have demoted Thomas and Headley's 'Malacca' (i.e. Aslian) and Nicobarese from coordinate families of Mon-Khmer to branches within Mon-Khmer, added a few minor new discoveries like Mang and Palyu (Lai), and proposed some possible higher-level divisions (Northern, Eastern, Southern, Vietic), but the basic Mon-Khmer branches remain identical to Thomas and Headley's original proposal: Viet-Muong, Khasi, Palaungic, Monic, Khmuic, Katuic, Bahnaric, Khmer, and Pearic.

particular target language families. Therefore, lexicostatistical methods, if applied with due caution and without extravagant claims,[46] may still serve as *subsidiary* tools for detecting probable subgrouping patterns.

Although the non-existence of genetic relations between languages is unverifiable in principle, it is possible to ascertain whether any two given members in a group of related languages share a *particularly close* relationship. However, this cannot be done simply by listing random similarities, because alternative explanations (borrowing, areal features, shared substratum, common retention, etc.) are not ruled out. Even if regular sound correspondences in the basic vocabulary are demonstrated, the special relation between the two languages remains unproven, for such equations can, by definition, be established between any two genetically related languages anyway.[47] What we need to do, obviously, is to single out *uniquely shared linguistic features* which set these languages apart from all others, enough to 'tip the scale against any contrary hypothesis which sets the relationship merely at the level of the underlying proto-language' (Bauman 1976:26).

However, sorting out the linguistic relations between Tani and its possible next of kin in Tibeto-Burman poses a currently insurmountable problem: the study of the Tibeto-Burman languages of Arunachal Pradesh and the immediate environs, among which the close relatives of Tani are most likely to be found, is still in its infancy, and we simply do not have the amount of linguistic information required for such detailed comparative analysis. What we can do at the present stage is no more than offer a *process of elimination*, which narrows down potentially promising avenues for further research.

3.2. A Lexicostatistic Test

A lexicostatistic study has been conducted (see the comparative table in the Appendix below) with the aim of assessing degrees of lexical affinities between Tani and four possible close relatives surviving the preliminary screening of the previous section: Taraon, Kaman,[48] Lepcha,[49] and

---

[46] Such as the controversial application of lexicostatistics to dating proto-languages (*glottochronology*).

[47] Thus, the sound correspondences between such language pairs as rGyarong-AMD (Nagano 1984), Lepcha-Adi, and Lepcha-Nung (Bodman 1988) alone do not constitute sufficient proof that these languages are more closely related.

[48] The Taraon and Kaman data are cited mostly from Sun et al. 1980 and from ZMYYC. Forms missing from these sources are supplemented from Chakravarty et al. 1963 for Taraon and Boro 1979 for Kaman.

[49] Lepcha forms are taken from Mainwaring-Grünwedel 1979. Root forms (enclosed in square brackets as in the original source) are cited where available; e.g. the root [krî], rather than the suffixed adjectival form a-krîm, is given for the gloss 'bitter'. Loanwords (chiefly from Tibetan) are marked with an asterisk in the dictionary; such forms are avoided herein except in the rare cases where the asterisked forms turn out to be the only ones listed for the given meaning.

Dhammai.[50] Written Tibetan, Written Burmese, and Garo, which have never been suspected of being *intimately* related to Tani, are added as control languages. The modest objective of this pilot study is to eliminate dubious candidates according to a simple and, we trust, reasonable principle: if a language is a true next of kin of Tani, then there should at the very least be a *significantly higher* percentage of shared core vocabulary between this language and Tani than that between Tani and languages from separate major divisions of Tibeto-Burman, in this case Written Tibetan (Bodish), Written Burmese (Lolo-Burmese), and Garo (Bodo-Garo). The test wordlist used in this study is based on the CALMSEA 200-word list[51] proposed in Matisoff 1978a: 284-96. For some CALMSEA glosses, however, no PT reconstructions are presently obtainable; either because extreme internal variation precludes positing uniform PT roots (e.g. 'descend', 'bamboo', 'sweat'), or Indic loanwords are suspected (e.g. 'needle', 'silver'), or simply because the gloss is not realized by distinct roots in most Tani languages (e.g. 'twenty'). In such cases (thirty-seven in total), CALMSEA glosses are replaced with the following items, mostly body part terms and common verbs: 'angry', 'borrow', 'call/cry', 'come', 'dead body', 'count', 'do', 'door', 'dry/wither', 'duck', 'exit', 'face', 'fireplace', 'float', 'flow', 'fly (insect)', 'gall', 'grandfather', 'grandmother', 'hungry', 'kidney', 'knee', 'language', 'melt', 'nest', 'placenta', 'rot', 'seed', 'shoulder', 'soul', 'suck', 'swallow (v.)', 'take', 'think', 'tired', 'tiger', and 'wet'. The resultant compromise list, we hope, contains few glosses that are arguably not part of the lexical core of the target

---

[50] Dhammai forms are based on Simon 1979. The sound system of Dhammai is retranscribed as follows (phonetic symbols used in the original are enclosed within parentheses):

1. Vowels: a, e, ɯ (i), i, o, u
2. Consonants:

p	t	tɜ	č (c)		k	ʔ
ph	th	tsh	čh (ch)	kh		
b	d	dz	ǰ (j)		g	
f	θ	ɜ	š (sh)	h		
v	δ	z	ž (zh)			
m	n		ñ		ŋ (ng)	
	l					
	ɫ					
	r					
w		ɉ (y)				

Remarks: (1) Dhammai may have contrastive vowel length and phonemic tone; neither gets marked in the main body of this source. (2) The glottal stop is a phonemic syllable coda, represented in the source by -h. (3) Dhammai has a peculiar lateral consonant symbolized by Simon as ll, which he describes as being 'articulated with the tongue rolled'. This is probably the retroflexed lateral ɭ.

[51] Abbreviated from Culturally Appropriate Lexicostatistical Model for South-East Asia, this list represents Prof. Matisoff's revision of the Swadesh basic vocabulary list to make it culturally more appropriate for Southeast Asian languages.

languages. Our cognacy judgement[52] with respect to WT, WB, and Lepcha should be relatively uncontroversial, for much is known about the historical phonology of these languages, and expert guidance is readily available from STC and various other works on Sino-Tibetan reconstruction. The same can be said of Garo, the best known of all Baric languages, not only because it was one of the principal languages on which the PTB reconstructions in STC were based, but also thanks to a series of treatises on Baric contributed by Robbins Burling (especially 1959, 1983, and 1992).[53] Cognate detection involving other target languages is much more difficult. In the case of Taraon and Kaman, although we are lucky to have access to mutually complementary Indian and Chinese sources (the accuracy of the latter is quite impeccable), the phonological developments of these languages, especially the less conservative Taraon language, are not yet well-known.[54] Dhammai is even more troublesome in terms of data reliability and cognate identification. Furthermore, thirty-three test items are missing from the word list in Simon 1979 (the only available substantial source on this important language), although it is not clear to what extent the incomplete data may cause the *averaged* cognate percentage to be skewed.

3.3. Results and Discussion

Each of the languages compared contains a number of forms of indeterminate cognacy with the corresponding PT roots. Such is the case, for instance, between PT *kɯ 'dove/pigeon' and WT 'aŋ-gu 'pigeon'.[55] A more conservative estimate may discount these doubtful cases, a bolder count would

---

[52] Cognate identification in Tibeto-Burman is an extremely risky undertaking. Our general attitude is to be more willing to *err on the conservative side*, for our knowledge of the various languages involved (except perhaps Tibetan) is not sufficient to allow bold speculation. In this study, forms are treated as cognate only if they are considered to descend from one and the same *proto-allofam* (i.e. variants of the same proto-word-family, Matisoff 1978a:17). Thus, WB klok~kyok and PT *luŋ 'stone' are not directly cognate even though they may come from related proto-allofams. By the same token, Taraon pia⁵³kɹau³⁵ and Kaman tɕi⁵⁵khɹuŋ⁵⁵ (< PTB *(m-)krəw 'dove', STC #118) are not cognate with PT *kɯ 'dove/pigeon' (< PTB *(m-)kəw 'pigeon' STC #495; note that PT normally kept the PTB *kr- cluster), for they are derived from related but distinct PTB etyma. Of course, such subtle distinctions are not always possible with languages the sound laws of which are not yet well-known.

[53] The Garo data are taken mainly from Burling 1983. Supplementary forms, marked by #-, are added from Momin: no date. Transcription of Garo is based on the 'combining' (i.e. non-final) form, which is etymologically more basic (Burling 1983:69-70). Garo-Tani cognate determination is greatly facilitated by the etymological tables in Burling 1983, where the PTB etyma of many Garo roots are provided.

[54] Initial efforts have been made to inspect the sound laws of Taraon, but a full-scale comparative study of Taraon and its close kin Idu has not been attempted.

[55] WT 'aŋ-gu is more common in Central Tibetan. In Khams Tibetan, ɳug-gu is used instead. The normal Classical Tibetan word is phug-ron. While PT *kɯ is clearly a reflex of PTB (*m-)kəw 'pigeon' (STC #495), WT 'aŋ-gu shows an unexpected voiced initial g- (although WT -u regularly reflects PTB *-əw).

include them all, while the cognate figure closest to reality may lie somewhere in between. These two different figures, then, represent the *range* of possible cognation between the given language and PT. Since, for example, WT shows two doubtful cognates (the other being PT *be, WT spre 'monkey') and fifty-six sound ones, the cognate ratio between PT and WT ranges from 56/200 (or 28%, conservative estimate) to 58/200 (or 29%, less conservative estimate). The much larger percentage of such uncertainty for Taraon is a function of the phonological deviancy of the language. The output of this study can be summarized in the following table:

	WT	Garo	WB	Taraon	Kaman	Dham-mai	Lepcha
available forms	200	194	200	200	200	167	200
cognate count	56-58	46-50	54-57	59-76	43-50	43-49	47-49
percentages	28-29	24-26	27-28.5	29.5-38	21.5-25	26-29	23.5-24.5
average percentage	28.5	25	28	33.75	23.3	27.5	24

Table 4.
Cognate Figures Between Tani and Seven Tibeto-Burman Languages

The output obtained from this pilot study has a number of noteworthy implications for the phylogenetic position of Tani.

First, this lexicostatistic test has indeed accomplished its unpretentious mission of *separating off problematic candidates* from among the possible close relatives of Tani. The cognate figures of PT with both Lepcha and Kaman are *lower* than those between PT and the three control languages. In particular, the PT-Kaman cognate percentage is the lowest of all figures obtained. If core vocabulary is reliable at all as an index of relative genetic distance, then these facts should constitute strong disproof of any intimate relation between either of these languages and Tani. As for the lexical similarities between Lepcha and Tani observed by Bodman 1988, alternative explanations must be sought, such as shared substratum,[56] or early contact (in southern Tibet?) of the two language groups before their migration to the present locations. In short, our findings support Bodman's conclusion that although Adi may be among the TB

---

[56] Consider for example PT *luk, Lepcha lyăk, cf. PTB *lay 'exchange' (STC #283). The PT and Lepcha forms may be related rather to Mon-Khmer, cf. Proto-Wa-Lawa *ʔlɔh (Diffloth 1980), Kammu (Yuan dialect) lɛɛk 'exchange' (Lindell 1974:200). The PT and Lepcha words for 'excrement' may also be of Mon-Khmer origin (Forrest 1962). The considerable Mon-Khmer contact vocabulary in Tani languages will be explored in a separate paper.

languages which are more similar in lexicon to Lepcha,[57] the relationship between them is not very close (Bodman op. cit.:4).

Compared with Lepcha and Kaman, Dhammai shares a higher cognate percentage with PT, yet, this figure is still lower than that between PT and WT. Although we are not well-informed enough about the linguistic structures of the Hrusish languages to say anything definite about the relation between Hrusish and Tani, we do suspect that the similarities between them[58] may be the consequence of prolonged contact rather than exclusively shared linguistic history, and that the true roots of Hrusish may lie somewhere else in Tibeto-Burman.

Cognate percentages between PT and the three control languages run between 24 and 29. The close clustering of these figures indicates that Tani indeed forms a distinct division in Tibeto-Burman, coordinate with other major nuclei in the family. The lower Tani-Garo figure suggests that Tani is more akin to WB (Lolo-Burmese) and WT (Bodic) than to Garo (Baric), corroborating Benedict's inclusion of Miri on the non-Baric side of the 'basic cleavage line' in Tibeto-Burman. This also shows that subgrouping Tani under Baric (e.g. DeLancey 1991a) may not be advisable. Furthermore, Tani shares almost as many cognates with WB as with WT, a finding which is all the more remarkable since Lolo-Burmese and Tani (or for that matter any Arunachal Tibeto-Burman groups except perhaps Singpo) have never been known to be in close areal contact. This calls into question Egerod's decision to classify Tani directly under Tibetic (Egerod 1974).

The language that stands out with the highest cognate figure with Tani is Taraon (29.5%-37.5%, average **33.75%**). This figure, interestingly, is higher even than that of the Taraon-Kaman pair (30%-33%, average 31.5%).[59] The large gap between the more conservative (29.5%) vs. the bolder cognate estimate (37.5%), nevertheless, reflects our current inability to distinguish between true cognates, allofams, and chance look-alikes. However, as stated, we have made an attempt to uncover the elusive sound laws of this language,

---

57 Unfortunately, the Kuki-Chin-Naga and Kiranti-Tibetan-Kanauri links are not considered in Bodman 1988. Lepcha certainly seems to have as many good lexical comparisons with Mikir and Ao Naga as with Tani, on Bauman 1976's evidence.

58 There are two major subgroups within Tani: Western and Eastern (Sun 1993: chapter III). As may be expected, more parallels exist between Hrusish and *Western* Tani. For example, the Western Tani root *nam 'house' (as against Eastern Tani *kjum) is obviously related to Hrusish, cf. Dhammai nen, Bangru ne:⁵⁵, Hruso ñe 'house'.

59 The Taraon and Kaman forms for the following items are judged to be cognate: 'bear n.', 'bird', 'blood', 'brain' (?), 'borrow', 'burn' (?), 'child/son', 'cloud', 'day', 'die', 'dog', 'dove' (?), 'dream', 'eat', 'eight', 'extinguished', 'fat/stout', 'fat n.', 'excrement', 'fire', 'fireplace', 'fish', 'float' (?), 'flower' (?), 'four', 'full', 'gall', 'guts', 'head', 'horse', 'kidney', 'kill', 'knife', 'leech', 'lick', 'listen/hear', 'melt', 'moon', 'mortar', 'name', 'neck', 'otter', 'penis' (?), 'pig', 'poison', 'ripe', 'river', 'road', 'round', 'seed', 'sharp-edged', 'smoke n.', 'stone', 'tail', 'thick', 'thin', 'thou', 'three', 'tiger', 'tongue', 'village', 'vomit', 'water', 'weave', 'wet', 'wing', and 'wood'.

and our cognacy judgments, we contend, are at worst educated guesses rather than wild speculations.

4. More Thoughts on the Tani-Digarish Relationship

A major outcome of this study is that Digarish (Taraon and Idu) may be the Tibeto-Burman group most similar in lexicon to Tani. However, before jumping to the conclusion that Digarish and Tani are collateral relatives in Tibeto-Burman, we should be reminded that the fundamental research necessary to prove such an intimate connection has not been done, and alternative accounts of such lexical parallels cannot yet be ruled out. Since to adequately pursue this line of research would involve at least another dissertation-length study, we will have to content ourselves with suggesting a few interesting Tani-Taraon parallels in other linguistic subcomponents.

With regard to shared *peculiar* phonological innovations, the development of PTB $*dz-$ to PT $*d-$ is paralleled by Taraon th-; e.g. PTB $*dza$, PT $*do$, Taraon tha^{53} 'eat'. Elsewhere in Tibeto-Burman, PTB $*dz-$ usually either survives as affricates (e.g. Mawo Qiang $dzə$; WB $câ$ 'eat') or spirantized (e.g. WT za; Jingpo $\int a^{55}$ 'eat').[60] Another possible example of common phonological aberrancy is the irregular *palatalized* initial in the following roots: PT $*rjam$, Taraon $liu\eta^{53}-gie^{31}$ < PTB $*la(:)m$ 'fathom'; PT $*rjum$ 'dusk/evening', Taraon $liu\eta^{53}$ 'night', < PTB $*rum \sim *rim$ 'dusk' (STC #401); PT $*\tilde{n}a-$; Taraon $xa^{31}nia^{53}puun^{55}$ < PTB $*s-na$ 'nose' (STC #101).

The remarkable lexical affinities between Taraon and Tani are not restricted to content words. Some *grammatical* morphemes may also be cognate:

- 'comparative auxiliary' PT $*ja\eta$; Taraon $jo\eta^{53}$ [61]

---

[60] The development to stops is not uniquely shared by Tani and Digarish, however. Matisoff 1978b:11 reports, for instance, that PTB $*ts-$ and $*dz-$ went respectively to th- and t- in Mpi, a southern Loloish language of Thailand. Cf. also the Queyu (Qiangic) word for 'eat' $kə^{35}tə^{53}$ (ZMYYC).

[61] For usage, consider the illustrative sentences below:

Bokar OY (Ouyang 1985: 71)

ši: laato a:to-joŋ-da
this road far-more-declarative
'This road is farther.'

Taraon (Sun et al. 1980:219)

$tçe^{55}$ $xa\eta^{35}-do\eta^{31}go^{31}$ $lau^{55}dzo\eta^{55}$ $pia^{55}-jo\eta^{35}$
s/he I-than learn good-more
'S/he learns better than I do.'

- 'imperative suffix' PT *to; Taraon t i o[53]
- 'prohibitive suffix' PT *jo; Taraon ja[53] [62]
- 'experiential aspect marker' PT *kɯ; Taraon koŋ[35]

The morphosyntactic structures of the two groups have not been carefully explored, but some prima facie resemblances exist here as well. In both groups, pronominal verb agreement is lacking. Further, the predominant verbal morphology in both cases is suffixal. Digarish languages, like languages of the Tani group, also seem to exemplify the 'anti-ergative' case-marking type, where patient and recipient nominals receive *identical* marking while agents are seldom case-marked.

On the other hand, the differences between the two groups seem to overshadow their similarities. Apart from their overall lexical differences, many of the characteristic Tani lexical items and phonological developments (such as PTB *-a > PT *-o, and the shift of all PTB diphthongs into PT monophthongs) find no counterparts in Digarish. The overwhelming majority of grammatical morphemes in Tani and Digarish are also unrelated. From the few available syntactic descriptions, the two groups also show important disparities in morphosyntax. For instance, Digarish languages use separate existential verbs depending on the *animacy* of the subject, a distinction unattested in any known Tani languages. As stated, although some Tani languages do contrast different existential verbs, the relevant distinctions are usually *polarity* (e.g. Bengni S do: 'exist/have'; ka:-ma: 'not exist/have') or even *posture* (Apatani A da 'exist (referent standing)'; du 'exist (referent sitting)'; do 'exist (referent lying)') of the predicated nominal (Abraham 1985:70-3). Moreover, relative clauses in Taraon are formed simply by gapping, without first nominalizing the embedded clause as is usually the case in the Tani languages.[63]

In summary, even though Digarish and Tani bear some striking resemblances, their equally impressive differences make it doubtful that this relationship could be an intimate one, even if future studies could establish an exclusively shared genetic relationship between them.

---

[62] There is an interesting look-alike in Tai: Proto-Tai *ʔjaa[A1] 'prohibitive; negative imperative' (F.K. Li 1977:181). [Ed.]

[63] Observe the example below, taken from Sastry 1984:189 (tone marks omitted):

hã [hɨbaŋ bo-ya jyinaŋ]REL Ø-dõ kitab haŋ-de
I forest go-impf cousin   Ø-obj book give-impf
'I give the book to (my) cousin who goes to the forest.'

REFERENCES

Anonymous. 1989. *Luobazu shehui lishi diaocha* ('Investigations on the society and history of the Luoba nationality"'. Vol. II. Lhasa: Tibetan People's Press.

Anonymous. 1991. *Zangmianyu Yuyin he Cihui* ('Tibeto–Burman phonology and lexicon"'. Beijing: Chinese Academy of Social Sciences Press. ("ZMYYC")

Bauman, James. 1976. "An issue in the subgrouping of the Tibeto-Burman languages: Lepcha and Mikir." Paper presented at the 9th International Conference on Sino-Tibetan Languages and Linguistics. Copenhagen, Oct. 22-24, 1976.

Benedict, Paul K. 1972. *Sino–Tibetan: a Conspectus.* Contributing editor: James A. Matisoff. Cambridge: Cambridge University Press.

———. 1976. "Sino–Tibetan: another look." *Journal of the American Oriental Society* 96.2:167-97.

Bodman, Nicolas C. 1988. "On the place of Lepcha in Tibeto–Burman: a lexical comparison." *Linguistics of the Tibeto–Burman Area* 11.1:1-26.

Bor, N. L. 1938. "Yano Dafla grammar and vocabulary". *Journal of the Royal Asiatic Society of Bengal* 4:217-81.

Boro, A. 1978. *Miju Dictionary.* Shillong: Research Department, Arunachal Pradesh Administration.

Bradley, David. 1993. "The subgrouping of Proto-Tibeto-Burman." Paper presented at the 26th ICSTLL. Osaka, Japan, September 13-17.

Burling, Robbins. 1959. "Proto-Bodo." *Language* 35.3:433-53.

———. 1971. "The historical place of Jinghpaw in Tibeto-Burman." Papers on Tibeto-Burman historical and comparative linguistics from the Second Annual Meeting on Sino-Tibetan Reconstruction. (=OPWSTBL Vol II), ed. by F. K. Lehman, 1-54. Urbana: University of Illinois.

———. 1983. "The Sal languages." *Linguistics of the Tibeto-Burman Area* 7.2: 1-31.

———. 1992. "Garo as a minimal tone language." *Linguistics of the Tibeto-Burman Area* 15.1:1-12.

Campbell, G. (ed.) 1874. *Specimens of languages of India.* Calcutta: Bengal Secretariat Press.

Chakravarty, L. N., et al. 1963. *A Dictionary of the Taraon language.* Shillong: Philology Section, Research Department, North-East Frontier Agency.

Chhangte, Thangi. 1990. "Nishi grammar sampler." Paper presented at the 23rd ICSTLL. University of Texas, Arlington, October 5-7.

———. 1992. "Phonology of some Nishi (Dafla) dialects." Paper presented at the 25th ICSTLL. UC Berkeley, California, October 14-18.

Dai, Qingxia. 1989. "Guanyu woguo Zangmian yuzu xishu fenlei wenti" ('On the issue of genetically classifying the Tibeto-Burman languages of China'). *Yunnan Minzu Xuebao* 3:82-92.
DeLancey, Scott C. 1989. "New vs. assimilated knowledge as a semantic and grammatical category." Paper presented at the Winter LSA Meeting. Washington, D. C., December 30th, 1989.
_____ . 1991. "Sino-Tibetan languages." *Oxford Encyclopedia of Linguistics*, Vol. 4, ed. by William Bright, 445-9. New York: Oxford University Press.
Diffloth, Gérard. 1980. "Proto-Waic." *Linguistics of the Tibeto–Burman Area* 5.2:1-182.
Diffloth, Gérard and Norman Zide. 1991. "Austro-Asiatic languages. " *International Encyclopedia of Lnguistics*, ed. by William Bright, 137-42. New York: Oxford University Press.
Egerod, Søren. 1974. "Sino–Tibetan languages." *Encyclopedia Britannica* 16:796-806.
Forrest, R. A. D. 1962. "The linguistic position of Róng (Lepcha)." *Journal of the American Oriental Society* 82:331-5.
French, Walter T. 1983. *Northern Naga: a Tibeto-Burman mesolanguage*. New York: City University of New York dissertation.
Hale, Austin, ed. 1982. *Research on Tibeto-Burman Languages*. Berlin, New York, Amsterdam: Mouton.
Hodgson, B. H. 1847. *Essay the first on the Kocch, Bodo, and Dhimal tribes in three parts*. Calcutta: J. Thomas Baptist Mission Press.
Huang, Bufan. 1991. "Qiangyu zhi" (The Qiangic branch'). *Hanzangyu gailun* ('Survey of Sino-Tibetan'), ed. by Ma Xueliang et al., 208-369. Beijing: Beijing University Press.
Konow, Sten. 1909. "North Assam group." *Linguistic Survey of India*, ed. by G. A. Grierson, Vol 3, pt. 1: 568-72. Calcutta: Superintendent of Government Printing.
LaPolla Randy J. 1992. "Anti–ergative marking in Tibeto–Burman." *Linguistics of the Tibeto–Burman Area* 15.1-1-9.
Li, Fang Kuei. 1977. *A Handbook of Comparative Tai*. Honolulu: University of Hawaii Press.
Lindell, Kristina. 1974. "A vocabulary of the Yuan dialect of the Kammu language." *Acta Orientalia* 36:191–207.
Lorrain, J. H. 1907. *A Dictionary of the Abor-Miri Language*. Shillong: Eastern Bengal and Assam Secretariat Printing Office.
Mainwaring, G. B. (Revised and completed by Albert Grünwedel). 1979. *Dictionary of the Lepcha language* (Revised and completed edition). Kathmandu: Ratna Pustak Bhandar.
Marrison, Geoffrey E. 1988. "The Adi-Dafla group of languages of North–East India: a sketch." *Prosodic analysis and Asian linguistics, to honour R. K.*

Sprigg, ed. by David Bradley, Eugénie J. A. Henderson and Martine Mazaudon, 20-22. Canberra: Australian National University.

Matisoff, James A. 1972. *The Loloish Tonal Split Revisited*. (Center for South and Southeast Asia Studies Research Monograph No. 7). Berkeley: Center for South and Southeast Asia Studies, UC Berkeley.

_____. 1974. "The tones of Jinghpaw and Lolo-Burmese: common origin vs. independent development." *Acta Linguistica Hafniensia* 15.2:153-212.

_____. 1978a. *Variational Semantics in Tibeto-Burman: The 'organic' approach to linguistic comparison*. OPWSTBL Vol. VI, ed. by F. K. Lehman. Philadelphia: Institute for the Study of Human Issues.

_____. 1978b. "Mpi and Lolo-Burmese microlinguistics." Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

_____. 1985a. "Out on a limb: *arm, hand,* and *wing* in Sino-Tibetan." *Linguistics of the Sino-Tibetan Area: the State of the Art*; papers presented to Paul K, Benedict on his 71th birthday, ed. by Graham Thurgood, James A Matisoff, and David Bradley, 421-50. Canberra: Australian National University.

_____. 1985b. "God and the Sino-Tibetan copula, with some good news concerning selected Tibeto-Burman rhymes." *Journal of Asian and African Studies* (Tokyo Foreign Language University) 29:1-81.

_____. 1991. "Sino-Tibetan linguistics: Present state and future prospects." *Annual Review of Anthropology* 20:469-504.

Momin, K. W. No date. *English-Achikku Dictionary*. No place of publication.

Nagano, Yasuhiko. 1984. *A Historical Study of the rGyarong Verb System*. Tokyo: Seishido.

Nichols, Johanna. 1986. "Head-marking and dependent-marking grammar." *Language* 62:56-119.

Nishida, Tatsuo. 1979. "Robago no keitō" ('The system of the Luoba language'). *Gengo* 7: 70-77.

_____. 1984. "Review of Hale 1982." *Tōkyō Gakuhō* 65.3-4.

Ouyang, Jueya. 1985. *Luobazu yuyan jianzhi: Bengni-Boga'eryu* ('Brief description of a language of the Luoba nationality: the Bengni-Bokar language'). Beijing: Minzu Chubanshe.

Qu Ai-tang. 1984. "Jiarongyu gaikuang" ('Brief description of rGyarong'). *Minzu Yuwen* 2: 67-80.

Sanyal, Chandra Charu. 1955. "The Ṭoṭos, a Sub-Himalayan tribe." *Journal of the Asiatic Society-Science* 21:97-126.

Sastry, G. Devi Prasada. 1984. *(Taraon) Mishmi grammar*. Mysore: Central Institute of Indian Languages.

Shafer, Robert. 1947. "Hruso." *Bulletin of the School of Oriental and African Studies* 12:184-96.

_____. 1950. "Classification of some languages of the Himalayas." *Journal of the Bihar Research Society* 34:192–214.

_____. 1955. "Classification of the Sino–Tibetan languages." *Word* 11:94-111.

Simon, I. M. 1979. *Miji Language Guide.* Shillong: Philological Section, Directorate of Research, Government of Arunachal Pradesh.

Sun, Hongkai. 1982. "Qiangyu zhishu wenti chutan" ('Preliminary explorations on the linguistic classification of Qiangic'). *Minzu yuwen yanjiu wenji* ('Collection of papers on minority language studies'),189-224. Xining: Qinghai Minzu Chubanshe.

_____. 1983. "Liujiang liuyu de minzu yuyan ji qi xishu fenlei" ('On the minority languages of the Six-River Region and their genetic classification'). *Minzu Xuebao* ('Yunnan Minzu Chubanshe') 3: 99-273.

_____. 1988. "Shilun zhongguo jingnei Zangmianyu de puxi fenlei" ('On the genetic classification of the Tibeto-Burman languages of China'). *Languages and History in East Asia: Festschrift for Tatsuo Nishida on the occasion of his 60th birthday,* edited by Akihiro Sato, 61-73.

Sun, Hongkai et al. 1980. *Menba, Luoba, Dengren de yuyan* ('Languages of the Menba, Luoba, and Deng peoples'). Beijing: Zhongguo Shehui Kexue Chubanshe.

Sun, Jackson T.–S. 1993. *A Historical–Comparative Study of the Tani (Mirish) Branch in Tibeto–Burman.* Berkeley: University of California dissertation.

_____. 1994. "Linguistic characteristics of the Tani (Mirish) branch in Tibeto–Burman." *Bulletin of the Institute of History and Philology* 65.1: 175-220.

Thomas, David and Robert K. Headley, Jr. 1970. "More on Mon-Khmer subgroupings." *Lingua* 25:398-418.

Thurgood, Graham. 1985. "The 'Rung' languages: notes on their proto-morphosyntax and subgrouping." *Acta Orientalia* 46: 79-99.

_____. 1986. "A brief comment on Das Gupta's works." *Linguistics of the Tibeto-Burman Area* 9.2:91-6.

Weidert, Alfons. 1987. *Tibeto-Burman Tonology.* Amsterdam/Philadelphia: John Benjamins.

Wolfenden, Stuart N. 1936. "Notes on the Jyârong dialect of eastern Tibet." *T'oung-Pao ser.* 2.32:167-204.

APPENDIX63
Comparison of 200-Word Core-Vocabulary in Eight Tibeto-Burman Languages

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
alive	*tur	gson-pa	taŋ-	hraŋ	a³¹suŋ⁵⁵	kɯ³¹ʐaŋ³⁵	ʃun	zu
angry64	*haŋ-fak	ʼkhro; ʼtshig; sdang	ka-o-naŋ	cit-chûi; mjak	khɯm⁵⁵ mioŋ⁵⁵	sun⁵⁵ dut⁵⁵ lat⁵⁵	nen	a-mlem nók non; li; sak lyak
ant	*ruk ~*rup	grog-ma	---	pu-rʋak	kɯ³¹jɯ⁵³	tɕu³¹ krik⁵³	---	tŭk-fyil
arrow	*puk	mda'	*bra	hmrá	pɯ⁵⁵	a³¹vat⁵⁵	nu	tsóñ
ascend	*ĉaŋ	ʼdzeg	ga-kat	tak	tu³¹dzi³⁵ noŋ⁵⁵	lun⁵⁵ xai⁵⁵	khun?	hróñ
awake (v.i.)	*ɦut²	gnyid sad	*mik-rak	nûi	dzɯ⁵⁵a⁵⁵	kʐaŋ⁵⁵	phrɯ-u	ʃí
banana	*ko-pak	skyes-sdong	te-rik	ŋgak	pha³¹ dzi⁵⁵	xa³¹ biuŋ⁵⁵	ru-ɓaŋ; ru-ɬaŋ	-blo
bear (n.)65	*tum	dom	map-il	ʋak ʋaɰ̃	ta³¹ma⁵⁵	kum⁵⁵	ʃu-tsaŋ	sɜ-na
belly	*kri	grod khog	ok	puik	kɯ³¹juŋ⁵⁵	dăk⁵³	rug	(tɜ̆-)băk
bird	*taŋ	bya	do?	ŋgak	pia⁵⁵	wa³⁵	bu-zu(?)	fo
bite	*gam~ *gjam	so brgyab	cik	kuik	tie⁵⁵	săk⁵⁵	tha?;ʃu-wrai?	tsuk;ran
bitter66	*kai-	kha	ka	kha'	khlai⁵⁵	kha⁵³	mu-khu?	kri

63 Probable cognates are bolded; uncertain cognates are both boldfaced and italicized, to be taken account of separately in the cognacy calculation.

64 Many of the 'angry' forms here are compounds with a first element meaning 'mind': e.g. PT *haŋ-, WB cit-, and Lepcha sak,- (which looks deceptively like the main PT 'angry' root *fak).

65 For Taraon ta³¹ma⁵⁵ cf. the more transparent form ta:-hom in Chakravarty 1963.

66 The Dhammai form mu-khu? exemplifies a regular sound change PTB *-a > Dhammai -u, cf. also bu-ŋu 'five'; tsu? 'eat'; lu 'month/moon'; zu 'son'; thu 'tooth'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammal	Lepcha
blood	*vī:	khrag	a?n-ci	svê	xa³¹ rual⁵³	a³¹ rui³⁵	žai	vî (nyo)
blow	*mut	ʼbud	spo-	hmut	muŋ⁵³	thut⁵⁵	---	mắt~ mũt
bone	*loŋ	rus-pa; gdung	greŋ	a'rui	rɯ³¹ boŋ³⁵	çiŋ⁵⁵ rak⁵⁵	(mɯ-) ljaŋ	a-hrǎt
borrow 67	*nar	g.yar; skyi; brnyan	ra?-cak	hŋâ; khyê	xa³¹ ŋa⁵⁵	a³¹ ŋat⁵⁵, lu⁵³	---	*nyó-lyã
bow (n.)	*rji	gzhu	*cri	lê	a³¹ lai⁵³	gaŋ³⁵	gu-ri?	sǎ-lí
brain	*pʸk-ni	klad-pa	ta-niŋ	û-hnok	pu³¹ rɯm⁵⁵	nun⁵³	---	a-tʸyak yón; a-yǎñ
branch	*hak	yal-ga; gel-pa	*cek-si	a'-khak	xa³¹ ra⁵³	ŋkhai³⁵	ou dɯ- tsaŋ	a-kón; a- nǔñ
breath	*sak, ŋa	rngam	raŋ-sit	(e-) sak	çu⁵⁵	ntshon³⁵	dɯ-thu	sóm
burn (v.i.)	*gu	ʼbar	kam	tok	xraw⁵³	gri³⁵, xw³¹ nai³⁵	phrjaŋ; rau?	mi dyak
buy	*ro	byo	bre	way	brai³⁵	çip⁵⁵	phun?	par
call/cry	*grok	grags~ grogs	o-kam; crik	hac; khaw	xa³¹ tiɯŋ⁵⁵ a³¹. #grã:	bɯɯ⁵³	then	lik
child/ son	*ño	bu	bi?-sa	sǎ	a⁵⁵ (ju⁵⁵ a⁵⁵)	sa⁵⁵ wai⁵³	zu	a-kúp
cloud	*muk ~ *mek	sprin-pa	a-ram	tim	a³¹m⁵⁵	ka⁵⁵ mãi³⁵	mei-miv	-byoñ
come	*vaŋ	ʼong	re?-ba?	la	xa³¹ naŋ⁵⁵	xu⁵³	dai	di; lat; tʼi

67 Sino-Tibetan languages generally do not lexicalize directionality of the loaning transaction, thus 'borrow' and 'lend' are usually expressed by identical roots. Instead, many Tibeto-Burman languages make a different distinction based on the nature of the loaned object: thus 'borrow/lend something that must itself be returned' and 'borrow/lend something that can be returned in kind' involve distinct roots, e.g. Tibetan g.yar vs. skyi: Burmese hŋâ vs. khyê; Kaman a³¹ŋat⁵⁵ vs. lu⁵³ in the table. This contrast has not been detected in any Tani language.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
count	*kru	brong	*chan	rañ'	ta³¹ tsɛi⁵⁵	xɐ³¹ tsɯt⁵⁵	---	fróń
day[68]	*lo	nyi-ma	sal	rak	kɯ³¹ŋ⁵³	ŋin⁵³	wu	nyí
dead body	*si-maŋ	ro	maŋ gi-si	ə-lóŋ	thuŋ⁵⁵	dzal³⁵	---	(a-) fúń
die	*si	'chi	si	se	ɕi⁵⁵	ɕi⁵³	ći	mɛk
dig	*dul¹; *ko²	rko²; 'bru	co?l	tɯ̂l	ua⁵⁵	gua³⁵; son³⁵	thau?l	dul·; byol
do	*rjɯl¹; ŋo²	byed; spyod	dak	lup; ɯu²	ba⁵³	pɐm³⁵	ru¹	mɐt; zuk; fɐt
dog[69]	*ki:	khyi	a-chak	khwḛ	kuau⁵³	kui⁵⁵	ɕa-ži?	ka-ju (pɐ-li)
door	*rjɐp	sgo	do-ga	tɐm-khá	kɐ³¹ lɯŋ³⁵	mphun⁵³	ban-phi?	(tǔn-) vyeń
dove	*ku	'aŋ-gu	do-kru	khui	pia⁵³ kɹau³⁵	tɕi⁵⁵ khɹuŋ⁵⁵	bjuŋ-lo	fɐ-vu-fo
dream	*jup-maŋ	rmi-lam; rmaŋ	ju-maŋ	ip-mak	ja⁵⁵ mo⁵³	kɐ³¹ maŋ³⁵	---	móń
drink	*tuŋ	'thuŋ	riŋ	thok	tim³⁵	taŋ⁵⁵	thuŋ	t'ǎn- t'óń; báp
dry/wither	*san	skam-po	ra?n	khrok	gog³⁵	sal⁵³	mɯ-khjaŋ	a-ɕin; a-són; a-jep
duck	*jap	ngur-ba	do-gep	bhai	ma³¹tɕi⁵³ pia⁵³	kɹai³⁵ pit⁵⁵	ŋu-so	*dɐm-byó
ear[70]	*ɓa-ruŋ	rna	na-cir	ná-rwak	kɹu⁵³ naŋ³⁵	iŋ⁵⁵	žo?	a-nyor

68 For the ZMYYC Kaman form ŋin⁵³ cf. Boro 1979 ŋit; Weidert 1987:478 ŋit 'day'.
69 Note the secondary -k coda in the Taraon form kuau⁵³ (for -u < -k. cf. Chakravarty 1963 kuɐk; Sastry 1984 kwég).
70 The Taraon word for 'ear' is literally kɹu⁵³ 'head' + naŋ³⁵ 'leaf'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
eat	*do	za	ca?	cá	tha⁵³	tɕa⁵³; ɕa⁵³	tsu?	zo; wam-mat; tˀa
egg	*pu	sgo-nga	bit-ci; do?-ci	u'	na³¹ na⁵³	krai⁵⁵ sit⁵⁵	du-rin?	a-tí
eight	*pri-ñi	brgyad	cet	hrac	lium³⁵	i⁵⁵ lioŋ³⁵	su-gi?	ka-kŭ
excrement	*e:	skyag-pa	ki	khyê	klai⁵³	tu³¹ khwi⁵³	---	'ayít; ít; e
exit	*len	thon; 'byung	---	thwak	leŋ³⁵ bi³¹	xa⁵⁵ tha³⁵	---	zán
extinguished71	*mit	shi	*ki-mit	se	xa³¹ nun⁵³	mǎn⁵³; *mut	---	mí mak
eye	*mik	mig	mik-ron	myak-se'	bu³¹lum⁵⁵	min⁵⁵; *mik	mi?	a-mik
face	*mik-mo:	gdong; ngo;bzhin	mik-kaŋ	myak-hna	ŋaŋ⁵⁵	a³¹gul³⁵	gu-mja?	a-mlem
fall (from a height)	*ho	lthung	gak-on	kya'	blai⁵⁵ dau⁵⁵; *ga-lja:	mit⁵⁵ti³⁵ seu⁵⁵; *bral	du-jun	hlat; glo; klo
far	*do	rgyang-ring-po	ce?l	vê	dia⁵⁵	klam⁵⁵	mu-run?	[rŭ]
fat/ stout	*ʃuŋ	rgyags-pa; tsho-ba	mil	va'	diuŋ⁵³	kw³¹diaŋ⁵⁵	za?-mu-do	[šu]; a-tˀor ~ a-tˀyor; [ñur]
fat (n.)	*fu	snum-pa	mit-dim	chi	ta³¹so⁵³	ta³¹si⁵⁵	thai-bau	a-šut < [šu]
fear	*bV-so: ~*pV-so:	'jigs; zhed; dngang	ken-	krok; khrok	rai⁵⁵	ta³¹si⁵⁵ tup⁵³	(mu-) rin	[ro]

---

71 The Lepcha form is literally mí 'fire' + mak 'die'. Lepcha mak 'die' is unlikely to be cognate with PT *mit 'extinguished'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
finger[72]	*ke(ŋ)	mdzub-mo	jak-si	lak-hñúi	a³¹bɹuŋ⁵⁵	duŋ³⁵	gi-tso?	ká-jóm
fire	*me	me	waʔl	mí	na³¹mun⁵⁵	mǎi⁵³	mai?	mí
fireplace	*ram ~*rom	(me-) thab	cu-dap	mí-láŋ-phui	sai⁵³ gɹoŋ⁵³ gɹoŋ⁵⁵	sai⁵⁵ gɹoŋ⁵⁵	lo?	[kom]; [dap; dop]
fish	*ŋo	nya	naʔ-tok	ŋǎ	ta³¹ŋaŋ⁵³	a³¹ ŋa⁵⁵	thui; tčui	ŋo
five	*ŋo	lnga	boŋ-a	ŋǎ	ma³¹ŋa³⁵	kw³¹len⁵⁵	bu-ŋu	fá-ŋo
flee	*kat¹	ʼbros	#ke-ne kat	thwak- prě; hroŋ	lui⁵³	lun⁵⁵	---	tor; tet
float	*bjaŋ	lding	#git-cho; bal-bo	po	reu⁵⁵ a³¹	jau⁵³	---	plyuñ
flow	*but	ʼbaß; rgyug	#jo-kang; so-ol-ang	cí	#blum	#tai	---	dáñ; nóñ; yǔ
flower	*puŋ -pun	me-tog	bi-bal	ə-pwaŋ'	ta³¹pu⁵⁵	phan⁵³	ou-bo?	ríp; [bor]
fly (n.)	*jiŋ	sbrang-bu	tam-pi	yaŋ; phrut	ta³¹ liau³⁵	giul³⁵	bu-luŋ?	süm-bryoñ
fly (v.)	*bjar	ʼphur	bil	pyam̌	jim³⁵	phiuŋ⁵⁵	gu-nui	lám
foot	*le	rkang-pa	jaʔ	khre	gɹoŋ⁵³	pla⁵⁵	lai	(a-) t'óñ; (a-) dyañ
forget	*mit-pan	rjed	gu-al	me'	we⁵⁵ma³¹ sa⁵³	a³¹mlaŋ⁵⁵	thlaŋ	hryu; plón; myón; pán
four	*pri	bzhi	bri	lě	ka³¹pɹai⁵⁵	kw³¹ bɹun⁵³	b(w)-li	tá-li
fowl	*rok	bya-de	#do-o-rang	krak	tiu⁵³	kɹai⁵⁵	du-zu	hik (-kǔp)
frog	*tuk	sbal-ba	#beng-bek	phǎ	pa³¹rai⁵³	kaŋ⁵⁵ khɹik⁵⁵	---	tǎ-lǔk

72 The ká- 'hand' element in the Lepcha form seems unlikely to be cognate with PT *ke(ŋ) 'finger'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
fruit	*zo; *pu	shing-tog; ʼbras-bu	bi-te	ə-sí	ta³¹ pi⁵³	sit⁵³	ou-then	[pót]
full	*bruŋ	gang	gap	praŋ̂ʼ	bluŋ⁵⁵	phlǎŋ⁵⁵	---	a-blyaŋ̂
gall	*pu	mkhris-pa	ka?-kit	sáñ-khre	thu³¹-mun⁵⁵	măn⁵³	---	*kʼí-bo
give	*bi	ster; skur; sbyin	o?n	pê	ŋŋ³⁵	pi⁵⁵	bi(?)	byì(n); bí;bo
grand-father	*to	mes-po	a-cu	ə-phúi; ə-bhúi	a³¹ tia⁵⁵	kuŋ³⁵	a-luw	tʼi-kuŋ̂
grand-mother	*jo	phyi-mo; ma-mo	am-bi	phwâ; ə-bhwâ	a³¹ ja⁵⁵	mǎi³¹ ŋu³⁵	ə-žui	nyi-kuŋ̂; nyo-kuŋ̂
guts[73]	*kri	rgyu-ma	bi-bik	u	kw³¹ łɛi⁵⁵	xa³¹ łɛi³⁵	luŋ	tă-klí
hair (on body)[74]	*mut	spu	kin-i; kim-ir	ə-mwê	m̥⁵⁵	bul³⁵	phiw	myel
hand/arm[75]	*lak	lag-pa	jak	lak	a³¹ tio⁵⁵	rau⁵³	gi	kǎ; ká
have/exist[76]	*duŋ	yod; ʼdug	doŋ	hriʼ	i⁵⁵, əŋ⁵⁵	tɕau⁵³, kam³⁵	du	nyí

[73] The Dhammai form is also glossed 'heart'.

[74] For the phonologically reduced Taraon form m̥⁵⁵, cf. Chakravarty et al. 1963 um 'hair (on body)'.

[75] For (ZMYYC) Kaman rau⁵³, cf. also Boro 1979 rok; Weidert 1987:479 rǎuk 'arm'. The r- initial of these Kaman forms is perplexing, especially since Kaman apparently maintains the PTB contrast between *l- (e.g. lǎuŋ 'stone' < PTB r-luŋ; lap⁵³ 'leaf' < PTB *lap) and *r- (e.g. rɛm³⁵ 'otter' < PTB *s-rɛm; ruwl³⁵ 'snake' < PTB *b-ru:l).

[76] In both Taraon and Kaman, several existential verbs are distinguished: Taraon i⁵⁵ and Kaman tɕau⁵³ occur with animate subjects, Taraon əŋ⁵⁵ and Kaman kam³⁵ with inanimate ones, a third Kaman existential verb tun⁵⁵ applies only to abstract qualities (Sun et al. 1980). A different type of semantic differentiation of existential verbs is reported in Apatani A, based apparently on posture of the predicated subjects, but comparative data from other Tani languages is not sufficient for deciding whether this distinction should be pushed back to the PT level. The different Tibetan existential verbs reflect rather the *pragmatic* distinction of

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
head	*dum; *tuk	ngo	sko	khôn	kru⁵³	kɯɯ⁵³	u	a-t'yak; tok
heart (organ)	*puk	snying	ka?-toŋ	hna'-lûn	xa³¹po⁵⁵ tiai⁵³	lum³⁵	luŋ	a-lŭt
heavy	*ñit	ljid-po	#jrim	lê	wau⁵⁵ɐ⁵⁵	ka³¹lɐŋ³⁵	mɯ-lji?	li; bryón-nǎ; glám-lǎ
horn	*reŋ	rwa	groŋ	khyui	rau⁵⁵	kɹɐ̆ŋ³⁵	šu-žuŋ	(a-) rŏ̌ñ
horse	*ku	rta	#gu-re	mrán	ma³¹roŋ⁵⁵	pa³¹xoŋ³⁵	šu-gro	on; *ta
hundred	*luŋ	brgya	rit-ca	ra	ma³¹ lɯm⁵⁵	wa³¹je⁵³ mɯ⁵³	bɯ-loŋ	k'a fǎ-ño
hungry	*kV-noŋ	ltogs; bkren	#ok-kri a-ni	cha; mwat; ŋat	na³¹ tioŋ⁵³	di³¹ iŋ⁵³	ɫen-či	krit
I	*ŋo	nga	aŋ	ŋa	xaŋ³⁵	ki⁵³	ŋaŋ	kǎ-do; go
ill	*ki	na(d)	sa; jom	na	ŋaŋ³⁵	nat⁵⁵	no	dǎk
insect	*pum	'bu	jo?ŋ	pô	ta³¹ pɯm⁵⁵	klaɯŋ⁵⁵	bi-luŋ?	[bik]
iron	*rjok	lcags	sil	sañ	sai⁵³	tɯ³¹gli⁵³	sen	pŭn-jen; lǎñ-sǎ a-lŭt
itch⁷⁷	*ñak	'bun; za	#mi-to; ka-kit	yǎ	ma³¹ so⁵³	phuŋ⁵³	gu-dzu	jak
kidney	*krat-pjul	mkhal-ma	#gi-la; ko-rong-te	kjok-kap	çi⁵⁵	ntçhi⁵³	nɯ-gɯ-bau?	*k'a-dok
kill	*man	gsod	so?t	phyak; sat	se⁵⁵	sat⁵⁵	wai	sót

degrees of knowledge integration: ɣod for fully assimilated knowledge and 'dug for new, unassimilated knowledge (DeLancey 1989).

77 Taraon ma³¹so⁵³ is undoubtedly cognate with PT *ɫak, both reflecting PTB *m-sak 'itch' (STC # 465). For the equation PT *-ak <> Taraon -o, cf. also PT *rjak. Taraon lio⁵³ 'lick'; PT *jak, Taraon jo⁵³ 'fox-tail millet'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
knee	*lə-buŋ	pus-mo	jaʔ-sku	dù	pha³¹ bum⁵⁵	pa³¹pau³⁵	lai gu-phiw	tŭk-pă̆t
knife	*rjok	gri	a-te	thá	ta³¹ra⁵⁵	sot³⁵; kra³⁵	vai-	ban
know78	*ken	shes; mkhyen [hon.]	u-i	siʔ	ka³¹sa⁵³	ŋit³⁵	ñi; zu-u	tʼyak; yă
language	*gom	skad	ku-sik	bha-sa	khi⁵⁵tw³¹ kw⁵⁵	khi⁵⁵ lai⁵⁵	lau	a-riñ
laugh79	*ŋil	dgod	ka-diŋ	réi	ma³¹ra⁵⁵	krit⁵⁵	tho	tʼyăn; sak prok; zól
leaf	*na	lo-ma	bi-jak	aʼ-rwak	naŋ³⁵	lap⁵³	ou-leʔ	lóp; a-nyóm
leech (land)	*patˡ	pad-pa	ru-at	hnyoʼ	ka³¹pe⁵³	tw³¹vat⁵³	dw-veʔ	-fót; šŭm-pat
left-side	*lak-ke	g.yon	jak-a-si	lak-wái	tw³¹ kɪm⁵⁵	kw³¹vai⁵³	su-vjoʔ	vím
lick	*rjak	ldag	*cha-srak	yak	llo⁵³	lo⁵³	---	*lók
liquor	*poŋ	chang	cu	se	ju⁵³	si⁵³	čaŋ	čí
listen/ hear80	*tas; *tas-paŋ	nyan; thos	kin-a-	ná-thoŋ; krá	tha³¹ ruŋ⁵⁵; tha³¹ tiwŋ⁵³	ta⁵⁵ giat⁵⁵; tat⁵⁵	rui	tʼyo
liver	*zin	mchin-pa	bi-ka	aʼ-săñ	ruⁱ⁵⁵xa³¹ tiei⁵³	blai³¹ blai³³	mw-thun	a-byet

78 In the sense of 'have knowledge of'.

79 This PT root is quite unique in Tibeto-Burman. The only extra-Tani cognate known to us so far is Tshangla ŋar 'laugh'.

80 In languages that distinguish 'listen' from 'hear', forms for both meanings are given (in that order), separated by a semicolon. In Tani, the same root occurs for both meanings: the punctual, nonvolitional sense 'see' is expressed by adding to the root a resultative verbal particle -poŋ. This is true of such other pairs as 'listen' vs. 'hear'; 'search' vs. 'find'. The Garo form means 'hear'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
look/ see81	*kaŋ; *kaŋ-peŋ	lta; mthong; rig	ni-; nik-	krañ; mraŋ	ruɛŋ⁵³; kɑ³¹; tiuŋ⁵³	thoŋ⁵⁵; ŋuŋ⁵⁵	waŋ	ñak, ši; hyón
louse (head)	*ruk	shig	tik	sén		sa1⁵³	fi?	*šǎk
man (homo)	*mi	mi	man-de	lu	me³⁵	tsoŋ³⁵	ñi?	mǎ-ró
marrow82	*loŋ-kin	rkang; ngo-bo-nyid	*gheu	khraŋ-chi	ruʔ⁵³ suʔ⁵³	xiŋ⁵³	---	yǎñ; sǔñ-dǎk
meat	*dun	sha	be?n	(a)-sǎ	ta³¹; bleŋ⁵³	çin⁵³	šu-čuŋ	a-mǎn
melt	*jit ~ *jet	bzhu	*jron-gat	pyo	ji⁵³	jɐu⁵⁵; krɐ̆l⁵⁵	---	*jǔ; *šǔ
monkey83	*be:	spra; spre(’u)	---	myok	ta³¹ min⁵³	a³¹muŋ³⁵	šu-bo	sǎ-hǔ
moon	*po-lo	zla-ba	Ja-joŋ	la’	xa⁵⁵ lo⁵⁵	la1⁶³	lu	lǎ-vo
mortar	*par	sgog-ting	ca?-am	chuṃ	loŋ³⁵	gloŋ³⁵	dw-lo	[tsam]
mountain	*di	ri	a?-bri	ri’	thɯ1⁵⁵; ja⁵⁵	a³¹ dzau³⁵	phuŋ-	hlo; rók
mouth84	*nap-paŋ; gam	kha	ku-sik	pǎ-cap; mé-ce’	thɯ³¹; rɯɯ⁵³; bɯɯ³⁵	ntɕhɯɯ⁵³	go	a-boñ

81 In languages that distinguish 'look' and 'see', both forms are given (in that order) separated by a semicolon.

82 This is not considered cognate with PT *-kin, because the regular reflex of the PTB medial vowel *-i- seems to be -ǎ-. (i.e. short -a-) in Kaman (but *-i- or *-w- in PT; e.g. sǎŋ³⁵ 'tree' < PTB *siŋ; a³¹mǎŋ 'name' < PTB *r-miŋ; mǎn⁵³ < mǎt < PTB *mit 'extinguished'; ntšǎn 'claw' < PTB *m-(t)sin.

83 The -ŋ in the ZMYYC Kaman form a³¹muŋ³⁵ seems secondary; cf. Weidert 1987:358 ?mǔk; Boro 1978 a-mǔk, both keeping the original -k coda; the latter Kaman forms are cognate with PLB *myokᴸ (Matisoff 1972 #133) < PTB *mruk STC:112.

84 The Dhammai form go could not be cognate with PT *gam because the expected Dhammai equation to PT (and PTB) *-am is -en; e.g. Dhammai lem-baŋ (< len-) PT *lam 'road'; Dhammai nen, Western Tani *nam 'house'; Dhammai ñen, PT *nam 'smell v.'; cf. also Dhammai sen < PTB *šam 'iron' (STC #228).

				WB	Taraon	Kaman	Dhimmal	Lepcha
nail	*zin	sen-mo	#jak-skil	lak-séñ	a³¹ ʦɯŋ⁵⁵	rɐu⁵³ dzit⁵⁵	gi-thɯn	pŭn-ći
name	*mɯn	ming	bi-mɯŋ	na-mañ	a³¹ mɯŋ⁵⁵	a³¹ mɐŋ⁵⁵	min?	a-bryañ
neck85	*lɯŋ	ske; mgul; 'jing-pa	git-dok	lañ-péŋ	pa³¹ ŋ̊ŋ⁵⁵	xuŋ⁵⁵	---	[tok]; [liñ]
nest	*sup	tshang	bi-tip	suik	a³¹ jɯ⁵⁵ #pjaː-sag	mphɐu⁵³, #ŏ-wa sap	---	-šap
night	*jo	nam; mtshan-mo	vel	naˀ, nañˀ	kɯ³¹ jɐ⁵⁵	ŋal⁵³	jaŋ-gou	[nap]
nine	*kjo-naŋ	dgu	sik-u	kŭl	ka³¹ rɯŋ⁵⁵	nan⁵⁵ mu⁵³	sɯ-thɯn	kă-kyót
nose	*ña-pum; ña-buŋ	sna	giŋ-tiŋ	hna-khóŋ	xa³¹ nia⁵³ pɯm⁵⁵	min⁵⁵ nioŋ³⁵	ñi	[nóm]
old (of things)	*ku~ *kju	rnying-pa	git-cam	hóŋ	me⁵³	tauŋ³⁵	mɯ-šwo	[ño]; sŭk-kyor
one	*kon	gcig	se	tac	khɯn⁵⁵	kɯ³¹ mu⁵³	uŋ	ket
otter	*ram	sram	mat-tram	phyañ	xa³¹ xɯŋ³⁵	ram³⁵	---	sɐ-ryóm
palm	*lak-pro	lag-mthil; thal-mo	jak-pa	vá	#aː-tjo-kaː	#rok ta-pa	gi dɯ-luŋ	[lyók]
penis	*mrak	mje	---	lí	#mlɔ̃	#jaŋ	---	t'ik
pig	*rjek	phag-pa	vak	vak	bɯ³¹ liai⁵⁵	li⁵⁵	žo	món
placenta86	*mam	sha-ma	---	a-khyáŋ	a⁵⁵ po⁵⁵	sa⁵⁵ sap⁵⁵	---	kap-p-ŭñ; ˀayeñ-ćót (~ tyól)

85 For the Taraon form pa³¹ŋ̊ŋ⁵⁵. cf. Chakravarty et al. 1963 paː-haŋ.

86 The Taraon and Kaman words are composed respectively of 'child' + 'protect' and 'child' + 'nest'. As for the Lepcha forms, kap-pŭñ is literally 'covering', that which covers; while ˀayeñ-tyól is 'child' + 'accompany'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammal	Lepcha
poison[87]	*dɯk; *mro	dug	*bi-si	ə-chip	tha^{53}	tau^{53}	nu-phaŋ	[bo]; a-nyiñ
put	*pa	'jog	don-	thá	xa^{31}goss^{55}	kials^{55}	rou	dya; t'o
rain (n.)	*pV-doŋ; *mV-doŋ	char	mik-ka	mũi rwa	ka^{31}raS^{55}	a^{31}waŋ^{55}	phrjo	so
rat	*ko-buŋ	byi-ba; tsi-tsi	*me-se	krwak	ka^{31}tɕiS^{55}	Si^{55}nu^{53}	--	kă-lŏk
red	*luŋ	dmar-po	git-cak	ni	ɕi^{53}	kap^{31} sal^{35}	mu-tsu	a-hyir
rice[88]	*pim	'bras-chan	mi	tha'-mán	ta^{31}peŋ^{35}	ɕat^{53}	an tsa-vo	nŭm-or-mo
right-side	*lek-bruk	g.yas	jak-ra	ya	tu^{31}tɕa^{55}	ku^{31}jau^{53}	Si-dzin	gyóm
ripe	*min	smin-pa	min-	chim'; hmañ'	*ha:-muŋ	*shu-mm	min	[kru]; a-mǎn
river	*si; *buŋ	chu	ci-bi-ma	mrac	tu^{31}lwu^{35}	tu^{31}lo^{35}	vu-do	uñ kyoñ
road	*lam	lam	ra-ma	lám	a^{31}lim^{55}	bloŋ^{35}, lam^{55}	lem-baŋ; hlen	lŏm
root	*pur; *m(i)a	rtsa-ba; rtsad	ja?-dir	ə-mrac	xa^{31}raiS^{55}	kɹa^{53}	-khrin	a-fja; a-běn; [sén]
rot	*jaŋ	rul	ta?m-bi?	pup	tshuŋ^{55} xo^{31}	ram^{53}	---	byót
round (globular)	*lum	ril-ba; zlum-po		vũiŋ; lũi	geŋ^{55} weŋ^{55} dass^{55}	gass waŋ^{55} nass	mu-du-riu	a-blam; a-pŭm
salt[89]	*lo	tshva	ka-ri	chá	pla^{35}	tu^{31}min^{55}	lu	vŏm

87 Cf. the Chakravarty et al. 1963 tha:ik for Taraon and Boro 1979 tŏuk for Kaman, both retaining the -k coda.
88 More precisely 'cooked rice'. For the Kaman form ɕat^{53}. cf. Weidert 1987:479 má-syát 'boiled rice' (root = syá 'eat' plus nominalizing dental suffix -t).
89 The Taraon form pla^{35} seems to come from earlier *plaŋ (cf. Midu prã 'salt') and therefore phonetically quite distant from PT *lo.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammal	Lepcha
scratch	*ɓok	'phrug; phur	ku-ak	yak; phrok~phyok	va⁵⁵	glua³⁵	gu-fja?	hut
seed	*li	sa-bon; son	bit-cri	myúi-ce'	ta³¹ plai⁵⁵	xa³¹ lɯi³⁵	thei-žo	lí
sell	*pruk	'-tshong	pal	rôŋ	kha³¹ ji⁵⁵	xa³⁵	tsuŋ-ru	úl
seven	*kV-nat	bdun	sin-i	khu-nac	weŋ⁵³	nɯn⁵³	mja?	kă-kyăk
sew	*ɦom	'tshem	*sik;ko	khyup	*ru	taŋ⁵⁵ kɹap⁵⁵	bu-ča	hrap
sharp-edged	*rat	rno	mat	thak	ɹa⁵⁵	kɹat⁵⁵	---	lăt~let
shoot[90]	*ap	'phen	go	pac	o⁵³ ja³¹	top⁵⁵ kap³⁵	buw	óp
shoulder	*gor-	dpung-pa; phrag-pa	pak-re	pu'-khûm	khu⁵⁵ liuŋ⁵³ pa³⁵	a³¹ pho⁵⁵	pa-stuŋ	tŭk-puŋ
shy	*han-ñiŋ	skyeng; khrel; 'dzem	*kat-ca	hrak	*ha:-la:g-a:	*i-juk-rai	dai	uk;a-mlem glo
sit	*duŋ	sdod; 'dug	a-soŋ	thuiŋ	di⁵⁵	lăp⁵⁵	juŋ?	ńen
six	*kre(n)	drug	dok	khrok	ta³¹ xɹo⁵³	kɯ³¹ tam⁵³	re?	tă-răk
skin	*pin	(l)pags-pa; ko-ba	bi-gir	a-re	ko⁵⁵	uŋ³⁵	phri?	a-kap; a-t'un; a-pi
sleep[91]	*jup	nyal; gnyid-log	tu-si	ip	ŋ⁵³	ŋui⁵⁵	ji	mik krap
smell (v.)	*nam	snom	---	năm;hru	nuŋ³⁵	ntshiŋ⁵⁵	ñen	n(y)óm

90 The Taraon form o⁵³- is judged to be cognate with PT *-ap. For the equation PT -ap <> Taraon -o. cf. also PT *krap. Taraon khro 'weep'.

91 The resemblance between Dhammal ji to PT *jup is misleading, for the Dhammal form could originate from a nasal-final rhyme. cf. Bangru dʑe³³; Hruso jum 'sleep'. The Lepcha compound is literally mik 'eye' + krap 'hang down'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
smoke (n.)	*mu-ku	du(d)-ba	wa?l-ku	mí-khûl	ma³¹ khuul⁵³	ta³¹ khuul⁵³	thuŋ	mí-kan
snake	*bu	sbrul	cip-bu	mrwe	ta³¹bu⁵⁵	ruul³⁵	nu-buv	bŭ
soft	*mjak	mnyen; 'jam, snyi	#nom;ri-nok	pyo'	ñim⁵⁵ m̥⁵⁵	ka⁵⁵ niŋ³⁵	mu-bu-lja?	nŭm
son-in-law	*mak-bo	mag-pa	#ca-wa-ri	sá-mak	ku³¹mu⁵³	tsa⁵³	---	myók
soul/spirit	*ja-lo	nyam(s)	#jaŋ-gi sil-ci; gi-sik	lip-pra	ta³¹gia³⁵	ka³¹mau³⁵	---	a-pil; [jŭm]; hyit
sour	*kruŋ	skyur	me-seŋ	khyañ	xum⁵⁵	sa¹⁵⁵	mu-čuŋ	a-čór; rók-nón
spittle[92]	*kjul	mchil-ma	ku-ci	tam-twé	khu³¹lai³⁵	dzal³⁵	že?	dyuk
stand	*dak; *rop	'greng	ca-deŋ	rap	deŋ³⁵	loŋ⁵³	gjuŋ	diñ
star	*kar	skar-ma	a-ski	kray	kha³¹dun⁵⁵	ku³¹grun³⁵	do-tsuŋ	sã-hór
steal	*pjoŋ	rku	ca-u	khui	a³¹kau⁵³	ka¹⁵⁵ xuul⁵³	tsu-khu?	tŭk-mo mat
stone	*luŋ	rdo	ro?ŋ-te	kyok < klok	phlaŋ³⁵	lauŋ³⁵	gu-luŋ	lăñ
suck	*bruŋ	'jibs	op	cut; cui'	du⁵⁵	jip⁵⁵, #that	bu-nu	yup; háp
sun	*ñi	nyi-ma	sal	ne	run⁵³	min³⁵	jo; zu?	sã-tsŭk
swallow (v.)	*met	(khyur) mid	#mi-nok	myui	blai⁵³	biap⁵³	bu-lui	yop; hyul; am-mat
sweet[93]	*ti:	mngar~dngar	ci-	khyui	ɣau⁵⁵	tim³⁵	mu-jaŋ	a-klyam

92 The Garo word means 'saliva'; from ku 'mouth' + ci 'water'.
93 The Taraon form ɣau⁵⁵ seems to come from a checked syllable. cf. Chakravarty et al. 1963 shyeb 'sweet'.

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
swidden	*ruk	zhing-ka	a-ba	ley	kha³¹ liau⁵⁵	a³¹kuŋ⁵⁵	vav	nyót
tail	*mjo~ *me	rnga-ma	kiʔ-me	a'-mrî	lu³¹ mun⁵⁵	a³¹mɐ̃i⁵⁵	---	[ší]
take	*laŋ	'khyer; len~ loŋ	raʔ-; rim	yu	çi³⁵	ta³¹lɐt⁵⁵	luʔ	lyă; le; lyo
ten	*rjwɥ; *čam	bcu	ci-kiŋ	ə-chai	xa³¹ luŋ⁵⁵	kiɐp⁵⁵ mu⁵³	lin	kă-tí
thick (book)	*bruŋ	mthug	rit-caʔ-	thu	bi³¹ tçoŋ⁵⁵	bi³¹ tçoŋ⁵⁵	---	tăn
thin (book)	*bV-čor	srab	ba?-	pâ	ba⁵⁵ a⁵⁵	kɯ³¹pa³⁵	mu-dɯ-thaŋ	sap
think	*muŋ	sem(s); bsam	can-ci	thaŋ; cáñ	ta³¹we⁵⁵	ntshum⁵⁵	mjen; šu	(sak)čiñ
thou	*no	khyod; khyed; nyid [hon]	naʔŋ	naŋ	ɲoŋ³⁵	ɲo⁵³	ɦi	hó; a-do
three	*ɦum	gsum	git-tam	sûm	ka³¹ suŋ³⁵	kɯ³¹ sɐ̃m⁵³	gu-thun	sam
tiger	*mro (*mjo?); *paŋ-te	stag	mat-ca	kyâ	bo⁵⁵da⁵⁵, *ta:-mja	bo⁵⁵da⁵⁵	tiŋ-graŋ	să-t'ăn
tired	*pe	dub; thang chad	neʔŋ-	mô	giɐi⁵³; *he-ra:	çɐi⁵⁵; *min-jin	khaŋ-ru	pyăl
tongue	*rjo	lce	sre	hlya	thɯ³¹ lim⁵³ na³⁵	blai⁵³	že?-yi	a-lí
tooth	*fi:	so	wa-gam	swâ	laŋ³⁵	si⁵⁵	thu	a-fo; fo-ki
two	*ɦi	gnyis	gin-i	hnac	ka³¹n⁵⁵	kɯ³¹jin⁵³	gni	nyăt; nyi

Gloss	PT	WT	Garo	WB	Taraon	Kaman	Dhammai	Lepcha
urine	*sun, *si	gcin; (dri-) chu	su-bu	chí	ku³¹ tɕɯŋ⁵⁵	tɯ³¹ ɕit⁵⁵	brui?	jit
village	*nam-pom; dup-luŋ	yul-gru; grong	soŋ	rwa	ma³¹ tiɯŋ⁵⁵	mu³¹ təŋ⁵³	gu-bjaŋ	li brom; li broñ; li kyoñ
vomit	*bat~ *brat²	skyug	*ci-sat; wa-kal	am	me⁵³	phat⁵⁵	mu	mót; hluñ
wash body; bathe	*hur	'khru ~ 'khrud; chu rgal	a-u	khyui'	ma³¹ num⁵⁵ tsai⁵³	ta³¹ ruw1³⁵ lai⁵³	---	mŭ-tŭt; mŭ-cóñ
water	*si	chu	ci	re	ma³¹ tɕi⁵³	a³¹t i³⁵	vu	uñ
weave	*čun	'thag	dok	rak	ta³¹ tiɯ⁵⁵ tio⁵³	tho⁵⁵ tan⁵⁵ tho⁵⁵	čun	t'ok
weep[94]	*krap	ngu; shum; khrap	grap	ŋui	khro⁵³	ŋai⁵⁵	---	hryóp; prám mat
wet	*ju-jaŋ	rlon-pa	so-si	cui; cwat	pun⁵⁵	phom⁵⁵, #kan-sak	mu-gro?	šal
white	*pun~ *puŋ	dkar-po	gip-bok	phru	lio⁵³	kɯ³¹ mphlaŋ⁵⁵	mu-grjeŋ	[du]
wind	*rji	rdzi; rlung; lhag-pa	bal-wa	le	ra³¹ rɯŋ⁵⁵	baɯŋ³⁵	jo	sŭn-mŭt; so-mŭt
wing[95]	*lap	gshog-pa; 'dab-ma	graŋ	a'-toŋ	ta³¹ loŋ⁵⁵	ŋkhloŋ³⁵	gu-či	pă-ku; pŭñ-ku

94 WT khrap occurs only in the phrase khrap-khrap 'weeper, cry-baby'. The normal 'weep' meaning has been taken over by the ngu root.
95 WT 'dab-ma (< N + lap) is a direct cognate of PT *lap. The dental stop initial is transparently caused by the homorganic nasal prefix N- (represented orthographically by the achung). For more evidence of the effects of achung, cf. 'dom (< N + lom) 'fathom' < PTB *la(t)m (STC p.71); 'do (< N + lo)~ zlo 'say, repeat'; this view is also strongly supported by the identical delateralizing effect of the m- nasal prefix, cf. WT mda (< m + *la); PTB *mla~bla 'arrow' (STC fn. 313). For a different

187

wood	*suŋ	shiŋ	bol	sac	ma³¹ suŋ⁵³	saŋ³⁵ khliŋ⁵⁵	u	šaṅ̇; kuṅ̇
year⁹⁶	*hiŋ	lo; -niŋ	*bil-si	hnac	kɯ³¹ naŋ⁵⁵	leu⁵³	du-ren	nam (tum)

interpretation of the provenance of this WT form (owing perhaps to a different view on the phonetic nature of WT achung). cf.
Matisoff 1985a:443-4 as well as STC: 122-3; fn.338, 339.
96 In WT, the root -niŋ 'year' occurs only in compounds, such as na-niŋ 'last year'.