Sino-Tibetan Numerals and the Play of Prefixes

Sino-Tibetan Numerals and the Play of Prefixes

James A. MATISOFF[*]

SYMBOLS AND ABBREVIATIONS

A ⋇ B	A and B are co-allofams; A and B belong to the same word-family
AMD	Abor-Miri-Dafla
BIHP	Bulletin of the Institute of History and Philology (Canton; Taipei)
BMFEA	Bulletin of the Museum of Far Eastern Antiquities (Stockholm)
BSI	Bible Society of India
CSDPN	*Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal* [HALE (ed.) 1973]
GEM	Geoffrey E. Marrison 1967
GSR	*Grammata Serica Recensa* [KARLGREN 1957]
GSTC	"God and the Sino-Tibetan Copula" [MATISOFF 1985b]
Him.	Himalayish
HJAS	Harvard Journal of Asiatic Studies
JASB	Journal of the Asiatic Society of Bengal
Jg.	Jingpho
JRASB	Journal of the Royal Asiatic Society of Bengal
KCN	Kuki-Chin-Naga
LSI	*Linguistic Survey of India* [GRIERSON and KONOW (eds.) 1903-28]
LTBA	Linguistics of the Tibeto-Burman Area (Berkeley)
MC	Middle Chinese (= Karlgren's "Ancient Chinese")

---

* University of California, Berkeley

NBP	Nagaland Bhasha Parishad (Linguistic Circle of Nagaland, Kohima)
OC	Old Chinese (= Karlgren's "Archaic Chinese")
PIE	Proto-Indo-European
PLB	Proto-Lolo-Burmese (=Proto-Burmese-Yipho)
PNN	Proto-Northern-Naga
PST	Proto-Sino-Tibetan
PTB	Proto-Tibeto-Burman
STC	*Sino-Tibetan: a Conspectus* [Benedict 1972]
STEDT	Sino-Tibetan Etymological Dictionary and Thesaurus Project (Berkeley)
TB	Tibeto-Burman
TBL	*A Tibeto-Burman Lexicon* [Dai and Huang 1992]
TSR	*The Loloish Tonal Split Revisited* [Matisoff 1972a]
VSTB	*Variational Semantics in Tibeto-Burman* [Matisoff 1978a]
WB	Written Burmese
WT	Written Tibetan
ZMYYC	*Zang-Mianyu Yuyin he Cihui* [Chinese Academy of Social Sciences 1991]

1. INTRODUCTION

1.0 Background

This study was originally prepared for the 17th International Conference on Sino-Tibetan Languages and Linguistics at the University of Oregon (September 1984),[1] but has been languishing on the back burner for nearly a decade. Although it is the most extensive synchronic and diachronic treatment of the Tibeto-Burman numerals yet attempted, considerations of time and space have made the present revised version less complete than I would have wished. I have tried to include data from as many languages as feasible, but the coverage is far from exhaustive, and is of uneven depth with respect to

---

1) I would like to thank Nicholas C. Bodman and David Bradley for their cogent comments on the first version of this paper. My thanks are also due to Chang Kun for the numerals of the Zida (Tzuta) dialect of rGyarong; to Gérard Diffloth for information on numerals in Mon-Khmer languages; to Robert Goldman for helping me read a Hindi source on Pochury; to David Solnit for data on Kayah Karen; to Purna C. Thoudam, for sending me over a dozen translations of the Bible into TB languages of India; to Graham Thurgood for forms from Idu and Rawang; to Chhewang Rinzin for making clear tape recordings of the numerals in Sharchop and Dzongkha; and to Gabrielle Yablonsky for putting me in touch with Mr. Rinzin. For more recent helpful comments and practical assistance I am much obliged to Ives Goddard, Eric Hamp, Gary Holland, Adam Jacobs, Terrence S. Kaufman, Ian Maddieson, Eric Oey, Eve Sweetser — and of course the "STEDTniks", especially Leela Bilmes, Michael Brodhead, Jonathan Evans, Zev Handel, Matthew Juge, John B. Lowe, Pamela Morgan, and Ju Namkung.

Tibeto-Burman as a whole. Of the hundreds of TB languages and dialects,[2] the ones whose numeral systems have been examined for this paper are listed in the Index of Languages and Sources.

A fine-tuned subgrouping of the TB languages is an essential longterm goal, though for the moment it belongs in the realm of *Zukunftsmusik*-music of the future! For our present purposes we shall have to be satisfied with a schematic family-tree like the following, where each major subgroup of the family is portrayed as branching off coordinately from the proto-language. See Figure 1.

This scheme differs in several respects from the diagram presented in *Sino-Tibetan: a Conspectus* [BENEDICT 1972; henceforth "STC"].[3] In any event,

Figure 1   The Branches of Tibeto-Burman

---
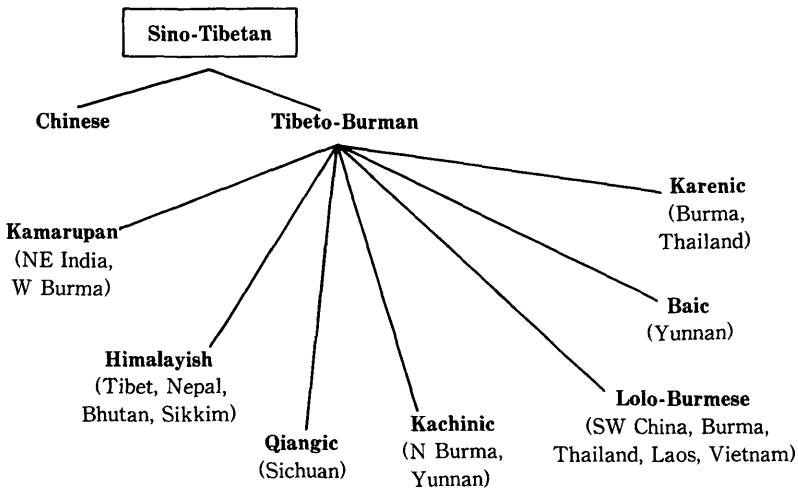
2) For a fairly complete alphabetical list of TB languages, where each is assigned to a subgroup of the family, see Matisoff 1986. A revised and expanded version of this list is to appear as Volume II of the STEDT Monograph Series (1995).

3) In the chart in STC (p.6) Karen is excluded from "TB proper", largely on syntactic grounds (since it is SVO, while TB proper is SOV). The rest of TB is indicated as radiating out of Kachin (=Jingpho), to which Benedict accords a genetically central position, both for geographical and lexico-phonological reasons. The STC does not recognize the Qiangic languages as a separate subgroup at all, not surprising since most of them have only recently been put into relief by Chinese scholars. I have suggested the term "Kamarupan" (from Kamarupa, the ancient Sanskrit name for Assam) as a neutral, overall geographical designation for the branches of TB spoken in NE India and adjoining regions, pending the vast increase in our knowledge that will be necessary before we can sort these dozens of languages out with more precision. (Kamarupa is home to more TB languages than any other region, and is the "center of diversification" of the entire family.) The genetic position of the highly Sinicized Bai language (formerly called "Minjia") is still controversial [see ZHAO 1982; WIERSMA 1990], though it seems safest for now to assign it to a subgroup of its own.

there are vast differences, both qualitative and quantitative, in the data available for particular TB languages or language-groups.

For the Naga languages we are still heavily dependent on G.E. Marrison ["GEM"] 1967, a rich source mined to excellent effect by W.T. French 1983.[4] When used with caution, the little glossaries produced by the Nagaland Bhasha Parishad ["NBP"] (Linguistic Circle of Nagaland) are also useful sources of information. For some Chin languages (Hmar, Gangte, Kom Rem, Kuki, Paite, Tiddim, Vaiphei) I have had to extract the numerals from translations of the Bible. (I would like to take this opportunity to sing the praises of the Book of Revelation as a numerological resource, with its Seven Seals, thousands of winged beings, and such invaluable passages as Rev. 21.19-20:*"And the foundations of the wall of the city were garnished with all manner of precious stones. The first foundation was jasper; the second, sapphire...the twelfth, an amethyst."*)

For the languages of the "Abor-Miri-Dafla" or "Mirish" group, the old data to be found in the *Linguistic Survey of India* ["LSI"] [GRIERSON and KONOW (eds.) 1903-28] has been largely superseded by recent work in Arunachal Pradesh, especially by K. Das Gupta and I. M. Simon, much of which has appeared in the modest journal *Resarun* (< **Res**earch **Arun**achal). A University of Calcutta dissertation by Shail Kumari Dubey [1983] contains useful material from several AMD languages. In China, the most important recent study of AMD languages is Sun, Lu, and Ouyang [1980], which presents highly accurate data on Monpa (Menba), Loba (Lhopa), and Darang Deng (Taraon). Most recently, Jackson T. Sun's dissertation [SUN 1993] has laid a firm foundation for the reconstruction of the "Tani" nucleus of this branch of TB.

A key compendium of data on the Himalayish languages of Nepal is Hale (ed.) 1973 ["CSDPN"]. For Hayu (= Vayu) the best modern source is Michailovsky [1981]. By happy chance, a recent treatment of the Kiranti group of Nepal TB languages [GVOZDANOVIĆ 1985] focusses directly on their numeral systems; although it appeared after the first version of this paper was written, it will be discussed in appropriate contexts below (2.1, 3.53, 4.02). Important older works on the Himalayish languages of Sikkim and Bhutan include Mainwaring and Grünwedel [1898] for Lepcha, and Sandberg [1895] for Sikkim Bhutia (= Danjongka = Dzongkha). I was able to use lists of the numerals of Sharchop (Tsangla) and Dzongkha specially tape-recorded by a native speaker, Mr. Chhewang Rinzin. An extremely interesting article on the ambiguous conceptual bases of the Dzongkha numeral system [MAZAUDON 1985] also appeared after the first version of this paper was composed (see below 3.534).

Reliable data on the Qiangic languages of Sichuan is now becoming available in quantity, thanks to the efforts of scholars like Sun Hongkai [e.g.

---

4) For full references to all works mentioned please see the BIBLIOGRAPHY.

SUN 1981, 1982a, 1985], and Lu Shaozun [e.g. LU 1983].

For Jingpho (Kachin) the classic source is Hanson [1906], now supplemented by Maran [in prep.] and two excellent dictionaries produced by Dai Qingxia, et al. (Chinese/Jingpho 1981; Jingpho/Chinese 1983). For Nungish, older sources like Barnard [1934] and Lo Ch'ang-p'ei [1942/1945] are now vastly enriched by Sun Hongkai 1982b (Dulong) and 1986 (Nung).

Not much new data has appeared on the Karenic branch of TB since Jones [1961], though important works are soon to appear (e.g. Henderson's dictionary of Bwe and Solnit's grammar of Kayah).

Lolo-Burmese,[5] perhaps the best-studied branch of TB, continues to receive its fair share of attention. On the Burmish side, Burling [1968] includes data from Atsi and Maru. More recently other first-class works have appeared on Atsi (= Zaiwa) by Yabu [1982] and Xu and Xu [1984], and on Achang by Dai Qingxia [1982]. Luce [1985] contains data on several Burmish languages, including Lashi, while Henderson [1986] refines data on Hpun collected long ago by Luce. On the Loloish (= Yi) side, useful data may be extracted from sources like Gao Huanian 1955 (Hani), 1958 (Nasu); He and Jiang 1985 (Naxi); Hu and Dai 1964 (Hani); Lewis 1968 (Akha); Ma Xueliang 1949 (Luquan), 1951 (Sani); Matisoff 1973a, 1988a (Lahu); Nishida 1966/1967 (Bisu); Srinuan 1976 (Mpi); Rock 1963 (Naxi); and Yuan Jiahua 1947 (Woni), 1953 (Axi). Comparative Lolo-Burmese studies include Nishida 1964; Burling 1968; Bradley 1978; and Matisoff 1972a, 1978b, 1979, 1994b.

1.1 Issues in the Reconstruction and Systematic Behavior of the Tibeto-Burman Numerals

In a sense this paper is a critique and expansion of the treatment of the PTB numerals presented in Benedict's *Sino-Tibetan: a Conspectus* ["STC"]. In that pioneering work, the emphasis is on the nuts-and-bolts of phonological reconstruction. Yet even a simple listing of the STC's proto-numerals[6] raises a variety of interesting morphophonemic and lexico-semantic issues. See Figure 2.

1.11 Proto-variation

Proto-variation must be recognized as just as much of a fact of life in Sino-Tibetan as in Indo-European. Reconstructed etyma should not be viewed as invariant monoliths, but rather as "word families": sets of morphophonemically and semantically related forms that cluster around a basic phonological shape

---

5) The term "Loloish" is now felt to be pejorative in China, where the term *Yi* is now politically correct (since it is no longer written with the character for 'barbarian'). The subgroup designation *Burmese-Yipho* has been suggested as a substitute for Lolo-Burmese, but for the moment I am sticking with the latter to avoid confusion.

6) See especially Section 16, pp. 93-95. References preceded by "#" refer to the numbered cognate sets in STC; other references are to page numbers.

Figure 2　Proto-Tibeto-Burman Numerals

ONE	*it [94, 162]; *kat [94]; *g-t(y)ik [84, 94, 169, 189]
TWO	*g-ni-s [#4]
THREE	*g-sum [#409]
FOUR	*b-liy = *b-ləy[7)] [#410]
FIVE	*l-ŋa ~ *b-ŋa [#78]
SIX	*d-ruk [#411]
SEVEN	*s-nis [#5]
EIGHT	*b-r-gyat ~ *b-g-ryat [#163]
NINE	*d-kuw = *d-kəw ~ *d-gaw [#13]
TEN	*gip [#16]; *ts(y)i(y) ~ *tsyay [#408]
TWENTY	*m-kul [#397]
HUNDRED	*r-gya [#164]
THOUSAND	*s-toŋ [#32]

and a core of meaning. Variability is observable in all parts of the TB syllable: rhymes, initials, prefixes, tones.[8)] To some extent STC is prepared to recognize cases of proto-variation, and its labyrinthine pages contain many more "allofamic reconstructions" than is at first apparent. An examination of Fig. 2 reveals several instances of putative proto-variation, either in the rhyme (NINE, TEN) or in the prefix (FIVE, EIGHT). Yet STC does not exactly "go the whole hog" and embrace the notion of proto-variability with enthusiasm. It is selective, sometimes even arbitrary, about which attested variants are ascribed to Proto-Tibeto-Burman and which are branded as "secondary" or explained away on other grounds.

1.111　Variation of proto-rhyme

The STC recognizes a PTB alternation *-uw ~ *-aw in NINE, on slender evidence, rejecting as secondary the better-attested variant in -wa (Lushai **pakua**, Angami **thepfə**).[9)] For TEN, the first version of STC recognized "vowel gradation" between *-ai and *-i to account for Written Burmese [WB] **ʔəchai** vs. e.g. Jg. **shi**. This view was later changed (n. 272) in favor of

---

7) Shortly before the publication of STC, Benedict changed his original reconstructions of the PTB rhymes *-iy and *-uw to *-əy and *-əw. These reconstructions are essentially equivalent (for some discussion see Matisoff 1985b, pp. 20-21), hence the equal-signs in the chart. Quite distinct from the above are cases where Benedict (explicitly or implicitly) recognizes phonological variation at the proto-stage. These are marked with a tilde in the chart. The case of TEN poses a special problem (below).

8) The theoretical framework for the analysis of variational phenomena in TB, including the notion of *allofam* (i.e. word-family alternant) has been developed at length in Matisoff 1978a, *Variational Semantics in Tibeto-Burman* ["VSTB"]. The symbol " ✕ " is there introduced to stand for the allofamic relationship: X ✕ Y 'X and Y are co-allofams; X and Y both belong to the same word-family'.

9) See the discussion in Matisoff 1980 (pp. 15-17), and below 4.24.

introducing a complication into the reconstruction of the initial consonant.[10]

TWO furnishes an example of variation of final consonant. Alongside the principal allofam in -s, a variant in *-k (underlying, e.g. WB **hnac** < *s-nik) is also attested independently in several branches of TB, but is denied PTB status in STC. (See below 4.11.)

In the course of this study, several new numerical etyma have been unearthed where the rhymes show such well-established variational patterns as alternation between homorganic final stops and nasals, or between the vocalic nuclei *-i- ⩲ *-ya-, or between the rhymes *-ay and *-an.[11]

1.112   Variation of proto-prefix

All the numerals from 2-9 are reconstructed with a prefixal element, to which no particular meaning may be assigned. In the case of FIVE, STC does admit proto-variation, positing alternation between the *b- and *l- prefixes at the PTB level.

EIGHT presents special problems, since it is a clear instance of a doubly-prefixed form even at the PTB stage. The daughter languages which retain segmental reflexes of two prefixes show wide variation both in the particular consonants "chosen" to serve as prefixes, and/or in their relative ordering. The STC recognizes this latter fact by positing proto-metathesis — i.e. metathetic co-variants that existed already at the PTB stage.

In other words, STC admits prefixal proto-variation for two numerals, FIVE and EIGHT, conceiving of this phenomenon in a paradigmatic sense for FIVE but in a syntagmatic sense for EIGHT.

However, proto-variation in prefix may with equal justice be imputed to at least two other numerals, SIX and NINE. Besides the *d- prefix for NINE, at least four branches of TB point to a sibilant prefix *s- which STC does not recognize (Garo **sku**, Kanauri **zgui**, Jingpho **džəkhû**, Pumi **sgiuh**). As for SIX, the initial velars in Himalayish, Jingpho and Lolo-Burmese (e.g. Magari **kruk**, Jg. **krúʔ**, WB **khrok**) are treated as secondary developments from the dental prefix before root-initial *r-, i.e. *d-r > kr. Yet tonal developments in Loloish, where the word appears in the LOW-stopped tone despite its voiceless initial (e.g. Lahu **khɔ̀ʔ**), clearly point to a doubly-prefixed prototype *d-k-rok.[12] This syntagmatic view of the prefixal dynamics in SIX makes its behavior analogous to that of EIGHT, another numeral with a "weak" liquid root-initial that was particularly conducive to repeated prefixation.

Of particular interest in the context of prefixal variation in numerals is the phenomenon we call "prefix runs" (below, 5.2 et seq.), whereby consecutive numerals acquire the same prefix. It is undeniable that many modern TB

---

10) Matisoff 1985b (pp. 5, 32), and below 3.22.
11) See below 3.11, 3.12, 3.14, 3.21, 3.22, 3.233, etc.
12) See Matisoff 1972a, *The Loloish Tonal Split Revisited*, pp. 35 and 71.

languages, especially in Kamarupa, have innovated by levelling out their numeral prefixes to produce runs, with the limiting case being languages like LUSHAI, where *all* the numerals from 1-9 have developed the same prefix, **pə-** (written "**pa-**"):

LUSHAI

'1'	pakhat	'2'	pahnih	'3'	pathum
'4'	pali	'5'	panga	'6'	paruk
'7'	pasarih	'8'	pariat	'9'	pakua[13]

Yet STC has to recognize two shorter prefix-runs already at the PTB level: the *g- in TWO *g-nis <=> THREE *g-sum, and the *b- in FOUR *b-ləy <=> FIVE *b-nga.[14] Does this mean that one of the two PTB prefixes posited for FIVE, *b-, might actually be "secondary", due to contamination from the *b- in FOUR, so that the "original" Sino-Tibetan prefix in FIVE was *l-? Must we assume that the further back we go, the fewer prefix runs we should find?

To me it seems more reasonable to conceive of the prefixation of numerals as a highly idiosyncratic and variable business "right from the beginning", with cyclical waves of analogical levelling and re-differentiation having occurred throughout (and before) the documentable history of the ST family.

1.12 Lexico-semantic issues

Numerals constitute a uniquely structured semantic field, both syntagmatically (because of their fixed linear order in counting), and paradigmatically (because of the multitude of mathematically precise relationships (additive, subtractive, multiplicative, etc.) in which they simultaneously participate.

Prefix runs are only one of the manifestations of the influence of adjacent numerals on each other. We shall point to cases where other areas of the syllable are clearly affected by inter-numerical contamination, including the *root-initial consonants* and *rhymes* (vowels, tones, and even suffixes) of consecutive numerals (below 4.01). In fact, it is not even necessary for numerals to be consecutive in order for them to influence each other's phonological shape. It is widely assumed that the complex initial consonant sequence in WT **brgya** 'hundred' is somehow modelled on the word for 'eight' (WT **brgyad**), though nobody has suggested any conceptual basis for this in terms of a

---

13) Lushai **pasarih** '7' is a doubly prefixed form, with the younger *pa- attached before the older sa- (< PTB *s-nis). Contra Matisoff 1980 (pp. 16-17), the Lushai form **pakua** '9' furnishes no support for a PTB *b- prefix with this numeral.

14) According to STC, SIX and NINE have the same prefix *d-, but there are no two consecutive higher numerals (6-7, 7-8, or 8-9) with the same prefix at the PTB level. For innovative runs in the higher numerals in Kuki-Naga, see 5.44, below.

mathematical relationship between '8' and '100'.[15]

Although wholesale borrowing of numerals is by no means unheard of,[16] and has reached critical proportions in many of the TB languages of Nepal (below 2.0-2.1), numerals are generally considered to be among the best specimens of core vocabulary. Indeed, the TB languages overwhelmingly reflect a single inherited etymon for each of the primary numerals from TWO to NINE.[17]

In striking contrast, there are multiple roots reconstructible for both ONE and TEN, with no single etymon distributed through all the branches of the family.[18] As we shall see, the proliferation of lexemes for TEN is undoubtedly connected to its special role as the "base" of most TB numeral systems, to its propensity for being confused or "transvalued" with ONE or TWENTY, and to its frequently ambiguous role as both a numeral and a classifier.

Several interesting issues may be raised concerning the relationship of the lower numerals (1-5) to the higher ones (6-10). First of all, from the viewpoint of language contact and lexical replaceability, the lower numerals seem much more resistant to outside influence than the higher ones. In areas like Nepal, where the local TB languages are under severe pressure from a prestigious majority language, it is common to find that the higher TB numerals have totally fallen into desuetude, while only a few of the lower ones are preserved.[19] To my knowledge no cases have ever been documented where a language has retained its inherited higher numerals, but replaced its lower numerals by borrowing.[20]

Language internally, the higher numerals may be conceptually secondary to the lower ones. A number of TB languages have lost their inherited forms for 6-9, replacing them with additive or multiplicative formations based on 1-5. Thus EIGHT may be expressed as "5 + 3", or as "4 × 2". (See below 4.20.)[21]

---

15) The STC does not exaggerate when it declares (n. 148, p. 45) that "This pair of numeral roots [EIGHT and HUNDRED] presents unusual difficulties both in TB and in Chinese."

16) As is well known, the Chinese numerals have been borrowed by Thai and Japanese, in the case of the former supplanting the native numerals almost entirely.

17) See the "profiles" of the primary numerals, below 4.1-4.24. This is not to say that isolated forms do not crop up here and there in one or another TB language or subgroup, a celebrated example being WT **bdun** '7' (below 4.22). The AMD branch of TB (below 1.25) has the most aberrant-looking numeral sets in the whole family.

18) Besides the three roots for ONE and the two roots for TEN that are reconstructed in STC (see Fig. 2), several additional etyma have been discovered for both. See below 3.15, 3.23.

19) Even Thai has retained its inherited word for ONE (**nŷŋ**), using the Chinese loan ʔèt only in compound numerals (11, 21...101).

20) This is certainly not to deny that a language may replace its lower numerals by some other means. Jingpho, while faithfully retaining its inherited etyma from 3 to 10, has introduced new lexemes for ONE (**ləŋâi**) and TWO (**ləkhôŋ**), that until recently have not been relatable to anything else. In Matisoff 1995 (to appear) I identify the former with the Jg. first-person pronoun **ŋâi** 'I'.

21) Subtractive formations are also occasionally encountered in higher numerals, e.g. EIGHT ex- ↗

A glance at Figure 2 reveals a similar phenomenon already at the PTB level: both TWO (*g-nis) and SEVEN (*s-nis) are reconstructed with identical roots, differing only in prefix. Every daughter language (even if it no longer retains any prefixes) manages to keep TWO and SEVEN distinct by one phonological means or another (below 4.11, 4.22), but it seems likely that the TB numeral system once related them conceptually.

This leads to the whole question of the "conceptual bases" of TB numeral systems. Besides the traces of QUINARITY just mentioned, there are strong indications that several other non-decimal bases have served as building-blocks for numeral systems at various stages in the history of the family. A mono-morphemic form for *twenty*, *m-kul, is reconstructible for PTB (see Fig. 2), and a number of modern languages have thoroughgoing VIGESIMAL systems of "round-number formation" (below 3.5). In many cases, however, there is hesitation between decimality and vigesimality within an individual language, sometimes involving change in referent or "transvaluation" of the lexeme for *twenty* (below 3.51, 3.534). Other, more exotic types to be found in one or another TB language include QUATERNARY (Boro, Kubhinde Dumi) and DUODECIMAL (Chepang) systems. In the case of Chepang, the system seems to have come into being through a transvaluation of the inherited root for *twenty* into the meaning *twelve* (below 3.535).

Actually the phenomenon of numeral transvaluation is surprisingly widespread in TB, a testimony to the multiple simultaneous conceptual intercon-nections among the numbers themselves (below 4.02).

In sum, this paper is concerned only tangentially with the refinement of the phonological reconstructions of the proto-numerals. At least equal attention will be paid to an appreciation of the internal workings of synchronic TB numeral systems. By studying the morphophonemic and conceptual vicissitudes that the inherited material has undergone in the various languages, we may arrive at something approaching a taxonomy or typology of TB numeral systems.

1.2 Overview of Sino-Tibetan Numeral Systems According to Subgroup

In general, it is the Kamarupan languages — especially the Kuki-Naga and Abor-Miri-Dafla groups — that best illustrate the complex "play of prefixes" with numeral roots (below 1.24, 5.4, 5.5). On the conceptual side, the Himalayish languages are of particular interest, especially because of the hesita-tion between decimality and vigesimality in their higher numerals (below 1.23, 3.53). Yet all the subgroups of the family have their characteristic numerological flavor, and it is worthwhile to do a quick rundown of the various branches, giving a representative specimen of the kinds of numeral systems to

---

pressed as "9 - 1". Below, *loc. cit.*

be found in each.

1.21   Lolo-Burmese and Karenic

These branches of TB have undergone radical simplification of initial consonant groups, and have thus lost most direct traces of prefixes with their numerals. An exception is the voiceless sonorants of Burmese (both in the anciently attested Written Burmese and in the modern dialects), which do directly reflect earlier prefixes, PLB *s- or *ʔ-, as in **hnac** '2', **hrac** '8' (< PLB *s-ni-t and *s-rit, respectively). See Figure 3.

Figure 3   Some Lolo-Burmese Numerals

	WRITTEN BURMESE	*LAHU (C. Loloish)*
ONE	tac	tê
TWO	hnac	nî
THREE	sûm	šêʔ ⋊ šɛ̄ [22]
FOUR	lê	ɔ̂
FIVE	ŋâ	ŋâ
SIX	khrok	khɔ̂ʔ
SEVEN	khu'-hnac	šɨ̄
EIGHT	hrac	hí
NINE	kûi	qɔ̂
TEN	tə-chay [23]	tê-chi
TWENTY	hnə-chay	nî-chi

Another route by which a prefix could survive was by "preempting" or driving out a weak (non-obstruental) root-initial,[24] as in SEVEN *s-nit > Lahu **šɨ̄**, where the root-initial *n- has fallen victim to the sibilant prefix. Another famous example is the Maru (Burmish) word for FOUR, **bìt** (< *b-liy), an isolated instance of the survival of the *b- prefix in Lolo-Burmese.[25]

Karen, like LB, shows no hint of vigesimality in its system of round numbers. Unlike LB, however, many Karenic languages have non-decimal multiplicative/additive formations for the numbers from 5 to 9.[26]   Compare

---

22)   šêʔ is the variant that occurs in counting, while the "etymologically correct" allofam šɛ̄ now appears only before certain classifiers. Morphophonemic alternations in Tibeto-Burman numerals, besides being triggered by classifiers, also typically occur in compound numerals (TEENS and ROUND NUMBERS), similarly to English five ⋊ fif-, or ten ⋊ -teen ⋊ -ty (below, 3.3).

23)   One characteristic type of morphophonemic change in numerals is *destressing* in non-final position in a collocation, as in Burmese TEN and TWENTY. Note that in these languages TEN is a classifier, not a numeral, i.e. '10' is expressed as "one tenworth", '20' as "two tensworth", etc. See below 3.32(B), and the Kayah form for ONE (Fig. 4).

24)   For the first use of the term *prefix preemption* , see Matisoff 1972b.

25)   The development of *-iy > Maru -it (as well as of *-uw > Maru -uk) is regular. See Burling 1968.

26)   Karen dialects mentioned as having such composite numerals in STC (p. 130) include ↗

Figure 4　Some Karen Numerals

	SGAW [Jones 1961]	E. KAYAH [Solnit 1984]
ONE	tɜ̄	tə-
TWO	khĭ[27]	nā̄
THREE	θɜ̄	sō
FOUR	lwì	lwī̄
FIVE	jɛ̀	ŋē ~ ñē
SIX	xȳ	sō swáʔ
SEVEN	nwĭ	sō swáʔ tə-
EIGHT	xɔ̀ʔ	lwī̄ swáʔ
NINE	khwĭ	lwī̄ swáʔ tə-
TEN	ʃĭ	chá ~ chā̄[28]

the decimal Sgaw system with the non-decimal system of Kayah (= Red Karen = Karenni) in Figure 4.

Since other Kayah dialects preserve the monomorphemic forms, Solnit (p.c. 1984) feels that these composite numerals are recent developments, and glosses **swá** as 'double'. (It also occurs in compounds with the meaning 'companion', as in **khō-bɛ́-swá** 'friend', **bɛ́-swá-rá** 'be companions with'.)[29] The Kayah numeral **tə-** 'one' is an always unstressed bound form, which must appear with a following classifier. Syntactically, Kayah SIX and EIGHT are preceded by their classifiers (e.g. **plɔ sō swáʔ** 'six round objects'), while with all the other numerals, including SEVEN and NINE, the classifier must follow (**sō swá tə-plɔ** 'seven round objects').

Other Karen dialects, especially Pa-O (Taungthu) have developed secondary dental suffixes with certain numerals: Pa-O **lĭt** '4', **ngāt** '5', **kŭt** '9'. In the case of **nŭt** '7' and **sɔ̄t** '8', the PTB forms themselves are reconstructed with final dentals (*-s and *-t respectively), but since Karen does not generally preserve final consonants, the -t in these forms also appears to be secondary.[30] We consider these final dentals to constitute a "suffix run", one of the many manifestations of the interinfluence of consecutive numerals (below 4.01). As we shall soon see [1.23], numeral suffixes are also characteristic of Himalayish, but there they tend to be fully syllabic.

↘ "White Karen, Bwe, Brek, Red Karen, Yintale, and Mano."

27) Note the preemption of the root-initial by the velar prefix, *g-nis > khĭ.

28) The rising-toned variant is basic, while the mid-tone occurs in the round numbers 20-90. As Solnit observes, this tonal difference has a practical disambiguating function. Compare, e.g. **chá sō swá** '16', i.e. $10+(3\times2)$, where TEN is in an *additive* relationship to the following numeral, with **chā̄ sō swá** '60', i.e. $10\times(3\times2)$, where TEN stands in a *multiplicative* relationship with it.

29) It seems likely that this morpheme is ultimately related to Chinese (Mand. **shuāng**) 'pair', which also underlies the Thai numeral **sɔ̌ɔŋ** 'two.'

30) See STC, p. 131, and Benedict 1979, pp. 18-20. For more discussion see below 4.223.

1.22   Kachin-Nung and Qiangic

Jingpho (= Kachin) has a lively proliferation of prefixal morphology, some of which is exploited for specific semantic ends. The negative morpheme *ma has been reduced to a syllabic nasal prefix, ṇ-. The old causative prefix *s- has been preserved and generalized as šə- ~ dža-. Younger strata of prefixation are much in evidence, with a tendency to create fully syllabic prefixes out of sub-syllabic ones, e.g. *m-raŋ 'horse' > Jg. gùm-ràŋ. With respect to numeral prefixation, Jingpho is relatively conservative, preserving the proto-prefixes rather well, though it does have a secondary "prefix run" from THREE to FIVE (see Fig. 5).

The Nungish languages seem generally quite close to Jingpho. However, unlike the sesquisyllabic Jingpho, which abounds in words beginning with prefixal "minor syllables" of the form Cə-[31], Nungish is strictly monosyllabic, so that only an occasional prefix survives before a non-obstruental root initial, as in Nusu (Central Nung: Sun and Liu 1986) vɹi³⁵ < *b-ləy 'four'.

Some Qiangic languages (the newly articulated branch of TB spoken in Sichuan) have complex initial consonant groups, often of demonstrably secondary origin. The Qiangic language with the most elaborate numeral prefixes seems to be Ergong [SUN 1985]. See Figure 5.

Note the impressive run of the prefix w- (< *b-) in the Ergong numerals from 2 to 6, even longer than the Jingpho run of mə- in 3-5. As mentioned above (n. 20), the Jingpho forms for ONE and TWO are innovations which require a special explanation.

Figure 5   Kachin-Nung and Qiangic Numerals

	JINGPHO	NUSU	ERGONG
ONE	ləŋâi	thi⁵³	ʐau
TWO	ləkhôŋ	m̩⁵⁵	wnɛ
THREE	məsūm	sɔ³⁵	wsu
FOUR	məlī	vɹi³⁵	wʐɛ
FIVE	məŋā	ŋɑ⁵⁵	wŋuɛ
SIX	krúʔ	khɹu⁵³	wtɕhau
SEVEN	sənìt	ŋɤ̃⁵⁵	sŋie
EIGHT	mətsát	ʂɑ⁵³	ɣie
NINE	jəkhû	guɯ³⁵	ŋgɛ
TEN	šī	tshe³⁵	zʁa/sqha

1.23   Himalayish and rGyarong

Himalayish shows fairly good preservation of the proto-prefixes, but by

---

31)   The term *sesquisyllabic*, referring to words "a syllable and a half" long, was introduced in Matisoff 1973b.

and large little innovation of secondary ones, so that "prefix runs" in the numerals are rare. The languages show variation and vacillation between decimality and vigesimality (below 3.534). In the case of many of the minority TB languages of Nepal, the higher native numerals are rapidly giving way to Indo-Aryan replacements from Nepali (below 2.1).

The numeral prefixes of Written Tibetan [WT] are taken (perhaps too uncritically) by STC as faithfully reflecting the most ancient stratum of prefixation in TB. In any case, WT is certainly much more conservative in this respect than younger Himalayish languages like, e.g. Tamang (Nepal) or Kanauri (Simla Hill States, Punjab), which only show prefixes with a few of the numerals. See Figure 6a.

Figure 6a　Some Himalayish Numerals

	WRITTEN TIBETAN	*KANAURI*	*TAMANG*
ONE	gćig	id	ki:h
TWO	gńis	niʃ	ɲi:h
THREE	gsum	ʃum	som
FOUR	bźi	pū	plih
FIVE	lŋa	ŋa	ŋa:h
SIX	drug	ʈuk	ʈu:h
SEVEN	bdun	stiʃ	ɲis
EIGHT	brgyad	rai	preht
NINE	dgu	zgui	ku
TEN	bću	sai	ci

Among the more prefixally innovative Himalayish languages is Lepcha (Sikkim), which not only preserves the "proto 4-5 run" as fə-, but has also innovated a kə- prefix for 7-10.[32] Even more exuberant in this respect is rGyarong (= Jiarong), which for several numerals not only retains the proto-prefix but adds a new one in front of it, creating a long velar-prefix run from 2 to 7. This is still another indication that rGyarong is not "core Himalayish" at all, but rather a transitional language, with suggested affinities to Abor-Miri-Dafla [see NAGANO 1984],[33] and/or (as maintained by Sun Hongkai 1985) with Qiangic.[34] See Figure 6b.

It is characteristic of many languages of Nepal to have *suffixes* attached to their numerals, e.g. Dumi -po (**tūk-po '1', sak-po '2', suk-po '3'**), Bantawa -pok (**ūk-pok '1', hūa-pok '2', sum-ka-pok '3'**), Yakkhaba -ci/-ji (**nic-ci '2',**

---

32) As we shall see (below 4.02,4.23,4.24), these Lepcha forms for EIGHT and NINE seem to have undergone an "etymological flipflop".

33) The possibility of a special AMD-rGyarong relationship is vigorously criticized in J. T. Sun 1993:379-389.

34) This rGyarong run is reminiscent of the 2-6 run of the w- prefix in the Qiangic language Ergong (above 1.22), though Ergong only shows one prefix per numeral.

Figure 6b   Himalayish Innovators: Lepcha and rGyarong

	LEPCHA	rGYARONG (Zida dialect)
ONE	kat	tšek
TWO	ɲət	kenes
THREE	sam	kesom ✕ kesam
FOUR	fəli	kewdži
FIVE	fəŋo	kemŋa
SIX	tərək	keʈa
SEVEN	kəkyək	keṣñit ✕ keṣñis ✕ keṣñes
EIGHT	kəku	warže(t)
NINE	kəkyot	keŋgu
TEN	kəti	ṣʈṣi

sum-ji '3', ri-ji '4'), etc. [GVOZDANOVIĆ 1985: 135-136]. These suffixes are fully syllabic (unlike those of Pa-O Karen, above 1.21), so one may surmise they are (or once were) classifiers, or even gender markers, rather than meaningless formatives. See below 2.1.

1.24 Kuki-Chin-Naga and Bodo-Garo

KCN shows good preservation of the proto-prefixes, but also a strong tendency toward innovative prefix runs. This is the branch of TB whose numeral prefixal behavior will be discussed in the most detail (below 5.4). Like Himalayish, Kuki-Chin-Naga shows a complex interplay of decimal and vigesimal characteristics (below 3.52).

Bodo-Garo (= Barish) displays occasional cases of reprefixation (e.g. Garo ge-gni '2'), but in general is not so extreme in this respect as KCN, Qiang, or rGyarong. BORO can definitely be shown to have a *quaternary* or 4-based numeral system, very unusual for TB.[35]

1.25 Abor-Miri-Dafla

This relatively obscure branch of TB harbors some of the strangest numeral systems of all from a comparative viewpoint, especially with respect to the "higher numerals" (7, 8, 9). Not only do we find roots that are hard to relate to anything else in TB, but the systems also reveal peculiarities of internal structure (e.g., "multiplicative" forms for EIGHT: see below 4.20, 4.237). Several new roots for ONE and TEN have been unearthed in this subgroup (below 3.15, 3.23).

The numeral prefixes that appear with the highest frequency in AMD consist of a vowel alone: **a-, o-, e-**. All other prefixes (e.g. **kV-, pV-, ra-**) are quite

---

35) See below 3.32(C) "Teen formation in Barish" and 4.201 "Multiplicative phenomena." Elsewhere in TB, the closest thing I have found to the Boro quaternary system is the *duodecimal* system of Chepang (below 3.535).

rare with AMD numerals.　See below 5.5.

For some indication of the bizarre appearance of some AMD numeral systems, consider those cited in Figure 7.

Figure 7　Some Aberrant Abor-Miri-Dafla Numeral Systems[36]

	AKA	MIJU	MILANG	SERDUKPEN
ONE	a	kumo	akan; atel	han
TWO	kshi	kinin	ne	n(y)ik
THREE	zu	ksam	ham	ung
FOUR	fi-ri	kambran	_pe	bi:si
FIVE	phum	klin	pangu	khu[37]
SIX	rieh	katam	sap	khit
SEVEN	mulh	nin	rangal	sit
EIGHT	sikzi	grin	rayeng	sargiat
NINE	sthö	natmo	kanyem	dikhi
TEN	rhi	kyapmo	hangtak	dokche

1.26　Chinese

Evidence for pre-Archaic Chinese prefixes is of course indirect, but it looks as if there may have been a run of the *s- prefix in the numerals from FOUR to SEVEN.[38]　See Figure 8.

Figure 8　Old Chinese and PTB Numerals

	OLD (= "ARCHAIC") CHINESE	PROTO-TIBETO-BURMAN
ONE	*ʔi̯ĕt [GSR #394]	*it
	*tśi̯äk [GSR #1260]	*g-t(y)ik
TWO	*ni̯ər [GSR #564]	*g-ni-s
THREE	*tsˈəm ~ *səm [GSR #647, 648]	*g-sum
FOUR	*si̯əd[39] [GSR #518]	*b-ləy
FIVE	*ngo[40] [GSR #58]	*l-/b-ŋa
SIX	*li̯ôk[41] [GSR #1032]	*d-ruk
SEVEN	*tsˈi̯ĕt [GSR #400]	*s-nis
EIGHT	*pwăt[42] [GSR #281]	*b-g-ryat/*b-r-gyat
NINE	*ki̯ug > MC *ki̯əw [GSR #992]	*d-kəw
TEN	*d̂i̯əp [GSR #686]	*g(y)ip
HUNDRED	*păk[43] [GSR #781]	*r-gya [but WT brgya]

---

36)　Aka/Hruso from LSI III.1, Miju from Das Gupta 1977a; Milang from Das Gupta 1980; Serdukpen from Dubey 1983.

37)　See below 4.217.

38)　Cf. the section "Chinese numerals" in STC, pp. 161-2.　Special studies have been devoted to the Chinese numerals ONE [BOLTZ 1969], TWO [BOLTZ 1977], and SIX [MEI and NORMAN 1968].

39)　STC derives this from pre-Archaic *p-səy.

2. LANGUAGE CONTACT AND THE WEIGHT OF NUMBERS

The numeral systems of majority languages may easily make profound incursions into those of less prestigious minority languages. Numbers prevail — a numerically dominant population will "make its numbers felt" in more ways than one! Differential numerical prestige is dramatically illustrated, e.g. in market situations, where speakers of minority languages come to town and have to bargain using the foreign numerals of the majority population.

The embattled indigenous languages of the Malay peninsula, belonging to the "Aslian" branch of Mon-Khmer, are a good case in point: "Mon-Khmer languages of Malaya, with the exception of Semelai and Semoq Beri, have not retained a complete set of Mon-Khmer numerals, but, above the numbers three or four, use Malay borrowings" [DIFFLOTH 1976:31].

Similarly, various Tai languages have exerted a decisive influence on the numerals of co-territorial TB languages, especially those spoken by very small populations. In Hpun, a moribund Burmish language of Kachin State, Henderson reports that "there was great uncertainty and much dispute among his informants over the numerals above three. Luce supposes that since the local bazaars are mostly run by Shans, Shan numerals have replaced the Hpun ones in general use."[44] In Bisu, a Southern Loloish language spoken in a few villages of Thailand, the original TB numerals 1-5 are still current, but above five only loans from Thai are found: '6' **hɔ̃k**, '7' **kjīt**, '8' **pɛt**, '9' **kãw**, '10' **sīp**.[45]

Going a step further up the totem pole of relative prestige, the Tai languages themselves have long ago replaced all their native numerals from 2-10 by Chinese ones.[46] In fact, the overwhelming influence of the Chinese numerals has been felt throughout the "Sinosphere", including Japanese, Korean, Vietnamese, Miao-Yao, and a number of the TB languages of China.

---

40) STC cites the very early loan into Proto-Tai, **ha*, as evidence for Pre-OC **hŋa* (ultimately < ****s-ŋa**).

41) STC posits a pre-Archaic **b-* prefix for SIX on *xie-sheng* evidence. Proto-Tai **hrok* also points to some sort of prefix in pre-OC, but not necessarily in my view to a labial prefix. **s-* in fact seems more likely (cf. Tho **sok**), though Benedict claims that Ong-Be **sok** points to **phr-* ("a regular shift").

42) STC (pp. 162 ,179) derives this from pre-Archaic **b-ryat* < **bryât*.

43) STC ingeniously but ad-hoc'ly derives this OC form "from **pak*(-rya) [with typical unvoicing of the prefix, then restressing of the prefixal vowel] < **b-grya* < **b-r-gya*."

44) Henderson 1986, p. 112.

45) See Nishida 1966/1967.

46) In Matisoff [in prep.], I suggest the term "Sinonumeric" to refer to those Tai-Kadai languages that have borrowed the Chinese numerals en masse. The inherited Austro-Tai numerals are only preserved in a few obscure "outlier Kadai" languages like Hlai (Hainan). For the introduction of the terms "Sinosphere" and "Indosphere", see Matisoff 1990a, 1991.

The same phenomenon is apparent in the "Indosphere" as well. Emeneau [1957] reports the massive influence of Indo-Aryan on the Dravidian numerals. Closer to home, the numerals of the Kamarupan and Himalayish branches of TB have undergone some influence from Indo-Aryan (Bengali, Assamese, Kashmiri, Hindi) — though the most dramatic inroads have been made by Nepali on the TB languages of Nepal. As we shall see (below 4.02), foreign incursions into a language's numeral system can lead to widespread *transvaluations*, or reinterpretations of the meaning of the individual elements in the system.

2.1　Nepali and the TB Languages of Nepal

Nepali is a member of the northern group of Indo-Aryan languages. Its numerals are displayed in Figure 9.

Figure 9　Nepali Numerals

ONE	ek	SIX	cha
TWO	dui	SEVEN	sa:ţh
THREE	tin	EIGHT	a:ţh
FOUR	ca:r	NINE	nau
FIVE	pa:nc	TEN	das
TWENTY	bis	HUNDRED	se
		THOUSAND	hajaar

The TB languages of Nepal are no exception to the principle that the lower a numeral is, the more likely it is to resist change.[47] Many languages (e.g. KHAM, SUNWAR, CHOURASE, MEWAHANG, ATHPARE) preserve only the TB numerals 1-3; MAGARI retains 1-4; CHEPANG and LOHORONG go so far as to keep 1-5. All other numerals in these languages are from Nepali, or else derived from extraneous morphemes meaning 'finger' or 'hand'. See Figure 10.

Similarly, Michailovsky [1981] reports that in HAYU (= Vayu), a language now on its last legs, "à partir de *cinq* (*quatre* pour la plupart des locuteurs) les numéraux et classificateurs nepali sont employés". Speaking in almost identical terms of the situation in Thulung Rai, Allen [1975] notes that "no Thulung that I met knew how to count in Thulung beyond four (many could only reach three)."

What accounts for the relative hardiness of the lower numerals? Gvozdanović attempts an explanation in terms of grammatical function, claiming that "the process of numeral decay is at each stage characterized by a language-specific cut-off point, defined by the highest numeral which is actively used in

---

47)　See above, 1.12, 2.0. For the exceptional situation in Jingpho in this regard, see note 20 and Figure 5.

Figure 10 Preserved TB Numerals in Some Languages of Nepal[48]

	KHAM	SUNWAR	MAGARI	CHEPANG
ONE	tobo	ka:	kat[49]	ya:t-jo?
TWO	nehplo	niksyi	nis	nis-jo?
THREE	sohmlo	sã:	som	sum-jo?
FOUR			puli[50]	play-jo?
FIVE				poŋa:-jo?

	CHOURASE	MEWAHANG	ATHPARE	LOHORONG
ONE	kolo/kwalo	ekku	ʈhik	ʈhikko
TWO	nimpha	hicci	ippok	ŋicci
THREE	summakha	sumji	sumbok	sumci
FOUR				ricci
FIVE				ŋaci

numeral constructions" [1985:140]. This is merely circular, however, since it amounts to saying that only the native numerals which are preserved are available to participate in native numeral constructions! Whether a language will preserve its original numerals only for 1 and 2, or whether it will keep 3, 4, and/or 5 as well is certainly not predictable from any independent grammatical. parameter (e.g. whether the language maintains a category of dual in its pronouns and verbs). The staying power of the lower numerals is best appreciated in a more common-sensical way. It is the lower numerals which have the highest real-life (pragmatic) frequency and saliency — things in the world come in two's and three's much more often than they do in seven's and eight's. Children learning their native language will have a clear conception of TWO and THREE long before they have the higher numbers figured out.[51] The lower numerals are apt to appear in many more idioms and collocations (set expressions) than the higher ones, which contributes to their survival value.[52]/[53]

48)   Data from the first four languages in the chart are from Hale (ed.), 1973 ("CSDPN"); forms in the other languages are from Gvozdanović 1985. Chourase **kollabremci '5'** and **nimphalabremci '10'** are derived from **la** 'hand' and **brem** 'finger'; Mewahang **ihuk '5'** (and perhaps **hukhu '10'**) are derivates of **huk** 'hand'. For the widespread association between FIVE and HAND, see below 4.15.

49)   I believe Magari **kat** and Chepang **ya:t** to be cognate, both descending from a PTB etymon *k-y-at that underlies two supposedly independent roots set up in STC, viz. *kat and *it. See the discussion of words for ONE, below 3.1.

50)   There is no trace of a velar prefix in TWO or THREE, but the labial prefix is preserved in Magari FOUR and Chepang FOUR and FIVE.

51)   For young children, big numbers are mysterious undifferentiated jumbles, so many "forty-levens'es."

52)   We do have a few idioms in English involving higher numerals (*at sixes and sevens, six of one and half a dozen of the other, a stitch in time saves nine*, etc.), but none of them are likely to be acquired by a child at an early age - certainly not until long after he has learned things like "*1, 2, 3 — go!*"

53)   Analogously, expressions like *give him an inch and he'll take a mile* will survive long after the English-speaking world converts completely to the metric system.

Figure 11 The Moribund TB-derived Numerals of Thulung Rai

ONE	ko	SIX	ru
TWO	nə	SEVEN	yet
THREE	sium	EIGHT	let[54]
FOUR	blə	NINE	gu
FIVE	ŋo	TEN	kodium

Irregularities and suppletions are quite tolerable with the high-frequency lower numerals, but tend to be quickly levelled out with the lower-frequency higher ones: we can readily accept the irregular ordinals *first* and *second*, since we have learned them by rote at such an early age, but we would not like it so much if it were, e.g. EIGHT and NINE that had irregular ordinal forms while the others were predictable from the corresponding cardinals.

In any event, loss and replacement of numerals can occur much more rapidly than a language's grammatical categories change. We have seen that by 1975 no speaker of Thulung Rai knew the TB numerals above FOUR. Yet Allen notes (pp. 102-103) that in a vocabulary compiled by Agami Singh Rai only 30 years before [1944], a full set of TB-derived Thulung numerals is given, including those in Figure 11.

In the case of Hayu, we can trace the breakdown of the traditional numeral system through a period of over a hundred years. It is interesting to compare the surviving TB numerals in Michailovsky's data [1981] with the forms to be found in Hodgson's 19th century material. See Figure 12.

What Hodgson found was considerably more elaborate than the vestigial system reported by Michailovsky after 120 more years of intense pressure from

Figure 12 Surviving TB Numerals in Hayu

(a) Michailovsky [1981:167]

ONE	kolu
TWO	nakpu (human)/naʔung (non-human)
THREE	tshukpu (human)/tshuʔung (non-human)
FOUR	b(l)iʔung

(b) Hodgson [ca. 1860, cited in LSI III.1:384.][55]

	Masculine	*Feminine*	*"Irrational"*
ONE	kom-pu/kwong-pu	kwo-mi/kwong-mi	ko-lu
TWO	naːk-pu	naːng-mi	naː-yung
THREE	chhuk-pu	chhung-mi	chhu-yung
FOUR	blik-pu	blig-mi	bli-ning

---

54) Note the convergence of the rhymes in SEVEN and EIGHT.
55) LSI also cites forms for FIVE and SIX, which appear, however, to be multiplicative in origin (below 4.201).

Nepali. In fact, the 3-way gender distinction in Hodgson's data furnishes a possible clue as to the original function of the *suffixes* which are such a characteristic feature of Himalayish numeral systems (above 1.23).

Several suffixes like these, which may once have been gender markers and/or classifiers, are still attested in more than one TB language of Nepal:

(a) **-lo/-lu.** Besides HAYU **ko-lu** '1', cf. CHOURASE **kolo/kwalo** '1'; KHAM **nehplo** '2' [with epenthetic **-p-**?] and **sohmlo** '3'; and YAK-KHA **kolok** '1' (with **-k** suffix).

(b) **-pu/-po.** Besides HAYU **nakpu** '2', **tshukpu** '3', **blikpu** '4', cf. KHAM **tobo** '1'; SUNWAR **saːhpu** '2' and **suhpu** '3'; and especially the Saptesar dialect of DUMI, which has generalized the **-po** with all the numerals from 1 to 9 (**tɨkpo** '1', **sakpo** '2' ... **ɔmpo** '8', **rekpo** '9'.[56])

(c) **-pok/-bok.** Possibly related to the previous suffix is a form with velar final that occurs in Athpare **ippok** '2', **sumbok** '3', and in some dialects of Bantawa (**ɨkpok** '1', **hɨapok** '2', **sumkapok** '3', **reʈkapok** '4').[57]

(d) **-ci/-ji.** Several languages have this numeral suffix, including Mewahang **hicci** '2', **sumji** '3'; Yakkha **hitci** '2', **sumci** '3'; and Lohorong **ŋicci** '2', **sumci** '3', **ricci** '4', **ŋaci** '5'. Sometimes it is found generalized with the whole set of numerals from 2-9 or 2-10, as in Yakkhaba (**nicci** '2' ... **nɔkci** '9') and Kulung (**nicci** '2' ... **nuci** '9', **bɔci** '10').

(e) **-si/-shi.** This suffix, which may well be etymologically related to the previous one, is found in Bahing **niksi** '2' and Sunwar **niikshi** '2'. In Limbu it has been generalized with all the numerals from 2-8 (**netsshi** '2', **sumsi** '3', **liisi** '4', **n(g)aasi** '5', **tuksi** '6', **nuusi** '7', **phangsi** '8'. There may also be an allofamic relationship with a velar-finalled suffix **-tsing** found in Sharchop and Monpa (Motuo) **ɲik-tsing** '2'.

Other suffixes, e.g. Chepang **-joʔ** (Fig. 10) and Hayu **-ʔung**, remain a mystery in comparative terms. The old Hayu feminine suffix **-mi**, however, is relatable to a general TB root *****mi(y)** 'woman, female' (e.g. Lahu **yâ-mî** 'daughter', **ɔ̂-mî-ma** 'wife').

The generalization of a particular suffix to a succession of adjacent numerals may be referred to as a "suffix run".[58]

---

56)   The Kubhinde dialect of Dumi also uses **-pu** with all its TB-derived numerals, though only 1-4 survive in this dialect: **tɔkpu** '1', **sɔkpu** '2', **bhlɔkpu** '3', **rɔkpu** '4'. As noted below [4.02] the Kubhinde words for '2' and '3' have been "transvalued" from their original meanings of '3' and '4', respectively.

57)   Other Bantawa dialects have a suffix with retroflex **ʈ**, e.g. **ɨkʈa(k)** '1', **hɨaʈa** '2', **sumkaʈ** '3', **reʈkaʈaʈ** '4'. See Gvozdanović, p. 155.

58)   See above 1.21, with respect to the non-syllabic dental-suffix run in Pa-O Karen.

3　ONE AND TEN AND TEENS AND TWENTIES

3.1　Profile of Number ONE

As STC observes (p. 94) there is no single general root for ONE or TEN in Tibeto-Burman, in sharp contradistinction to the "primary" numerals 2-9, for each of which a single etymon overwhelmingly predominates. The special importance of the concept ONE links it to many other semantic fields. As the most frequently occurring numeral, its constant use may lead to its semantic bleaching, until it becomes an indefinite article. Its high frequency encourages morphophonemic irregularity, and idiosyncratic fusions with other morphemes. (Cf. the multiple English allofams which all descend somehow from PIE *oino-: *one, an, once, only, alone, anon, onion, eleven* < ME **ellevene** < OE **endleofan** < *ain-lif-* "one left [beyond ten]".)

Sometimes a language maintains more than one ONE, one of which occurs as the independent numeral while the other survives only as a part of compound numerals, e.g. Thai **nỳŋ** '1', **sìp** '10', but **sìp-èt** '11', **rɔ́ɔj-èt** '101'. This -**èt**, like the rest of the Thai numerals from 2-9, is of Sino-Tibetan origin [below 3.11].

In Garo, three separate etyma for ONE have been preserved, each frozen into the numeral system in its own restricted context: **sa** '1' (independent), **chi-sa** '11', **ritcha-sa** '100'; but **kol-*grik*** '20' ("20×1"; below 3.14); and **chi-*kung*** '10' (presumably "10×1"; below 3.13b).

Many languages have an unrelated ("suppletive") form for the ordinal corresponding to ONE, e.g. Eng. *first*. This study does not deal with words like *first* or *single*, since they frequently come from unpredictable non-numerical semantic fields.

3.11　*it ✕ *yat

STC (p. 94) sets up a PTB etymon *it on the basis of only two forms, Kanawari **id** and WB **ac**, identifying it as cognate to Old Chinese *ʔiĕt (p. 162).

To these I would now like to add Chepang **yaːt(-joʔ)**, which agrees well with Chinese, and leads me to revise the PTB (and PST) reconstruction to *it ✕ *yat. I am thus claiming that this root displays the -i- ✕ -ya- variational pattern established independently for a number of non-numerical roots (e.g. EYE, PHEASANT[59]), and, strikingly enough, with several other numerical etyma as well, as we shall see.[60]

Several TB languages of Nepal have disyllabic forms for ONE where the first syllable has a superficial resemblance to the above forms, but these all seem

---

59) See Matisoff 1978a ("VSTB"), pp. 40-41.
60) Cf. *tik ✕ *tyak '1' (below 3.14), *gip ✕ *gyap '10' (below 3.21), and perhaps *ring ✕ *ryang '10' (below 3.233).

to be borrowings from Nepali **ek** (see Fig. 9):

Mehawang   *ek*-**ku**[61]
Yakkhaba   *ik*-**ko**
Kulung   *i*-**bum** ∼ *i*-**bim**
Bantawa   *ʉk*- (as in *ʉk*-**ʈai**, *ʉk*-**ʈak**, *ʉk*-**pok**, *ʉk*-**ʈa**, all meaning 'ONE' in various dialects; see Gvozdanović 1985: 188)

We are now able to relate the root **it* ⪥ **yat* to another set of forms that STC sets up as an independent etymon, **kat* (next section).

3.12   **k-(y)at* ⪥ **k-(y)it* ⪥ ⪥ **k-yan* ⪥ **k-(y)in*

STC laconically sets up a PTB root **kat* on the basis of "Lepcha **kat** and Kuki-Naga **khat*" (p. 94). More specifically, we may cite the following forms from Kuki-Naga languages:

Zeme and Zeliang (**hang**)**kat**; Kom Rem **inkhat**; Lushai, Hmar, and Vaiphei **pakhat**; Gangte, Khoirao, Maring, Paite, and Puiron **khat**; Liangmai **khad**; Thado **xʌt**; Nruanghmei **khüt**.

To the Lepcha form, we may add another cognate from a Himalayish language, Magari **kaʈ** (with unexplained retroflex ʈ).

3.121   **kya-n* ⪥ **kya-t*

Many other Kamarupan forms with front vowels may reflect a medial -y-: Mzieme **ket**; Sangtam **khe** (also **khürü**); Pochury **khe**; Meluri **ke** (also **kesü**); Sema **khe** (also **laki**); Mishmi [DUBEY] **khege**; Chulikata **eːkheː**; Idu **khe-ge** (also **kheng-ge**).

Other Abor-Miri-Dafla languages have a final nasal after the front vowel (note the variation in Idu):

Idu **kheng-ge** (also **khe-ge**); Gallong **aken** (also **ako**); Lhopa **aken** (also **ako**); Padam **akem** (also **atel**); Tagin **akin**; Dafla **akkin** (E. Dafla, HAMILTON 1900), **aking** (also **aku**) [DAS GUPTA 1969], **akhin** (Yano Dafla, BOR 1938), **aːkin** [ROBINSON 1851]; Taraon (**eː-**)**khing** (Digaru Mishmi, LSI 3.1, 623).

I would like to relate all these forms in a word family like **k-(y)at* ⪥ **k-(y) it** ⪥ ⪥ **k-yan* ⪥ **k-(y)in*, showing variation both between -i- and -ya-, and between final homorganic stop and nasal. Also to be accommodated here are the two forms cited in STC #34: WT **rkyaŋ-pa**, WB **khyâŋ** 'single'.

A similar variational pattern in TWO is suggested by the Lepcha doublet ɲi ⪥ ɲăt. See below 4.114.

---

61)   For the second element in these apparently pleonastic Mewahang and Yakkhaba formations, see 3.13 below.

3.13　*ka and *ko

The forms in these groups, with velar initial and non-front vowel, may or may not be etymologically related to the forms cited in 3.11 and 3.12. Several Kamarupan languages have two velar-initial words for ONE, one with a back vowel and the other with a front one (e.g. Gallong/Lhopa **aken** and **ako**).

(a)　With **a**-vocalism and no overt trace of a final consonant:

Ao **ka** (Chungli), **akha** (Mongsen); Lotha **ekha**; Lakher **mia-kha** (also **sa-**, below 3.152); Yacham-Tengsa **kha-tu** (for second syllable see below 3.143); Tangkhul **akha, khatkha**;[62] Sunwar (Himalayish) **ka:-**.

(b)　With **-o** or **-u** vocalism and no following nasal element:

Abor-Miri-Dafla

Abor-Miri **a-ko** 'one', **-ko** 'general numeral suffix'; Lhopa **a-ko** (also **aken**); Gallong **a-ko** (also **aken**); Dafla **aku** (also **akkin, aking**); Miju **-ko** 'one; -teen' (see below 3.4).

Himalayish

Thulung Rai **ko** '1', **ko-** '-teen' (see below 3.4); Newari **-gu:** 'general numeral suffix'; Hayu *ko*-**lu** '1'; Yakkha *ko*-**lok**, Chourase *ko*-**lo**, *kwa*-**lo**; Mewahang **ek-**ku, Yakkhaba **ik-**ko; Lohorong **ṭhik-**ko (all meaning ONE).

The first syllables in the Mewahang and Yakkhaba forms seem to be loans from Nepali **ek** (above 3.11); if the second syllables also mean ONE, these are redundant or pleonastic formations (as in Lohorong, where the first syllable descends from another native root for ONE [below 3.14]). In these languages the second syllables have evidently been bleached to suffixal status, as in Abor-Miri or Newari, devoid of anything but a weak meaning like 'unit' (below 3.16).

(c)　With non-front vowel and following nasal element:

Abor-Miri-Dafla

Milang **akan**; Minyong **akon** (also **atîr, ayirr**); Darang Deng **k'ɯn**[55]; Apatani **kun** (nonhumans), **kon** (humans); Idu **khɯn**[55] [SUN 1983:69].[63]

Other TB forms which seem to belong here are Bahing (Himalayish) **kong** '1', and Garo (Bodo-Garo) **chi-**kung '10' (lit. "10×1"). Michailovsky (p.c. 1995) sets up *koŋ for W. Kiranti, based on Bahing and Hayu **koŋ**, Sunwar **ka:**, and Thulung **ko(ŋ)**.

3.14　*g-t(y)i-k ≍ *tya-k and *d/tay ≍ *d/tan

STC reconstructs an etymon *g-t(y)ik ≍ *tyak 'one' on the basis of WB **tac**, Nung **thi**, and a group of forms from Himalayish (WT **gcig**, Chingtang

---

62)　The Tangkhul variant **khatkha** is hard to evaluate in the light of our present knowledge. Is it reduplicative? Or does each syllable represent a quite separate etymon, *ka vs. *kat? Or is the **-t** a suffix (*ka-t)?

63)　J.T. Sun has now reconstructed a Proto-Tani root *kon, on the basis of forms he cites as Apatani **kū** ≍ **kõ**, Bengni **a-kin**, Bokar **a-ken**, and Padam-Mising **a-kon** [1993:183].

thit-ta, and "Rai" **tik-pu**).[64] To these may be added Dumi **tɪk-po, tɔk-pu**; Lohorong **ʈhik-ko** (for the second element see above 3.13); Athpare **ʈhik**; Limbu **lot-*thik*** 'only one'; Dzongkha **ci**; Kaike **ʈi**; and the second element of Sikkim Bhutia **khe-*chik*** '20', lit. "20 × 1".

Allofamic variations involve all parts of this etymon:

(a) A velar prefix is reconstructed on the basis of WT and other Himalayish languages (e.g. rGyarong **kətɛk** [ZMYYC #911]), but other prefixes are attested elsewhere (e.g. Qiang **petʂʰi** [WEN YU 1950]).

(b) The root-initial consonant shows hesitation between a dental stop and a palatal affricate (natural enough before a high front vowel), both at the proto-level (compare WT **gcig** and WB **tac**) and at much more recent time-depths (e.g. in Nungish, where Rawang has **hti** (= **thi**) [BARNARD 1934] but a Nujiang dialect has **tɕi⁵⁵**).

(c) The Himalayish languages of the Tamang-Gurung-Thakali nucleus seem to point to medial **-r-* rather than **-y-*: Gurung **grihq**, Tamang **kiːh** (with preemption), Thakali **ʈih**.

(d) The vocalic nucleus also shows proto-variation between **-i-* and **-ya-*, a mysterious property of several other numerical roots as well (above 3.11). The variant with **-ya-* vocalism, ****tya-k*** [STC, n.271, p. 94], is reflected by Chinese 隻 ****tśiäk*** 'one, single' [GSR #1260c], to which we may add a number of putative TB cognates: Bumthang **thek, tek** [NISHI 1982]; Monpa (Cuona) **t'eʔ⁵⁴** [SUN et al. 1980]; Bai **tia** [DELL 1981:61].[65] Several forms with affricate initials probably also descend from the allofam **-ya-*, with no direct reflex of a final stop: Newari **cha**; Konyak Naga **ja**; Chang Naga **chie**.[66]

(e) Many daughter languages show no trace of an original final stop. Sometimes this is undoubtedly the regular fate of the **-ik* rhyme, but often (e.g. in Lolo-Burmese) we are forced to recognize a proto-variant with no final consonant. If we indicate this in our reconstruction by putting a hyphen before the **-k*, the resultant **t(y)i-k* then looks a lot like one of the main TB roots for TEN **ts(y)iy* ✕ **tsyay* (below 3.22), a resemblance that may be more than accidental. (See below 3.4, "Inter-

---

64) See STC, pp. 84, 94, 169, 189. It is amusing to note that J. Greenberg [1987:112] has seized upon this reconstructed PST root as a good candidate for his "Proto-World" or "Proto-Sapiens" lexicon, claiming it is genetically related to (among others) Proto-Indo-European **deik-* 'to point', Amerindian forms like Karok **tiːk** 'hand; finger', Yagua **tiki** 'one', and Eskimo **tik-iq** 'index finger', as well as Nilo-Saharan forms like Maba **tek**, Fur **dik** 'one'. For a critique of Greenberg's unfettered approach to linguistic comparison, see Matisoff 1990a.

65) Other possible reflexes of **tyak* in languages of Nagaland are Phom **hük**, and the 2nd syllable of Wancho **tu-ta**. For the first syllable of this Wancho form, see below.

66) French [1983:529] sets up a Proto-Northern-Naga root **-kla*, to which he assigns the Konyak and Wancho forms, as well as the second syllable of Yogli **śa kha** '100' ("100 × 1"), suggesting a connection with Jg. **maʔ⁵⁵khra³¹** 'all; whole' [DAI et al. 1983:455]. An obvious alternative source for the Yogli syllable would be **ka* (above 3.13).

change and confusion between ONE and TEN".)

Many Loloish languages have forms meaning 'one' or 'only' with dental stop or palatal affricate initials and high front vowels, but microlinguistic work reveals a confusing array of variants already at the PLB stage. Some modern forms reflect final *stops, others do not. The vocalism appears to vary among *-i-, *-ay, and *-ey, suggesting that this etymon was often unstressed and hence of unstable vowel quality.[67] Any given language is likely to have developed several co-existent variants (much like English *one, an, only,* etc.; see above 3.1), e.g. Lahu **tê** 'one', **dê-dê** 'all', **tí** 'only', *tèʔ*-**chí** 'nothing', **a-cí** 'little bit' < PLB *day, *nday, *ʔdik, *dek, and *ʔgyik, respectively.[68]/[69]

It is actually far from certain that Lahu **tê** 'one; whole; a/an' and **dê-dê** 'all' are relatable at all to the other forms in the group just cited. As explained in GSTC #148, they are more plausibly to be derived from a newly reconstructed PST root *day ⋈ *tay, underlying forms like Jingpho **tai** 'single', **ətai** 'one, as of a pair', **guntai** 'single', **shingtai** 'only'; Boro **otay** 'whole'; and Lakher **dei** 'only, alone'.[70] The affinities of this etymon seem to lie not with the *tyik family, but rather with the nasal-finalled Chinese morpheme *tân 單 'single, simple' (GSR #147a-d).[71]

A group of forms with **tu** are perhaps distinct from the above:[72]

(Himalayish)	Khaling	**tu** [high tone]
(Naga)	Yacham-Tengsa	**kha-tu**
	Wancho	**tu-ta**

---

67) In Modern Burmese the fully stressed form **tiʔ** (the regular reflex of WB **tac**) appears only in isolation, while the unstressed variant **tə-** occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form **tiʔ**, with constricted vowel, is used in counting, while a low-tone open syllable **tì** occurs otherwise. See also the unstressed form for ONE in Kayah (above 1.21), as well as Dulung **tiʔ**[55] [SUN 1982b:244-245], and Karenic forms like Pa-O *tə*ʔ-**ba**, Palaychi **tə-**, Sgaw **tə-**.

68) Supporting forms for *ʔgyik 'little bit' offered in TSR #70 include WT **cig** 'a little, few, some', WB **kyac** 'be diminutive, smaller than ordinary', Lahu **a-cí**, Akha **á-cýq**, and Moso **tçi**[55].

69) Several PLB variants are reconstructed in Matisoff 1972a ("TSR") #31/#48 and #70: *-tik ⋈ *ti ⋈ *ʔdik ⋈ *ʔ-gyik ⋈ *kyik. A revised analysis is offered in Matisoff 1985b ("GSTC") #148, where there is reconstructed a new PST word-family of the shape *day ⋈ *tay ⋈ ⋈ *dan ⋈ *tan. Still another allofam *tey is recognized in GSTC to underlie forms like WB **thî** 'single, alone'. Much work remains to be done in this complex word-family, which challenges our understanding of Lolo-Burmese vocalism in general.

70) Other candidates for membership in the *day ⋈ *tay group include Monpa [Dubey] **thee**, Nocte **wan-***the* [DUBEY, DAS GUPTA], **van-***the* [GEM] (for the 2nd syllable see *haŋ, below 3.153), Ersu **te**[55], Pumi **ti**[13] (Qinghua), **ti**[45] (Taoba). For an alternative etymology of the Nocte form see below 3.152.

71) It is demonstrated in GSTC that Lahu **-e** is the normal reflex of PTB *-ay (as well as of *-an). The variation between TB *-ay and Chinese *-an that is hypothesized for this etymon is shown to be paralleled in several other cases. For detailed further discussion of these complex word families, see Matisoff 1989/1995b, set (27).

72) Cf. also Kham *to*-**bo**. Michailovsky observes that Khaling **-u** can be a reflex of *-ik (p.c. 1995).

Miri

Padam

Minyong

Mon

Mon

Shar

We

var

Tang

Tang

Tang

Kim

Boro

N

Nten

Dim

Mik

3.15 New roots for ONE in Abor-Miri-Dafla and elsewhere in Tibeto-Burman

3.151 *tir ⋈ *tur

The AMD group and a few geographically close Bodish languages have a group of forms for ONE with dental initials, high vowels, and liquid finals, which seem independent of the other roots we have discussed with dental onsets (*tyi-k ⋈ *tyak, *t/day, or *t/dan):

Mising (=Miri)	a-ter
Padam (=Abor)	a-tel (also akem [q.v.])
Minyong	atîr ∼ ayirr (also akon [q.v.])
Milang	atel (also akan [q.v.])
Monpa (Motuo)	t'or
Monpa (Central)	thur [DAS GUPTA]
Sharchop/Tsangla	thur [Chhewang RINZIN (p.c. 1984); also NISHI 1982]

We reconstruct this etymon as *tir ⋈ *tur (-u- ⋈ -i- is a well-established variational pattern in TB; see VSTB pp. 41-42).[73]

3.152 *(t)se

This group of Kamarupan forms meaning 'ONE' seems to have undergone "contamination" with a root meaning 'TEN'. (See below 3.22, 3.4.)

Tangsa (Moshang)	ashi (GEM); ashe [DAS GUPTA 1978]
Tangsa (Muklom)	ase [DUBEY]
Tangsa (Yogli)	ashi (GEM)
Kimsing	ashi [DAS GUPTA 1978]
Boro	-she (LSI); se [BHAT 1968]
"N. Monpa A"	hi [NISHI 1982]
Ntenyi	kesü (with prefixal k-)
Dimasa	se[74]
Mikir	isi

Two higher Mikir numerals contain this morpheme in interesting combinations: **throk-si** '7', an *additive* formation based on **throk** '6' ("6 + 1") and **sir-kep** '9', a *subtractive* formation based on **kep** '10' ("1 from 10"); see below 4.20.

Distinct from the above is another group of Kamarupan forms with **-a**

---

73) J.T. Sun [1993:234] reconstructs Proto-Eastern-Tani *tel on the basis of the Mising-Padam and Milang forms. The Monpa and Sharchop forms belong with WT *thor-bu* 'single; separate ' (Jäschke 289).

74) Dimasa and Mikir s- frequently reflect the PTB affricate *ts- (STC p.28). French [1983:529] reconstructs Proto-Northern-Naga *-tse on the basis of the Yogli and Muklom forms just cited, as well as Chang Naga **cie** and Nocte **van-*the***. An alternative etymology for this Nocte form is suggested above (note 70).

vocalism, apparently from *sa or *tsa:

 Garo   sa [BURLING, PHILLIPS]; **gesa** [MOMIN]

 Kokborok **-cha** ~ **-sa** ~ **-ca**

 Lakher **sá** 'one'[75]

This Lakher morpheme is also used as a prefix before all the numerals 2-10, e.g. **sa-pali** '4' (literally "1 × 4"), **sa-pangaw** '5' ("1 × 5"). Also perhaps reflecting this etymon are Tiddim **a-ma-sa** 'first', Lotha **ma-tsa-nga** 'one'.

3.153　*han or *haŋ

 Serdukpen **han** [DUBEY]

 Zeme   **hangkat**

 Maram **hangline**

This morpheme seems to function as a fully syllabic prefix[76] with the lower numerals in a few languages:

 Nocte **van-the** 'one', **vanyi** 'two' (< *van-nyi), **van-ram** 'three'

 Maram **hang-line** 'one', **hang-na** 'two', **hang-tum** 'three'

3.154　*a

This "minimalist" morpheme has only been unearthed in a couple of languages so far, but seems to represent a genuine etymon:

 Aka (Hruso)   **a**

 Qiang (Taoping)[77] **a²¹** [SUN 1981:217]

 Qiang (Mawo) **a** [SUN 1981:217]

3.155　(k-)lV(N)

A number of forms with lateral initials look as if they are related somehow, though their vowels cannot yet be reconstructed. A couple of these words for ONE have final nasals, which make them look suspiciously like a root for TEN reconstructed as *l/riŋ � l/ryaŋ (below 3.233, 3.4):

 Pwo Karen **làn** [JONES 1961:618]

 Yimchungru **khü-lang** (first syllable is a prefix)

 Sangtam   **khürü** (*also* **khe**)

 Kezhama   **kele** (**ke-** is a prefix, part of a 1-3 run)

 Mao   **kali** (**ka-** is a prefix, part of a 1-3 run)

 Sema   **la ki** (also **khe**)

 Maram   **hang-li-ne** (cf. *haŋ, above 3.153)

---

75) Lorrain [1951:59] notes that the word for 'one' (here written **sá**) is "high-pitched", while the nearly homophonous Lakher word for 'thousand' is "low-pitched" (**sà**). Undoubtedly cognate to these Kamarupan forms is Jingpho **šà** 'only' [DAI et al. 1983:726].

76) See below 5.446, "Where the lowest run has a fully syllabic CVC- prefix".

77) These Qiang dialects have other allomorphs for ONE (which occur in compound numerals like ELEVEN): Taoping **tʃi³³**, Mawo **tɕi** [SUN 1981:217], clearly from the *t(y)i- family (above 3.14).

It is not clear whether these forms are relatable to a group of phonological-ly similar Himalayish words for ONE, e.g. CHOURASE **kolo/kwalo**; Yakkha **kolok**; HAYU [HODGSON] **ko-lu**.   See above 3.13b.

3.156   Miscellaneous residual forms
(a)   A few Naga languages of the Angami group have words for ONE with initial **p-** and a back vowel:

Angami	**puo** (Kohima), **po** (Khonoma)
Chokri	**pü**
Chakhesang	**püh**

(b)   A few Kamarupan languages have forms with initial **m-** and (except for Rengma) a non-front vowel:

Deng Geman	**kɯ³¹ mu⁵³** [SUN et al. 1980:252]
Kaman (Miju Mishmi)	**ku-mo** [DAS GUPTA 1977a]
	kmo: ~ kŏmo: (LSI 3.1:623)[78]
Rengma	**me** 'one'
Tiddim	*a-ma*-**sa** 'first'
Lotha	*ma*-**tsa-nga** 'one'
Meithei	**ama** 'one'

Cf. also Meithei *ma*-**pan** '9', a subtractive formation from TEN ("one from ten"), alongside **ni-pal ~ ni-pan** '8' ("2 from 10"); cf. *****ban** ⋊ *****bal** 'ten', below 4.203.[79]

(c)   The Jirel form for ONE given in CSDPN is **dok-pei**.   It is tempting to compare the first syllable with Chinese 獨 'alone; only' (OC *****d'uk** [GSR #1224i]), but since Jirel is a Bodish dialect, it would be well to find a cognate in WT before going out on a limb.

3.16   ONE as indefinite article or general numeral affix
ABOR-MIRI shows a clear picture of semantic interchange between ONE and a kind of indefinite article or general numeral suffix: AM **a-ko** 'one', **-ko** 'general numeral suffix' (above 3.13b).   This same etymon appears as a suffix in the vestigial numeral systems of Kiranti languages like MEWAHANG, YAK-KHABA, and LOHORONG, and has been generalized with all the numerals in NEWARI (**cha-gu** 'one', **ni-gu** 'two', **swa-gu** 'three'...**jhi-gu** 'ten').   We have also seen Lakher **sa-** used as a prefix with all the numerals from 1 to 10; this etymon appears as the independent word for ONE in Garo (above 3.152).

---

78)   **kɯ-/ku-/k-** is now a meaningless prefix, part of a secondary 1-6 "prefix run" (below 5.5), but is perhaps itself a reflex of *****ko** 'one' discussed above (3.13). Cf. also Mishmi (DUBEY) **mu-ou** '10', especially in the context of the interchange between ONE and TEN (below 3.4).

79)   As a longshot we might compare these forms with the Lahu 'general classifier for objects', **mà** [MATISOFF 1973a:91-92, 1988a:975-976].

In fact nothing is more natural than for a language to develop a generalized counter or an indefinite article by semantic bleaching of the numeral for ONE. This is of course what has happened in English,[80] and a similar process is now well advanced in Israeli Hebrew, where the numeral *exad* 'one' is rapidly developing into an indefinite article.

3.2 Profile of Number TEN

As STC observes (p. 94), "extreme variation obtains" in TB with respect to etyma for the number TEN. The special importance and salience of TEN in decimal systems sets it apart from the ordinary numerals 2-9. Since a morpheme meaning '10' normally occurs in all compound numerals (both the TEENS and the ROUND NUMBERS), there is frequently morphophonemic variation as it interacts with its fellow constituents.[81] Often a language will maintain several etymologically distinct morphemes for '10', one used as the independent numeral, and the other(s) for the TEENS and/or ROUND NUMBERS.

3.21 *gip ⋇ *gyap

In STC #16, a PST etymon *gip 'ten' is reconstructed, based on Limbu gip (in comp.), Miju kap ~ kyep, Mikir kep, Maring tśip, Yawdwin (S. Kukish) gyip (in comp.), WB (ə)kyip.

In fact, however, this seems to be still another root where we must posit -i- ⋇ -ya- variation,[82] as witness these forms from an AMD language: Kaman (Miju Mishmi) kyap-mo (LSI has kap), Deng Geman kiap⁵⁵ mu⁵⁴.

The obvious Chinese cognate is 十, reconstructed as OC *d̥i̯əp in GSR #686 (see STC p.175).

3.22 *ts(y)i(y) ⋇ *tsyay

One other root for TEN is reconstructed as *ts(y)i(y) in STC (#408 and pp. 131, 136), based on the following forms:

Jingpho tśi ~ śi, Namsang (= Nocte) i-tśi, Moshang rok-śi,[83] Garo tśi, Dimasa dźi, Miju si (in comp.), Karen (Taungthu) tśi, (Pwo and Sgaw) shi.

The vocalism of WB əchai poses a problem,[84] which the original version of

---

80) English *an* derives from the unstressed variant of *one*, just as the preposition *of* is historically an unstressed version of *off*. The schoolchild's chant "*a one, and a two, and a three, let's go!*" is perhaps the closest English equivalent to the TB penchant for modifying all the numerals by a form of the number ONE.

81) In much the same way as the English *ten* has the allofams *-teen* (<OE *-te:ne, -ty:ne*), and *-ty* (e.g. *twenty* < OE *twe:gentig* 'twice ten' < *-tig* '10').

82) See *tik ⋇ *tyak (above 3.14), *it ⋇ *yat (above 3.11).

83) The MOSHANG (= Tangsa) form cited in STC seems to be an error, since the second syllable means ONE, not TEN (cf. **rok-ni** '20', **aśi** 'one', **ani** 'two'); **rok** is from a distinct ↗

STC glosses over with the remark (p. 94) that it "appears to be related to this root through vowel gradation". In a new footnote (n.272), Benedict suggests that the solution is to change the reconstruction to *tsyay, but I feel it is better to recognize both allofams at the proto-level, *ts(y)i(y) ✕ *tsyay, giving us yet another instance of the -i- ✕ -ya- variational pattern in numerals.

Many other forms may be added in support of this reconstruction:

(*Himalayish*) Newari **jhi**, Tamang **ci**, Sherpa **ci-thambaːq**, Tsangla/Sharchop **s(h)e**, Kanawari **sáí**.[85] Here belong several other Himalayish forms with **-u** vocalism:[86] WT **béu** (Lhasa **cu**), Kaike **chyu**, Gurung **cyuq**, Thakali **cyu**, Jirel **cyu-taːmbaːq**, Sikkim Bhutia **chu-tamba**, Dzongkha [Mazaudon] **cu-thãm**.[87]

(*Kamarupan*) Monpa (Cuona) **tɕi⁵³**, [Dubey] **chi**; Monpa (Motuo) **se**; Garo **chi-kung** [Phillips], **ci-king** [Burling]; Kokborok **ći**

(*Baic*) Bai **tsw⁸** [Dell]

(*Qiangic*) Ersu **tshe⁵⁵**, Proto-rGyarong *sy**ʈ**sye [Nagano 1984] < *s-tsyiy [JAM]

(*Loloish*) Proto-Loloish *tši¹ > Lahu **chi**, Akha **tsé**, Lisu **htsi⁴**, Phunoi **təsé** (**tə-** 'one'), etc.

(*Nungish*) Dulung **tsäl⁵⁵**, Rawang **hti** *sel*, Nujiang **ts'i⁵⁵** **tshãn⁵⁵** [88] (first syllables mean 'one'). Have these curious Nungish forms with final **-l** developed from *-y, or do they point to an allofam *tsyal?

3.23   New roots for TEN in Kuki-Chin-Naga, Abor-Miri-Dafla and elsewhere

3.231   Proto-Kuki-Chin *som (< *tsom)

This root is widespread in Kuki-Chin,[89] both as the independent numeral for TEN, and as the first constituent in higher multiples thereof:

Kom Rem, Kuki, Puiron **som**; Lushai **shom** 'ten', **shom-hniʔ** 'twenty';

Moshang root meaning TEN, not mentioned in STC (below 3.233c). Ultimately, however, I believe that the meaning ONE for forms like **shi**, **she** might actually be a transference from an original meaning TEN (see above 3.152).

84)   Note that Burmese has reflexes of both *gip (3.21) and *tsyay, with some repartition of function. According to Judson [1953:215], "**kyip** is substituted for **chai** in the numbering of rational beings."

85)   Perhaps ✕ Kanawari **sa'e-** '10 in additive higher round numbers'; see below 3.533[D].

86)   Michailovsky and Mazaudon [1992] point out that WT and other Himalayish **-u** corresponding to yodated vowels elsewhere is paralleled in several other roots (e.g. 'bow' PST *d-ləy (STC #463), but WT **gźu**), and may be viewed as a quasi-regular (dissimilatory?) development after palatal affricate initials.

87)   The morpheme **-t(h)am-** in many of these forms is to be referred to PTB *dyam ✕ *tyam 'full' (STC #226); for a detailed discussion of this root see Matisoff 1988b. See below 3.235.

88)   Undoubtedly this **-n** is from an earlier lateral *-l.

89)   See Ono 1965.

Gangte, Hmar, Paite, Tiddim, Vaiphei **sawm**; Anal, Lai, Laizo, Ngawn, Thado **sɔm**; Maring **chip** 'ten' (< *gip [3.21 above], but **som-thum** '30', **som-li** '40';[90] Zotung **suŋ**

There is evidence that this etymon may be more widespread, at least in the Kamarupan nucleus of TB. One likely relative is the GARO bound morpheme for TEN (**sot-**) in the round numbers from '40' on up, e.g. **sot-bri** 'forty', **sot-bonga** 'fifty' (below 3.51). J.T. Sun [1993:277] proposes a relationship between the KC forms and his Proto-Tani *čam, also a bound morpheme occurring in multiples of ten (e.g. BENGNI **čam-ni** 'twenty', **čam-pi** 'forty'; Hill Miri **čom-oum** 'thirty', **čaŋ-ŋo** 'fifty', **čem-piŋ** 'eighty').[91]

For now we reconstruct this etymon as Proto-Kamarupan (maybe ultimately PTB) *tsom.

3.232   *pal or *bal

Several forms meaning TEN in Northern Naga languages (Chang **an**, Phom **an**, Konyak **pen**, Wancho **ban**) led W.T. French [1983:565-566] to set up PNN *boːn, though he suggests that this might be a "loan from Austroasiatic into Northern Naga", citing KHASI **ši pón** 'ten' (**ši** means 'one'[92]).

I consider this loan origin highly unlikely, however, in view of a pair of very interesting forms in MEITHEI: **nipal** ∼ **nipan** 'eight', **mapan** 'nine'. These are both subtractive formations from TEN, meaning respectively "2 from 10" and "1 from 10" (Meithei **ani** '2', **ama** '1'). [See below 4.20.] These forms seem to indicate that the original final consonant in this root was *-l.[93]

Also undoubtedly to be assigned to this etymon are Phom **püan-** ("plus ten"), used in the odd round numbers of its vigesimal system, e.g. **pinyi-püan** '50' ("[2 × 20] + 10"); and perhaps also Ntenyi **apyam-**, used in the decimal formation of its round numbers from 60 to 90 (see below 3.522).

3.233   *s-r/liŋ ⋊ *s-r/lyaŋ

The AMD languages clearly point to an etymon for TEN with liquid initial (it is not easy to decide whether it was *r- or *l-), velar nasal final, and a vocalic nucleus that displays the familiar *-i- ⋊ *-ya- variational pattern:

Abor-Miri **eying-ko, iying-ko** (**-ko** 'one'); Minyong **e'ying**; Tagin **ering**; Nishi **aring**, Nishing/Dafla **eriŋ** ∼ **erjaŋ** (Das Gupta; note the intralingual variation of rhyme), **reng-cheng** (Yano), **raːng** (Robinson), **il-lyi** (E.

---

90)   GEM (gives Maring **som-nga** for both '20' (p.279) and '50' (p.79), but the former seems to be an error. See below 3.513, 4.14.

91)   Note that this morpheme for TEN precedes the unit both in KC and in Tani. There is another etymon for multiples of ten in Tani (PT *rjɯŋ), but it *follows* the unit. See below 3.233.

92)   Note the fortuitous similarity of this Khasi morpheme to some of the TB forms for ONE cited in 3.152.

93)   The independent word for TEN in Meithei is **tara**, whose affiliations are elsewhere (below 3.233a).

Dafla), Apatani **alyã** (for humans) ✕ **lya** (for nonhumans; < *lya-ŋ); Gallong **i'ri°** ~ **i'yi°**; Padam (Dubey) **i:yi, i:i**; Aka (Hruso) **rhi, ru**; Taraon **ha:long**, Darang Deng **xɑ⁵⁵luŋ⁵⁵**; Idu **hũ** [TALUKDAR et al. 1962], **hoŋ⁵⁵ɦoŋ⁵³** (Luoba: Sun 1983); Chulikata **hush** (< *hu-shV, with vowel of second syllable apocopated[94])

J.T. Sun [1993:144] sets up Proto-Tani *rjuŋ on the basis of Bengni **ɯ-rjuŋ**, Lhopa/Bokar **ujuŋ**, and the above Abor-Miri (=Padam-Mising) forms, also citing Dhammai **lin**, Bangru **rəŋ⁵³**, and Idu/Luoba (ZMYYC) **hɹoŋ⁵⁵** (used in multiples of ten, e.g. **ni⁵⁵hɹoŋ⁵⁵** '20', **ɑ³¹soŋ³⁵hɹoŋ⁵⁵** '30').[95]

Several of the above AMD forms with **h-** or voiceless sonorant initials point to a possible *s- prefix on this root (Aka **rhi**, Idu **hũ** ✕ **hɹoŋ⁵⁵**), and the same is true of an apparently solid Sema Naga cognate, **lho-** 'combining form in multiples of ten', as in **lho-bidi** 'forty'. Weidert [1987:249] reconstructs a Proto-North Assam etymon *lhyaŋ 'ten (in decimal counting' [i.e. in multiples of ten]), to which he assigns this Sema form, along with KEZHA(MA) **lha-** (e.g. **lha-pangu** '50'[96]), TANGKHUL **həŋ-** (e.g. **həŋ-phənga** '50'), Southern Rengma **hẽ** (e.g. **hem-pfũ** '50'), Angami (Kohima) **hie-** (e.g. **hie-pengou** '50'), and Chokrü (=Chokri) **he-** (e.g. **hie-püngu** '50' [GEM]; we may add Angami (Khonoma) **lhi-** (e.g. **lhi-pengu** '50').

We should now consider a large number of sesquisyllabic Kamarupan forms, mostly from the Naga group (as cited in "GEM", Marrison 1967), with dental or velar prefix followed by a full syllable with a liquid onset. Though they all seem to be related internally, the vocalic correspondence is obscure (partly due to the inadequate phonetic transcriptions of the forms available to GEM). In the present state of our knowledge, it is not clear whether to assign them to *riŋ ✕ *ryaŋ, or rather to the stop-finalled PNN etymon *rok discussed below (3.234):

(a) With velar prefix:

Angami **kerü**, Chokri **küri**, Chakhesang **keri**, Liangmai **kariu**, Maram **kero**, Mzieme/Zeliang **kerei**, Zeme **kereu**, Mikir **kre-**[97]

(b) With dental prefix:

Ao (Mongsen) **tera**, Ao (Chungli) **ter** (with apocope)[98]; Khoirao **sara**, Lotha **taro**, Meithei **tara**, Meluri **tera**, Ntenyi **dagha, ta'a**[99] Pochury **türa**, Rengma **tsarü**, Sangtam **thüre**, Tangkhul **thara**, Yacham-Tengsa **thelu**,

---

94) This is an apocopating language. Cf. Chulikata **kāsh** 'three' < *g-sum (below 4.12, 5.131).

95) These AMD forms are phonologically quite similar to another, probably distinct root for HUNDRED, below 3.546.

96) All these illustrative forms meaning '50' are from GEM, not Weidert.

97) This is the Mikir combining form for teens, as in **kre-isi** '11', **kre-hini** '12'; the independent Mikir numeral '10' is **kep** (above 3.21).

98) This form bears a merely accidental resemblance to some AMD forms descending from *tir 'ONE', e.g. Mising **a-ter**, Padam **a-tel** (above 3.151).

99) The apostrophe probably means glottal stop.

Yimchungru **thürü**

(c)   With palatal affricate word-initially:

Kezhama **chiro**, Mao **chüro**, Sema **chü***ghi*

The first syllables in these last three forms require some comment. On the one hand, they bear a superficial resemblance to reflexes of **tsyiy ⪥ tsyay* (above 3.22). A closer look convinces us that they are merely prefixal. This is especially clear in Mao where all the higher numerals (6-10) participate in a prefix run with a palatal pre-syllable (**choro** '6', **chani** '7', **chacha** '8', **choku** '9', **chüro** '10'). The second syllable of Sema **chü***ghi* (where the "gh" presumably stands for [ɣ]) agrees well with other Naga forms (e.g. Chakhesang **keri**), and might well be an intralingual co-allofam of the Sema combining form (above). Alternatively, **lho-** might better be assigned to **s-ryak ⪥ *s-rwak* (next section).

As a possible Himalayish connection to this etymon, we should mention KAIKE **phera:ng**, used in its vigesimal system of round numbers to express "minus-ten" from the next higher multiple of 20 (below 3.533).

3.234   **s-ryak ⪥ *s-rwak*

French [1983:565] sets up a PNN etymon **ro:k* on the basis of several combining forms for multiples of ten in Northern Naga languages:

Tangsa (Moshang) **rok-shi**[100] '10' ("10×1"), **rok-tachat** '80'; Tangsa (Yogli) **rauk-shi** '10', **rauk-tüchat** '80'; Kimsing **ro-shi** '10', **ro-bangi** '50' [DAS GUPTA 1978]; Nocte **i-chi** '10' (< **tsyiy*), but **ruak-banga** '50', **ruakisat** '80'

However, a better PNN reconstruction would be **rwak*, in view of a number of forms from other Naga languages that point to **ryak*:

Zeme **riak-seruk** '60', Liangmei **ria-charuk** '60', Nruanghmei (=Rongmei) **rek-cüruk** '60'[101]

Somewhere in this word family (probably under the **rwak* allofam) we must also include the Nruanghmei independent numeral **ruh** '10', as well as the LAKHER morpheme **-hraw** '10', which apparently must always be preceded by one of three semantically equivalent prefixes: **sa-***hraw*, **pa-***hraw*, **mia-***hraw* '10'.[102] The Lakher voiceless liquid clearly points to an **s-* prefix at an earlier stage.

We may thus combine the etyma discussed in 3.233-3.234 into a single word family comprising both nasal- and stop-finalled allofams, and displaying both -i- ⪥ -ya- and -y- ⪥ -w- variation:

---

100) This form was cited in STC #408, but the first syllable was not related to anything else, and the second syllable was misinterpreted as TEN, not ONE.

101) See Weidert 1987:413 and below 3.5212.

102) Lakher also has a multiplicative combining form for the multiples of ten, **sy-** (e.g. **sy-pali** '40'), hence '10' can also be expressed as **sy-kha** (**kha** 'one').

*s-riŋ ⋈ ⋈ *s-ryaŋ ⋈ *s-ryak ⋈ *s-rwak

There is some evidence of phono-semantic interchange between TEN and HUN-DRED/THOUSAND in this root, which once might have meant something more vague, like "BIG NUMBER." See below 3.546, 3.547.

3.235   *d(y)am ⋈ *t(y)am 'ten; a full decade'

We have already mentioned (above 3.22, n. 87) a morpheme meaning FULL that occurs in several Himalayish compounds for TEN, evidently signify-ing something like the *completion of a full decade*, e.g. Sherpa ci-*tham*-baːq, Jirel cyu-*taːm*-baːq, Sikkim Bhutia chu-*tam*-ba.[103]

Several other Himalayish languages have words for TEN with a similar-looking morpheme, though a connection with the concept FULL has yet to be demonstrated: Bahing kudum '10' [GVOZDANOVIĆ 135]; Khaling tadam (the first syllable looks like a reduction of tu 'one'); Thulung Rai ko-dium (glossed "one-zero" in Allen 1975); Lepcha kɑ-ti (kat 'one'; see above 3.12).[104] Note that in these languages the first element means ONE, whereas in the Bodish languages the first element means TEN.

This morpheme for TEN, perhaps bleached of any synchronic association with FULL, seems also to occur in Qiangic: Pumi (Taoba) ka^{55}tī^{55}, Pumi (Qinghua) qa^{55}stiɛ̃^{55}, Qiang (Taoping) xɑ^{21}dy^{33}, Qiang (Mawo) hədiu.

3.236   *p/boŋ

Several Kiranti languages (E. Nepal) have multiplicative morphemes that occur in compounds for the multiples of ten, reflecting Proto-Kiranti *poŋ or *boŋ:

> Kulung ik-poŋ '10', ngi-poŋ '20', etc.; Yakkhaba ip-poŋ '10'; Limbu thi-booŋ '10', ni-booŋ '20', sum-booŋ '30', etc. [GVOZDANOVIĆ 136, 146, 162]

Limbu also has an interesting form i-booŋ 'NINE', which looks as if it may have been transvalued or "downstepped" from an earlier meaning of TEN (compare Yakkhaba ip-poŋ). The words for ONE in Limbu and Yakkhaba are thik (above 3.14) and ik-ko (above 3.11), respectively. See below 4.02.

This root *p/boŋ is distinct from the general, meaningless suffix -pok/-bok attached to whole sets of Kiranti numerals (above 2.1).

3.237   *Apparent isolates*

There remain a number of isolated forms meaning TEN in individual

---

103) This root has several reflexes in Tibetan, e.g. WT ltams-pa 'be full', tham-pa ⋈ them-pa 'com-plete, full', ldem-pa 'straight, upright'. As demonstrated in JAM 1988b, STC #226 'full' and #227 'straight/flat' really represent one and the same etymon. The presence of the "infinitive" or nominalizing suffix -ba/-pa in these Bodish forms indicates that the preceding morpheme is in-herently verbal.

104) Perhaps allofamically related to this root is Lepcha t'ăp (< *tap, with homorganic final stop), used in teen-formation, e.g. sam-t'ăp '13', tărăk-t'ăp '16'; see below 3.32[A].

languages that so far resist attempts at etymologization, including:
(a) Boro *khao*-**she** '10' (-**she** 'one') [105]
(b) Mishmi **muou** '10'
      Is the **mu-** segmentable off with the meaning 'one'? Cf. Geman Deng
      kiɑp-*mu* 'ten' (for the first syllable, see above 3.21).
(c) Milang **hang-tak** '10' (**hang-** 'one')
(d) Apatani **khrā** '10'
(e) Damu **pət** '10'

3.3   TEEN Formation: From 10 to 20

As we shall see, "teen problems" are not limited to acne and sexual awakening. For our purposes, the "teens" include all the numerals from 11 to 19 — it is only an accident of English morphophonemics that ELEVEN and TWELVE lack the *-teen* suffix. The teens are almost always morphemically complex, i.e. combinations of a morpheme for TEN and one for the particular unit from one to nine.

Parameters to consider in analyzing teen-systems include:
• Does the TEN morpheme come before or after the UNIT morpheme? E.g., is '19' TEN + NINE (like French *dix-neuf*) or NINE + TEN (like German *neunzehn*)?
• Is this TEN morpheme identical to the independent simple numeral for '10' (as in French *dix-sept, dix-huit, dix-neuf*)?
• If it is not, is it merely an allofam (morphophonemic variant) of the ordinary independent numeral for '10' (like Eng.-*teen*), or is it a totally separate etymon (e.g. Mikir **kep** '10', **kre-** '-teen')?
• Is the complex numeral agglutinative, easily segmentable into the TEN part and the UNIT part (Fr. *dix-huit*, Germ. *achtzehn*), or is it fusional (Fr. *onze, douze, treize, quatorze, quinze*)? Does the UNIT morpheme undergo morphophonemic change when combined with the TEN morpheme (e.g. Eng. five /fayv/ but fifteen /fif-/; three /θriy/ but thirteen /θər-/)? Does an epenthetic sound get inserted at the morpheme boundary?
• Are the TEN and UNIT morphemes combined by simple juxtaposition, or is the additive combination explicitly marked by a linking morpheme?
• Do any teens occur that are not simple additive combinations of TEN and UNIT? For the higher teens (16-19 or 17-19), are there any *subtractive* formations based on TWENTY (e.g. Latin *duodēvīginti* '18', *undēvīginti* '19')?
• Strictly speaking the concept of TEENS is only relevant to decimal systems of numerals. In vigesimal systems, the numbers 1-19 correspond to the "units" 1-9 of decimal systems. In vigesimal systems, it is advantageous to be able to

---

105) There is an interesting lookalike in Hmongic, e.g. White Hmong: **kaum** [kaoʔ²¹] '10', (nees) **nkaum** [nkaoʔ²¹] '20', (peb) **cɑug** [kyɑo²¹] '30'. See Heimbach 1969:9, 77, 152.

refer to the numbers from 21-39 as a group — I suggest the term *TWEN-TEENS*. The twenteens 21-39 of vigesimal systems correspond to the teens 11-19 of decimal ones. (See, e.g. the discussion of Sherpa, below 3.534[B].)

• In the excessively rare duodecimal type of system represented by Chepang (below 3.535), the numbers 13-23 correspond to the teens of decimal systems. We might as well call them the *TWELVEENS*.

3.31   Teen formation in Kuki-Chin-Naga[106]

In virtually all languages of this group so far examined, the TEEN morpheme precedes the UNIT, the only apparent exception being Maram (below).

3.311   Where the combining form ("-TEEN") is identical or morphophonemically related to the independent numeral TEN

	TEN	-TEEN	3	13	5	15
Angami (Kho.)	keru	kero-	se	kerose	pengu	keropengu
Angami (Koh.)	keru	kere-	se	kerose	pengou	kerepengou
Ao (Chungli)	ter	ter(i)-	asem	terasem	pungu	teripungu
Ao (Mongsen)	tera	tera-	asam	teraasam	phanga	teraphanga
Chokri	kuri	kuri-	su	kurisu	pungu	kuripungu
Meluri	tera	tera-	keche	terakeche	manga	teramanga
Nocte	ichi	ichi-	vanram	ichivanram	banga	ichibanga
Ntenyi[107]	ta'a	ta'a-	keching	ta'akecham	munga	ta'amanga
Sangtam	thure	thure-	asang	thureasang	munga	thuremunga
Sema	chughi	chughi-	kuthu	chughikuthu	pongu	chughipongu

3.312   Where a linking morpheme occurs between the TEN and the UNIT

	TEN	LINKER	3	13	5	15
Chang	an	-tak-	sam	antaksam	ngau	antakngau
Khoirao[108]	sara	-na-	kathum	charanakasum		
Konyak	pen	-me-	lem	penmelem	nga	penmenga
Lotha	taro	-si-	etham	tarosietham	mungo	tarosimungo
Moshang	rokshi	-ra-	atum	rokshiraatum	banga	rokshirabanga
Nruanghmei	ruh	-na-	kathum	ruhnakathum	pangu	ruhnapangu
Phom	an	-pu-	jam	anpujam	nga	anpunga
Tangkhul	thara	-da-	kathum	tharadakathum	phanga	tharadaphanga

---

106) All data in this section is from Marrison (GEM), except for the Pochury forms (for which see Nagaland Bhasha Parishad 1972b).

107) The independent Ntenyi form for THREE is either **keching** or **keshang**, both different from the combining form **-kecham**.

108) '13' and '12' are the only Khoirao teens to be found in Marrison. Note the change in the consonant of the UNIT morpheme (**kathum** > **-kasum**), paralleled also in '12' **charanakachi** (**kati** '2'), as well as the change in the prefix of the TEN morpheme (**sara** > **chara-**).

Wancho	ban	-ba-	ajam	banbajam	aga	banbaga
Y-Tengsa	thelu	-le-	asam	talulesam[109]	phungu	talulephungu
Yimchungru	thuru	-kheak-	asam	thurukheakasam	phungu	thurukheakphungu
Zeme	kereu	-ze-	kechum	kereuzekechum	mengeu	kereuzemengeu

3.313 Where the linking morpheme comes after the UNIT

	TEN	LINKER	3	13	5	15
Liangmai	kariu	-kiu	shum	kariushumkiu	mangiu	kariumangiukiu
Mao	churo	-o	kosu	churokosu-o	pongo	churopongo-o
Meithei[110]	ara	-thoi	ahum	tarahumthoi	manga	taramanga
Mzieme[111]	kerei	-ngkei	ketsum	karieketsumngkei		
Puiron[112]	som	-to	thum	somthumto		
Rengma	tsaru	-chu	keshan	tsarukeshanchu	pfu	tsarupfuchu

3.314 Where the combining form ("-TEEN") is a different etymon from the independent numeral TEN

MIKIR

10	-TEEN	3	13	5	15
kep	kre-	kethom	kre-kethom	phongo	kre-phongo

3.315 Where the TEEN morpheme follows the UNIT:

MARAM

1	11	2	12	3	13
hangline	kerui-kaniko	hang-na	nangko	hang-tum	tumko

Unfortunately 11-13 are the only Maram teens given in Marrison. The independent word for TEN is **kero**, which evidently is the basis for the first element in ELEVEN. The **-ni-** of **kaniko** may be an allofam of the **-ne** of **hangline**.[113] The morpheme **-ko** apparently means '-*teen*'. Note the intrusive **-ng-** in TWELVE.

3.316 Subtractive higher teens

The phenomenon of subtractivity in the formation of the higher teens has no genetic significance — dialects of the same language may differ in this respect. Thus in Marrison's data Angami (Khonoma) and Ao (Mongsen)

---

109) Note the combining form **talu-**.

110) Meithei has the linking **-thoi** suffix only in 11-13; the rest of its teens are formed by simple juxtaposition (e.g. 15).

111) 11-14 are the only Mzieme teens that appear in Marrison. Note the combining form **kerie-** vs. independent **kerei** '10'.

112) Puiron '15' is lacking in Marrison, but 11-14 are **somkhatto, somkhanito, somthumto, somlito** (< **khat, kani, thum, mali** '1-4').

113) **-kani-** looks like it should mean TWO (cf. Puiron **kani** '2', **somkhanito** '12'), though that does not fit the meaning!

have subtractive higher teens, but Angami (Kohima) and Ao (Chungli) do not.

	ANGAMI (Khonoma)	AO (Mongsen)	MELURI
SIXTEEN	[kerosuru]	mukyi*mupen*terok	mukwe*shun*taro
SEVENTEEN	meku*pomo*thena	mukyi*mupen*teni	mukwe*shun*teru
EIGHTEEN	meku*pomo*thetha	mukyi*mupen*tsit	mukwe*shun*tuze
NINETEEN	meku*pomo*theku	mukyi*mupen*tuku	mukwe*shun*tokhu
TWENTY	meku	mukyi	mukwe
	NTENYI	POCHURY	RENGMA
SIXTEEN	kwu*she*tuo	mke*shun*toro	nki*pamo*tsaro
SEVENTEEN	kwu*she*tughu	mke*shun*turu	nki*pamo*tsanu
EIGHTEEN	kwu*she*tuza	mke*shun*tuze	nki*pamo*tutse
NINETEEN	kwu*she*tukhu	mke*shun*toku	nki*pamo*tukhu
TWENTY	mekweru/mukwung	mke	nki

These formations are subtractive in a different sense from, e.g., Latin *duodēvīginti* '18' and *undēvīginti* '19', lit. "two from twenty" and "one from twenty", respectively. The last morphemes in these Naga words for 16-19 are not the lower numerals 4,3,2,1, but rather the additively appropriate higher numerals 6,7,8,9. That is, the expressions mean something like *the six before twenty, the 7 that comes before 20,* etc.[114]

We therefore assume that the linking morphemes in these numerals (i.e. Angami **-pomo-**, Rengma **-pamo-**, Ao **-mupen-**, Meluri and Pochury **-shun-**, Ntenyi **-she-**) mean something like "before."[115]

3.32   Teen formation elsewhere in Tibeto-Burman
[A]   HIMALAYISH

Evidently the norm in Himalayish teen-formation is to have the morpheme for TEN precede the UNIT. Usually there is no overt marker of the conjunctive relationship between the TEN and the UNIT (as in, e.g. Tibetan and its dialects, Newari, Thakali, Kanauri, Dzongkha). Lepcha seems to be an exception to both of these generalizations:

LEPCHA

ONE	kat	THIRTEEN	sam-t'əp
TEN	kəti	FOURTEEN	fəli-t'əp
ELEVEN	kəti-kat-t'əp	SIXTEEN	tərək-t'əp
TWELVE	kəti-nyət-t'əp	NINETEEN	dəkyot-t'əp

Lepcha '11' and '12' are formed with the linking morpheme **t'əp** added to the independent numeral for TEN plus the units ONE (**kat**) and TWO (**nyət**). However, in the higher teens the independent morpheme for TEN disappears,

---

114) This is explicitly stated (in Hindi) in the Pochury source (p. 16).
115) See below 4.20, "Additive, subtractive, and multiplicative formations".

and the **t'əp** takes over its semantic load.[116] Note that now the order of constituents is reversed, so that the UNIT precedes the TEN.

In KANAWARI and DZONGKHA/SIKKIM BHUTIA, the combining form for TEN undergoes morphophonemic changes, though it is not clear how big a role free variation and/or vowel harmony are playing. Thus, Kanawari **sai** '10', **sanish** '12', **sorum** '13', **sapü** '14', **songa** '15', etc.; Sikkim Bhutia **chu** '10', **chu-sum** '13', **chegye** '18'. In SHARCHOP, the independent numeral for TEN is **she**, but the combining form for -TEEN is **song-**, e.g. **song-sam** '13', **song-zon** '17'.

[B] LOLO-BURMESE

In Loloish the morpheme for TEN is often a classifier, not a numeral — i.e. '10' may be expressed multiplicatively as "1 × 10" (e.g. Lahu **tê** '1', **tê-chi** '10'), in the same way as the other round numbers (e.g. Lahu **nî** '2', **nî-chi** '20'). Teens are then formed by adding the UNIT morpheme after the classifier for TEN, e.g. Lahu:

tê chi tê	'11'	"[1 × 10] + 1"
tê chi khɔ̂ʔ	'16'	"[1 × 10] + 6"

In Burmese, TEN is also a classifier with respect to the round numbers: WB **tac** '1', **tə-chay** '10'; **hnac** '2', **hnə-chay** '20' (**tə-** and **hnə-** are unstressed combining forms of '1' and '2'). However, the teens contain only the root **chay-** '10' itself, without the ONE morpheme: **chay-tac** '11', **chay-hnac** '12', **chay-khrok** '16'.

Alternatively, TEN may function as an ordinary numeral, as in AKHA **tshé** '10' (with ONE not expressed), **tshé-tîʔ** '11', **tshé-kòʔ** '16'.

[C] BODO-GARO

In this branch of TB, the teens are usually formed simply by juxtaposing TEN to the UNIT:

GARO

chi-kung '10' (**kung** seems to mean ONE: see above 3.11)

chi-sa '11' (**sa** is the independent numeral ONE)

chi-gni '12', **chi-gatham** '13', **chi-dok** '16', **chi-sni** '17'

DIMASA

ji '10', **ji-se** '11', **ji-gini** '12', **ji-gatham** '13', **ji-biri** '14'; but there is a special form for '15', **je-ra**, which looks unrelated to the independent numeral **bonga** '5'. Dimasa is also a language with a special form for FIFTY **dan**, which looks related to the KCN group discussed below (3.522).

---

116) I suspect that this **t'əp** is actually from the root *tap 'fold, layer, place atop one another, order, succession' [STC #493, p. 184]. (Beware, there is a different root also numbered #493 on p. 173!) This may in fact also be the etymological home for the nasal-finalled morpheme *tam glossed '10' (above 3.235).

KOKBOROK[117]

c(h)a ~ -sa '1', -ci '10', kay-ci-cha '11', kay-ci-ba '15', kay-ci-cuku '19'.

I believe this **kay-** is related to the Boro morpheme in **zokkay** 'group of four' [below], and means something like 'group' or 'unit'.[118]

BORO [BHAT 1968]

Boro has the most thoroughgoing QUATERNARY (four-based) system that I am aware of in TB.[119]/[120] As D.N.S. Bhat says, "The system is basically a quadruplous one" [1968:29].

ONE	se	SIX	zokkay-se kanəy
TWO	nəy	SEVEN	zokkay-se katam
THREE	tam	EIGHT	zokkay-nəy
FOUR	brə	NINE	zokkay-nəy kase
FIVE	ba	TEN	dos [< Indo-Aryan]

The numbers 1-5 are the usual Barish set of inherited forms (with preemption by the prefix in FIVE). The higher numerals, however, are formed on the basis of groups of four (**zokkay**). Even multiples of 4 (8, 12, 16 ...) are expressed by **zokkay** plus the appropriate UNIT (2, 3, 4 ...). Unfortunately, Bhat does not provide the word for '11' — is it **dos kase** or **zokkay-nəy katam**? Numerals between multiples of 4 are expressed by velar-prefixed forms of ONE, TWO, and THREE postposed to the next lower multiple of 4, so that the TEENS probably are as follows (the only teen actually to appear in Bhat is '14'):

THIRTEEN	zokkay-tam kase	"[4×3] + 1"
FOURTEEN	zokkay-tam kanəy	"[4×3] + 2"
FIFTEEN	zokkay-tam katam	"[4×3] + 3"
SIXTEEN	zokkay-brə	"4×4"
SEVENTEEN	zokkay-brə kase	"[4×4] + 1"
EIGHTEEN	zokkay-brə kanəy	"[4×4] + 2"
NINETEEN	zokkay-brə katam	"[4×4] + 3"

Note that in a quaternary language like this, the concept of TEEN is quite meaningless!

The inherited TB numerals for 6-9 seem to survive as ordinals: **thai-*do*-nia** '6th', **thai-*shni*-nia** '7th', **thai-*dang*-nia** '8th', **thai-*ne*-nia** '9th' [Revelation 21:19-20].

However, the quaternary system seems to be on the way out in Boro, judg-

---

117) Data from Pushpa Pai Karapurkar 1976.

118) Cf. the mysterious first syllable of the etymologically distinct though semantically similar WB form **khu'-hnac** '7', alongside **hnac** '2'.

119) According to Ian Maddieson (p.c. 1984), quaternary numeral systems are widespread in African cultures that hold markets every four days.

120) Other possible reflections of an original quaternary system are Abor-Miri-Dafla multiplicative formations for EIGHT of the form "4×2" (see below 4.237).

ing from the Bible translation [1972], where the teens are formed decimally: **khao-she '10'** (**-she '1'**), **khao-she-thai-she '11'**, **khao-she-thai-noi '12'**, **khao-she-thai-broi '14'**. Note the conjunctive marker **thai**, which certainly belongs with the root meaning *big*, reconstructed as PTB *tay in Matisoff 1985b ("God and the ST Copula" #68). Compare Tangkhul **kətay** 'be extra', **khəmətay** 'increase, multiply', **akətay** 'remnant'; Wancho **a-tai** 'far', **tai-hu** 'many', etc.

[D]　ABOR-MIRI-DAFLA

My data on teen formation in AMD is quite limited, though a couple of points may be noted.

ABOR-MIRI, GALLONG, and IDU form teens of the type TEN + LINK + UNIT, with cognate linking morphemes (**-lang-/-la-/-lo-**):

> Abor-Miri **eying-ko '10'**, **eying-ko-lang-ater-ko '11'**, **eying-ko-lang-akeng-ko '16'**, **eying-ko-lang-pinyi-ko '18'**
> Gallong **iri-go '10'**, **iri-go-la-ken '12'**, **iri-go-la-um '13'**
> Idu **hũ '10'**, **ho-lo-ke '11'** (**ke '1'**)

Note the relative prolixity of the Abor-Miri formation, where both the TEN and the UNIT morpheme take a prefix (**e-**, **a-**) and the suffix **-ko**. In Gallong only the TEN takes the suffix **-go**, while in Idu neither the TEN nor the UNIT morpheme is suffixed.

This suffix **-ko** is extremely interesting. In Abor-Miri **a-ko** is still one of the independent words for ONE (along with **ater-ko**), but AM has also developed it into a general suffix used with all numerals, exactly like the cognate Newari morpheme **-guː** (above 3.13). However, in Kaman (Miju Mishmi) there is a different word for ONE, and this **-ko** has taken on the meaning TEEN: **kumu '1'**, **kumu-ko '11'**, **kinin '2'**, **kinin-ko '12'** ..., providing us with one of our most striking instances of the "interchange and confusion between ONE and TEN". (See next section, 3.4.)

3.4　Interchange and Confusion between ONE and TEN[121]

ONE and TEN both occupy unique places in decimal systems. When two morphemes, one meaning TEN and the other meaning ONE, are juxtaposed in a compound numeral, it can either mean ONE *TIMES* TEN = 10; or ONE *PLUS* TEN = 11; that is, either the end of the first decade (...10), or the beginning of the second decade (11...).[122]

Put another way, both the UNITS (1-9) and the TEENS (11-19) can claim

---

121) See above 3.14 *t(y)ik; 3.152 *s(h)e; 3.155 *(k-)IV(N); 3.233 *s-riŋ ✕ *s-ryaŋ; also below 3.235, 3.546, 3.547. The Proto-Mayan forms for ONE (*xuːn) and TEN (laːxuːn; lit. "end of one") are obviously related morphophonemically and conceptually (p.c. Terrence Kaufman 1994). According to Ives Goddard (p.c. 1994), there is also ONE/TEN interchange in Algonkian.

122) See below 3.512-3.513 "Decimal multiplicative TWENTY".

"firsts" — the UNITS are the first group of ten above zero; but the TEENS are the first decade where two digits are required to express the numbers.

More facetiously, there is literally nothing (i.e. zero) distinguishing ONE from TEN.

3.41   ko/ka in Abor-Miri-Dafla, Himalayish, and Naga

Abor-Miri **a-ko** '1', **-ko** 'general numeral suffix' corresponds neatly both phonologically and semantically with Newari **-gu**: 'numeral suffix' (above 3.13c). But **-ko** means something quite different in Miju Mishmi, viz. -TEEN: **kumo** '1', **kumu-ko** '11'; **kinin** '2', **kinin-ko** '12'. The development in Miju probably was via the notion "one more time around; once more coming back to the unit ONE" — i.e. ELEVEN in a sense is "ONE and ONE", as its graphic shape *11* implies. All that separates "ONE + ONE" from "ONE + TEN" is one zero — and that's nothing much.

In exactly analogous fashion, Thulung Rai (E. Nepal) **ko** means '1', but also functions as the morpheme for -TEEN in the formation of the numerals from 11 to 19: **ko-nə** '12', **ko-sium** '13', **ko-gu** '19'.

Similarly, Lotha **ekha** means ONE as an independent numeral. As a formative in the three highest round numbers, however, it means TEN (i.e. -TY): **ti-ing** '7', **ekha-ti-ing** '70'; **tiza** '8', **ekha-tiza** '80'; **toku** '9', **ekha-toku** '90'.

3.42   mu- in Mishmi

In Das Gupta's "Miju Mishmi" [1977a], **kumu** means '1', but the first syllable **ku-** is a meaningless prefix, part of a 1-6 run. The root is **-mu-** 'ONE'. The word for TEN is **kyap-mo** ("10×1"; for the first syllable see 3.21). In Shail Kumari Dubey's Mishmi [1983], with a set of numerals vastly different from those reported by Das Gupta for Miju, TEN is **muou** (above 3.156), probably to be segmented **mu-ou**, with one syllable meaning ONE and the other meaning TEN — but which is which?[123]

3.43   tšek / ṣ(ə)- in rGyarong

In the Zida dialect of rGyarong (data from Kun Chang), '10' is ṣtṣi ; what seems to be the same initial element occurs in ṣətšek '11', ṣənes '12', and presumably all the higher teens as well. It could well be that this prefix is a reduced form of the independent numeral **tšek** '1', so that ṣtṣi meant "1 × 10" (i.e. < *tšek-tṣi); in the teens, however, where it cooccurs with the UNIT morphemes, the ṣə- then came to mean '10' (-TEEN). (Note that this analysis implies that an older form of '11' was something like *tšek-tšek, with subsequently greater and greater destressing of the first syllable.)

---

123) This is the same problem STC faced (p.94) in analyzing Moshang **rok-shi** '10' as "1 × 10", when actually its structure is "10×1". See above 3.22, 3.234.

3.44 **s(h)e* '1' and **ts(y)iy* '10'

Given the high degree of phonetic similarity between **s(h)e* '1' (above 3.152) and **ts(y)iy* '10' (above 3.22), as well as the organic semantic connection between the concepts ONE and TEN, it is not surprising that these etyma now seem inextricably intertwined.

3.5 Round Number Formation: Decimal and Vigesimal Systems

3.51 *Vigesimality, in Tibeto-Burman and elsewhere*

Is vigesimality a primitive characteristic? Consider that well-known primitive language, French: *quatre vingts* '80' ["4×20"], *quatre vingt un* '81' ["(4×20) + 1"], *quatre vingt dix* '90' ["(4×20) + 10"], *quatre vingt quatorze* '94' ["(4×20) + 14"], etc. Many Francophones outside of France (Belgium, Switzerland, Québec) sensibly prefer decimal alternatives to the higher round numbers, viz. *septante* '70', *octante* '80', *nonante* '90'.[124]/[125]

The Danish numeral system is even more vigesimal than the French, and is so interesting that it is worth presenting in some detail:
DANISH [126]

TEN	ti	SIXTY	tres(indstyve)
TWENTY	tyve	SEVENTY	halvfjerds(indstyve)
THIRTY	tredive	EIGHTY	firs(indstyve)
FORTY	fyrre(tyve)	NINETY	halvfems(indstyve)
FIFTY	halvtreds(indstyve)	HUNDRED	hundrede

The complications presented by this system are due largely to the word for TWENTY itself (**tyve**), which consists of the morpheme for TEN (**ti**) plus an element **-ve** which once meant TWO, but which has now lost its independent morphemic identity. This leaves the way open for a transvaluation of **tyve** from TEN(S) to TWENTY. There is no problem with THIRTY, which is simply "three [times] ten" (with voicing of the initial of the second constituent). With the higher round numbers 40-90, however, a tendency to drop the last element in the numeral has led to total loss of transparency in the system. The word for FORTY, originally a decimal multiplicative formation **fyrretyve** "four [times] ten"), has been shortened to **fyrre** (lit. "four"). The remaining even round numbers, SIXTY (originally **tresindstyve**, lit. "3 times tens") and EIGHTY (originally **firsindstyve**, lit. "4 times tens") were once segmentable as **tre-sinds-tyve**, **fir-sinds-tyve** (**sinds** "times"), but have now been shortened

---

124) Note a bit of orthographic pedantry here: for the round number '80', the plural grapheme *-s* is used after *vingt* (*quatre vingts*), emphasizing the multiplicative nature of the numeral ("four twenties"). For the odd numbers (81, etc.) the *-s* is omitted: *quatre vingt un*, *quatre vingt deux*. This distinction is hailed by educated Frenchmen (e.g. Gérard Diffloth) as a particularly subtle and powerful triumph of Gallic logic, and is obviously pushed hard in the French educational system.

125) Eric Hamp believes that French vigesimality reflects a Celtic substratal influence (p.c. 1994).

126) Data from Koefoed 1958. My thanks to Gary Holland for some further elucidation.

(except in very formal, emphatic speech) into **tres** and **firs**, with incorporation of the first consonant of **sinds** into the unit numeral.   The transvaluation of **tyve** is apparent from comparing FORTY and EIGHTY; in FORTY, **tyve** means "ten"; in EIGHTY, **tyve** means "twenty".   It is only the morphological difference between **fyrre** "four" and **firs** (with incorporated -s from the following syllable) that keeps the shortened forms distinct.   The higher odd round numbers (50, 70, 90) are expressed in an even more indirect and opaque way. With the last elements -**sinds-tyve** expressed, these numerical expressions at least make sense in terms of their constituent morphemes: FIFTY **halvtredsind-styve**, lit. "half-from- three times twenty", i.e. "two and a half times twenty"; SEVENTY **halvfjerdsindstyve**, lit. "half-from-four times twenty", i.e. "three and a half times twenty"; NINETY **halvfemsindstyve**, lit. "half-from-five times twenty", i.e. "four and a half times twenty".   When the last elements are omitted, however, one is left with the paradoxical vigesimal sequence **fyrre** '40', **tres** '60', **firs** '80', etymologically "4", "3" and "4", respectively — i.e. "four (tens)", "three (times [twenty])", "four (times [twenty])"!

In Tibeto-Burman, hesitation between decimality and vigesimality is apparent in several subgroups — Himalayish, Barish (e.g. Garo), Kuki-Naga. Some languages have both kinds of systems in more or less free variation, with the vigesimal one apparently older.[127]

It is common to find systems (e.g. in Kuki-Naga or Bodo-Garo) with a unitary monomorphemic word for TWENTY (like archaic English *score*), but where the higher twenties (40, 60, 80) are formed on the basis of TEN not TWENTY, e.g. English *eighty* (< EIGHT - TEN) vs. *fourscore*.   In GARO, for example, '20' is **kol-grik** ("20 × 1") and '30' is **kol-a-chi** ("20 + 10"), but from '40' on up the system becomes decimal, using the bound morpheme **sot-** '-TY' before the unit: **sot-bri** '40', **sot-bonga** '50', **sot-dok** '60'.[128]   Conservative speakers use a vigesimal system throughout (e.g. **kol-chang-gni** or **wakma-gni** '40').   [See PHILLIPS 1904; MOMIN n.d.; BURLING 1961]

In a "pure decimal" system (e.g. NOCTE, below 3.513) even the word for TWENTY is analyzable into TEN and TWO.   The maximal contrast is furnished by a "super-vigesimal" system, where even the word for HUNDRED is expressed as TWENTY times FIVE (below 3.524).

3.511   TWENTY as a unitary, unanalyzable morpheme

STC reconstructs a monomorphemic, unanalyzable word for '20', like Eng. *score*, of the shape *m-kul,[129] on the basis of the following forms:

---

127) French itself belongs to this "mixed" category of languages.   All its round numbers are decimal except for '80' ["4 × 20"] and '90' [" (4 × 20) + 10"].   The round number '70' (*soixante-dix*) ["60 + 10"] resembles '90' in its additivity, but '60' is not itself based on twenty.   For a discussion of "vigesi-decimal vacillation", see below 3.533.

128) This **sot-** is possibly related to Chin **som** '10' (above 3.231).

JINGPHO **khun**, GARO **khol** ～ **khal**, DIMASA **khon**, MIKIR **ingkol** ～ **ing-koi**, SIYIN **kul**, HAKA **kul** ～ **kwe**.[130]

There is no trace of this root in Lolo-Burmese or Karen. On the other hand, it is very widely attested in Kuki-Chin-Naga, where the nasal prefix is faithfully preserved (data mostly from Marrison ["GEM"] 1967:279):

Angami (Khonoma)	**meku**	Meluri	**mukwe**
Angami (Kohima)	**mepfu**	Nruanghmei	**ncui**
Ao (Chungli)	**metsu**	Ntenyi	**makweru/mukwung**
Ao (Mongsen)	**mukyi**	Pochury	**mke**
Chokri	**mechi**	Rengma	**nki**
Khoirao	**machi**	Sangtam	**mukyu**
Liangmai	**makai**	Sema	**muku**
Lotha	**mekwi**	Tangkhul	**makɯ** [BHAT 1969][131]
Mao	**makei**	Yacham-Tengsa	**machi/tamong**
Maram	**make**	Yimchungru	**muku**
Meithei	**kul**	Zeme	**nkai**

Yet, interestingly enough, even though *m-kul is so widespread in Kuki-Naga, it is not used to form the higher twenties (40, 60, 80) anywhere in the family (i.e. '40' is not "2 times *m-kul").

In the original version of STC, Benedict had reconstructed a distinct root *kun 'all' (#10, p. 18), based on WT **kun** 'all' and WB **kun** 'come to an end; be used up', **əkun** 'all'. In the footnotes and indices of the published version (pp. 15, 18, 202), he changed his mind, and decided to group these forms under *m-kul, an etymon now assigned the broader gloss 'all; twenty',[132] with the linking notion being "all the fingers and toes are used when counting up to this number".

Things may not be so simple, however. A number of languages, mostly Himalayish (cf. 3.533, 3.534) seem to reflect a prototype with *-a-, i.e. *-kal:

SHERPA **khal-jik**; JIREL **khalq**; TAMANG **khaːl**; KHALING (k)**haːel**; LEPCHA **k'a**; and perhaps SIKKIM BHUTIA **khe-chik**, SHARCHOP **khye**.[133]

To these we should probably add MONPA (Motuo) [AMD group] **k'ai**, as in **k'ai-ŋa** '100' ("20×5"), and perhaps also WANCHO **ca**. In a couple of other languages, the initial is a dental rather than a velar:

KAIKE [Him.] **thaːl**; MIJU [AMD] **katal-mo** (-**mo** 'one', ka- is a secon-

129) STC #397, pp. 15, 18, 83, 119, 120.
130) This etymon is discussed in Matisoff 1980, "Stars, moon, and spirits...", pp. 17-18.
131) Marrison (loc. cit.) has Tangkhul **maga**.
132) If we accept this, the Burmese forms for 'all' would cause us to modify our above statement to read "there is no trace of this root *with the meaning 'twenty'* in Lolo-Burmese".
133) Mazaudon [1985:154] cites several additional Himalayish forms, including Gongar (Bhutan) **khay** /khe/, Dungkarpa **khɛ**, Thakali **khal**, and Tamang ⁴**pokal**, as well as Tipra (=Kokborok) **khol**.

dary prefix).

It should be noted that these Himalayish and AMD forms deriving from *kal (unlike the KCN forms < *m-kul), are used in multiplicative formations for the "higher twenties" 40, 60, 80.

As Mazaudon points out in her excellent study of the Dzongkha numeral system [1985:136], the WT cognate **khal** provides the semantic key to these forms, at least as far as Himalayish is concerned. This WT form is glossed with two main meanings: (1) 'burden, load' (✕ **sgal** 'load of a beast of burden', p. 114); and (2) 'bushel; a dry measure equal to 20 *bre* ; therefore a score or twenty things of the same kind'. (Also possibly related is WT *sgal-pa* 'small of the back'.) [JÄSCHKE 1881:40].

Two explanations are therefore possible. Either we assume that two totally unrelated etyma are involved, one with -u- vocalism (*m-kul 'all; twenty') and one with medial -a- (*kal 'load; bushel measure; group of twenty'). Alternatively, we can posit an earlier allofamic connection between these two roots (*m-kul ✕ *kal), and claim that the semantic developments have all sprung from the same original meaning, e.g. "a complete load; everything that can be placed on a beast of burden at one time". The variational pattern -a- ✕ -u- is grudgingly recognized even in STC (e.g. #405 *b-suŋ ✕ *b-saŋ 'fragrant').[134]/[135]

There are a couple of other monomorphemic etyma for TWENTY of much more restricted distribution, which should be mentioned:

[A] MEITHEI -phu; WANCHO pu-[136]; PHOM pü- ~ bü- ~ pi- ~ bet-;

[B] YACHAM-TENGSA **tamong** and **mesung**. The former seems to be the same etymon that means FIFTY in several other languages (below 3.522, 3.524).

These etyma are used multiplicatively to form the higher twenties (including HUNDRED) [below 3.542].

3.512 Decimal multiplicative TWENTY, with the unit first
"20 = 2×10"

This is the universal pattern in Lolo-Burmese, e.g. Lahu **nî chi**.

---

134) Several other such etyma are discussed in VSTB [Matisoff 1978a:43-44].. See also our posited allofamic alternation *sam ✕ *sum 'three' (below 4.12). Note that the Garo alternants cited in STC (**khol** ~ **khal**) confirm the reality of the variational pattern with this root, as perhaps do the Tangkhul variants **maga** and **makɯ** (n. 130).

135) Aficionados of worldwide lookalikes will be pleased to know that the reconstructed etyma for TWENTY in at least two Mesoamerican language families bear a striking resemblance to our TB forms: Proto-Mayan *k'ahl (with *preglottalized initial) and Proto-Otomanguean *kala! (Personal communication, Terrence S. Kaufman, April 1994.)

136) The independent WANCHO word for '20' is **ca** or **tsa** (above). W.T.French [1983:572] reconstructs Proto-Northern Naga *ja on the basis of Wancho **tsa**, Konyak **ta**, Phom **ta**, and Chang **sau** (see below 3.524), but perhaps these may all be referred back to an earlier **k(y)al.

3.513　Decimal multiplicative TWENTY, with the unit second
"20 = 10×2"

(a)　Where the morpheme for TEN is the same as the independent numeral '10':

	2	10	"-TY" 20	30	40	50
PUIRON	kani	som	som-	somni	somthum	somli	somnga

(b)　Where a special combining form for TEN is used, that has no etymological relationship to the independent numeral:

	2	10	"-TY" 20	30	40	50
NOCTE	vanyi	ichi	ruak-	ruaknyi	ruakram	ruakbeli	ruakbanga
MARING	khani	chip	som-	somni[137]	somthum	somli	somnga

Note that there are many Kuki-Chin languages where **som** is the independent numeral '10' (e.g. Lushai). In Maring, however, it is a bound morpheme occurring only in the round numbers from 20 to 90.

In this kind of PURE DECIMAL system, '20' is treated the same as all the other round numbers from 30 to 90. In languages where TEN is expressed as "10×1", TWENTY is of course also expressed as "10×2". See above 3.4, below 3.5213.

The Central Chin language LAKHER (= MARA) has no fewer than four alternative expressions for TWENTY (**mia-ki, sy-no, sa-ki, hlei-hraw**), each formed according to a different pattern:

LAKHER

  mia-ki 'twenty'
  　　　　mia- is a general prefix used with all numerals; **ki** is the inherited monomorphemic root;
  sy-no 'twenty'
  　　　　sy- '10; -TY', as in **sy-pali** '40'; **no** '2': "10×2"
  sa-ki 'twenty'
  　　　　sa- '1'; **ki** = '20' : "1×20"
  hlei-hraw 'twenty'
  　　　　-hraw '10'; **hlei** seems to mean 'pass, exceed, be extra'[138]

3.52　Round number formation in Kuki-Chin-Naga

3.521　Decimal systems of round number formation in KCN

3.5211　Where the morpheme for TEN in the compound numerals is the same as (or a morphophonemic variant of) the independent numeral '10'

---

137) Marrison (p. 279) has "somnga" for '20', though this certainly seems to be an error, since **somŋa** is glossed as '50' on p. 79. The form **somni** is my own guess.

138) This derives from a PTB root set up as *s-lay ⪥ *s-ley [MATISOFF 1985b:#58]. It is used in Lakher as a linking morpheme in teen-formation, e.g. **pa-hraw hlei no** '12' ("10 + 2").

	10	30	40	60	70	80	90
YIMCHUNGRU	thürü	samrü	yirü	rukrü	nierü	zharü	kurü

Yimchungru has a special word for '50' (below 3.522).

3.5212 Where the morpheme for TEN in the compound numerals is etymologically unrelated to the independent numeral

Here we must make a further distinction:

(*a*)   *Where 30 behaves differently from 40-90*

In many Kuki-Naga languages all the round numbers from 30 to 90 are multiplicative decimal constructions, but the formation of '30' is different from '40' and above; that is, THIRTY is expressed as *3 × 10*, with the morpheme for TEN based on the independent numeral '10', and the UNIT morpheme PRECEDING this TEN morpheme; *but* '40, 50...90' are expressed as *10 × 4, 10 × 5...10 × 9*, with the UNIT morpheme *following* this TEN morpheme. The morpheme for TEN used in composition is usually etymologically distinct from the independent numeral.

	10	3	30	"-TY"
ANGAMI (Khonoma)	keru	se	serü	lhi-
ANGAMI (Kohima)	keru	se	serü	hie-
CHOKRI	küri	sü	——	hie-
KEZHAMA	chiro	katsu	——	lha-
KHOIRAO	sara	kathum	thumra	ra-/re (k)
LIANGMAI[139]	kariu	shum	samriu	ri(a)-
MAO	chüro	kosü	shüro	ri-
MARAM	kero	hangtum	tumru	rag-/re (k)
MZIEME	kerei	ketsum	tsamrei	riak-
NRUANGHMEI	ruh	kathum	tümru	rek-
RENGMA	tsarü	keshan	shenrü	en-
SEMA[140]	chüghi	küthu	sheghi	lho-
TANGKHUL	thara	kathum	thumra	hang-
ZEME	kereu	kechum	himreu	he-/ re-/ riak-[141]

	40	50	60
ANGAMI (Khonoma)	lhida	lhipengu	lhisuru
ANGAMI (Kohima)	hiede	hiepengou	hiesorou
CHOKRI	hieda	hiepungu	hieshwuru
KEZHAMA	——	lhapangu	——

139) Liangmai '40' does not exactly fit the pattern; it has the prefix a- instead of ri(a)- ['4' is madai]; '70, 80, 90' all have ria- (riachania, riatachad, riachakiu).

140) Note küthu 'three' ≠ she- 'thir-'.

141) Zeme '70, 80, 90' have riak: riaksena, riakdesat, riaksekui.

KHOIRAO	ramri	renga	reksaruk
LIANGMAI	atai	ringiu	riacharuk
MAO	ridei	ripongo	richoro
MARAM	ragdai	rengo	reksaruk
MZIEME	riakdai	riangngei[142]	riakheruk
NRUANGHMEI	rekdai	rekngu	rekcüruk
RENGMA	henzi	hempfü	hentsaro
SEMA	lhobdhi	lhopongu	lhotsogho
TANGKHUL	hangmati	hangphanga	hangtharuk
ZEME	hedai	rengeu	riakseruk

(b)　*Where 30 behaves the same as the higher round numbers*

	10	*3*	*"-TY"*	*30*	*40*	*50*
NOCTE	ichi	vanram	ruak-	ruakram	ruakbeli	ruakbanga

In Nocte, however, not only is '30' formed the same as the higher round numbers, but so is '20'; i.e. instead of a monomorphemic word for '20', it too is composed of the special morpheme for TEN[143] (along with the unit morpheme TWO): **ruaknyi** '20'.

3.5213　*Where TEN is a classifier, not a numeral*

In this pattern, the number '10' itself is treated as a multiplicative construction "1 × 10", so that '10' is structurally identical to the higher round numbers '20', '30'...

(a)　This structure is common in Lolo-Burmese (see above 3.32[B]):

	1	*10*	*2*	*20*	*3*	*30*
LAHU	tê	tê chi	nî	nî chi	šɛ̂ʔ	šɛ̂ʔ chi

tê chi	'10' (Num + Clf)	"1 × 10"
nî chi	'20'	"2 × 10"
šɛ̂ʔ chi	'30'	"3 × 10"...

(b)　It is also found in at least one Naga language:

	1	*10*	*2*	*20*	*3*	*30*
TANGSA (Moshang)	ashi	rokshi	ani	rokni	atum	roktum
TANGSA (Yogli)	ashi	raukshi	anei	rauknei	adim	raukdim

The difference between the Tangsa and Lahu cases is simply one of word order. In Lolo-Burmese the numeral precedes the classifier TEN; in Tangsa the TEN precedes the numeral.

In most of Kuki-Naga the word for TEN is not analyzable into two morphemes, ONE and TEN, but merely consists of a unitary root for TEN preceded by a meaningless prefix.

---

142)　Note the assimilation of the final of **riak-** to the nasal root initial in FIVE.

143)　For the etymology of this special combining form for TEN, see above 3.233 (c).

It is noteworthy that in languages where TEN is expressed as *"ten times one"*, TWENTY is also expressed as *"ten times two"* — i.e. there is no unanalyzable monomorphemic word for '20'.

3.522   Languages with a special word for FIFTY

A number of Kuki-Chin-Naga and Barish languages have a disyllabic (but not easily analyzable) form for FIFTY, which is quite distinct from their words for FIVE, TEN, or TWENTY:

	5	10	20	50
AO (Chungli)	pungu	ter	metsü	tenem
AO (Mongsen)	phanga	tera	mükyi	tünam
LOTHA	mungo	taro	mekwi	ti-ingya[144]
MELURI	manga	tera	mükwe	teni
NTENYI	münga	dagha	mekweru	teni
POCHURY	mnga	türa	mke	tünie
SANGTAM	münga	thüre	mükyü	thünyang
YIMCHUNGRU	phüngü	thürü	muku	thünim
DIMASA	bonga	ji	khon	dan

The words for FIFTY in these languages have first syllables that begin with a dental stop, and later syllables that contain a nasal group (-n-, -ngy-, -ny-), and sometimes 2 nasals (**tenem, thünyang**).

Now it is reasonable to suppose that a disyllabic word meaning FIFTY should usually consist of components that mean TEN and FIVE (ordered either *10×5* or *5×10*).

(a)   *First syllable:*

There is some basis for hypothesizing that it is the first syllable in these forms that means TEN. Note that the independent numeral '10' in all these languages (except Dimasa) has a prefix with dental stop.[145] However, these eight languages have generalized a dental prefix with the numerals all the way from 6 to 10,[146] so we would have to suppose that an originally meaningless prefix came to take on the semantic value of the highest numeral with which it appeared (TEN). Lest this seems too farfetched, we shall soon see (below 3.523) how in Mikir the morpheme **throk**, etymologically 'SIX', has come to mean 'TEN'.

(b)   *Second syllable:*

We might suppose that the second syllable is some allofam of the ordinary

---

144) Cf. Lotha **ti-ing** '7', **ekhati-ing** '70'.

145) The Dimasa form **dan** looks as if the rhyme of an original second syllable was apocopated, so that the former prefix amalgamated with the former second syllable's initial to form a stressed monosyllable. See below 4.1411.

146) A slight exception is Sangtam, which lacks a dental prefix in '8' (**thuro, thunye, ke, tuku, thure**). See below 5.44.

numeral for FIVE *b-ŋa or *m-ŋa. This is not implausible phonologically, since *m-ŋa has two nasals, like the second element in many of the modern forms.

If this interpretation is correct (and it is certainly open to question[147]) we must note that the order of the constituents in FIFTY (TEN — FIVE) is the reverse of that for almost all the other round numbers from '30' to '90' in these languages (UNIT — TEN).[148]

	30 (3 × 10)	40 (4 × 10)	50 (10 × 5 ?)	60 (6 × 10)
AO (Chungli)	semer	lir	tenem	roker
AO (Mongsen)	samra	lira	tunam	rokra
LOTHA	*thamdro*	zuro	*tiingya*	rokro
MELURI	chera	zura	teni	rora
NTENYI	chagha	jugha; zua	teni	*apyampero*
POCHURY	chera	zura	tunie	rora
SANGTAM	sangre	zyure	thunyang	rore
YIMCHUNGRU	samru	yiru	thunim	rukru

	70 (7 × 10)	80 (8 × 10)	90 (9 × 10)
AO (Chungli)	neter	tir	tukur
AO (Mongsen)	nira	*lira-anekhi*	telangtuku
LOTHA	*ekhatiing*	*ekhatiza*	*ekhatoku*
MELURI	rura	zera	khura
NTENYI	*apyamtughu*	*apyamtuza*	*apyamtukhu*
POCHURY	rura	zera	kuru
SANGTAM	nyure	*zyurereanyu*	kure
YIMCHUNGRU	nieru	zharu	kuru

On the other hand, we may be barking up the wrong tree here. Perhaps these problematic words for FIFTY come from an entirely different semantic field. After all there is something intrinsically special about FIFTY, since it is situated at the midpoint of the nine two-digit round numbers, with four below (10, 20, 30, 40) and four above (60, 70, 80, 90).[149] A language that is instructive here is MEITHEI, which also has a special word for '50', though it does not begin with a dental prefix: **manga** '5', **tara** '10', **kul** '20', **yangkhei** '50'. According to Purna Chandra Thoudam, a native speaker, Meithei **yangkhəy** '50' "might have some affiliation with **yaŋ** 'backbone; middle of the back' or 'mid-

---

147) Among the puzzling sidelights here is the similarity between FIFTY (**ti-ingya**) and SEVEN (**ti-ing**) in Lotha. Is this merely accidental? See below 4.228.

148) This is similar to the case of those languages discussed above (3.5212a) where it was the word for THIRTY whose constituent order was out of step with all the higher round numbers.

149) *Wednesday* occupies an analogous position with respect to the other days of the week (cf. German *Mittwoch*, lit. "mid-week").

dle of the roof in houses' [i.e. 'ridgepole']."[150] It is possible that forms like Sangtam **thunyang** or Lotha **tiingya** are also to be analyzed as containing this element in their second syllables (**thun-yang, tiing-ya**).

　　Although these languages all have a unitary word for TWENTY, they do not form their "higher twenties" (40, 60, 80) vigesimally. These are either straightforward decimal formations, or else present other complications that have nothing to do with "twenty" as a structural unit:

- In Sangtam '80' is expressed multiplicatively as "40 (**zyure**) ×2 (**nyü**)": **zyure-re-anyü**. Ao Mongsen has an identical formation for '80': 40 (**lira**) ×2 (**anet**): **lira-anekhi**.

- The Ao Mongsen word for '90' contains a special allomorph of '10' (**telang**). The independent Mongsen numeral '10' is **tera**, with no final nasal. Note that **telang-tuku** has the structure "10×9", unlike all the other Mongsen round numbers (except of course '50'), which have the UNIT before the TEN.

- In Ntenyi, 60-90 have a special morpheme **apyam-** which must mean '10', since it is followed by the unit morphemes 6-9 (**-pero, -tughu, -tuza, -tukhu**). However, the '6' in '60' (**-pero**) is not the same as the independent numeral **togho**, though clearly related to it (**-ro ≍ -gho**). It is possible that this **apyam-** is related to the root *(**b**)**an** which appears as the independent numeral for TEN in several languages (above 3.232). In Lotha also, 70-90 contain a special morpheme **ekha** '10', distinct from the independent numeral (**taro**). What is especially interesting here is that **ekha** is also the ordinary Lotha word for ONE. This is a prime example of the interchangeability of ONE and TEN that we have already discussed (above 3.4). An alternate Lotha form for '80' (**zaro**) follows the simple regular pattern of the root for EIGHT (**za-**) plus the ordinary root for TEN (**-ro**).

- I am at a loss to explain Lotha **ti-ingya** '50',[151] which looks as if it has been influenced or contaminated by **ti-ing** '7'. Why '7' and '50' should enjoy a special relationship remains obscure (even though '50' is one more than 7-squared).

- Lotha **thamdro** '30' shows an interesting epenthetic **-d-** intervening between the **-m** of THREE and the **r-** of TEN. The position between a nasal and a liquid is a classic locus for an epenthetic stop.[152]

- The Ao Chungli round numbers (except '20' and '50') are all formed

---

150) Personal communication, July 18, 1988. This metaphor makes especially good sense if one thinks of the horizontal backbone of an animal on all fours. A more specific Meithei compound for 'backbone' is **yaŋ-len sәru** (where the last constituent means 'bone'). A rough analogy to this intrusion of a word from an outside semantic field into a system of round numbers is Russian **sórok** '40', which is said to derive from a word meaning "a batch of fur pelts."

151) It is perhaps to be analyzed as **tiing-ya** (above).

152) Cf. French *chambre* 'room' < Latin *camera* (Vulg. Lat. *camra*).

multiplicatively of UNIT × TEN. The TEN morpheme used as a combining form is -(e)r, a reduced version of the independent numeral **ter** (which itself is an apocopated form, as witness Ao Mongsen **tera**). Of special interest are the monosyllabic forms **lir** '40' and **tir** '80'; this parallel between '40' and '80' is also evident in Ao Chungli, where '80' is actually expressed as "40 × 2", perhaps to avoid the near-pernicious homophony found in Mongsen. We should note that the Chungli independent numeral '4' is **pezü**; a more ancient-looking form (< PTB *b-ləy) now survives only in **lir** '40'.

3.523 The case of MIKIR

The round number system of Mikir presents special problems because of the polymorphemic (additive or subtractive) structure of the independent numerals 7-9:

'1'	isi	'6'	throk	'60'	throk-kep
'2'	hini	'7'	throksi ("6 + 1")	'70'	throksi-kep
'3'	kethom	'8'	nirkep ("10 − 2")	'80'	throk-hir-kep
'4'	phir/phli	'9'	sirkep ("10 − 1")	'90'	throk-sir-kep
'5'	phongo	'10'	kep	'100'	pharo

The round numbers from '30' to '60' are simply formed decimally, with the independent numeral **kep** following the UNIT morpheme: **thom-kep, phli-kep, phongo-kep, throk-kep**. As one would expect, '70' has a similar structure: **throksi-kep**. However, with '80' and '90' the language has a severe problem. Since EIGHT and NINE already end in **-kep**, if their corresponding round numbers were formed "regularly" we would get *nirkep-kep and *sirkep-kep. Instead what we find is **throk-hir-kep** '80' and **throk-sir-kep** '90'.

This form for '80' is readily understandable. The first two syllables **throk-hir-** are an additive expression for '8' ("6 + 2"), where the morpheme for '2' is the same as the first syllable of the independent numeral **hini**. (Note the difference from the ordinary subtractive expression for '8' ("10 − 2"), where the morpheme for '2' is the same as the *second* syllable of **hini**.)

The form for '90' is more difficult to explain. The first two syllables **throk-sir-** do not stand in an additive relationship, as they do in **throksi-kep** '70'. Morpheme by morpheme the three syllables mean SIX - ONE - TEN — and there is no way these can be juggled to yield '90'. It looks to me as if the morpheme **throk-** '6', since it appears in '60', '70', and '80', has been included in '90' as well, where it has acquired the meaning TEN by a process of false analogy! If this is what has happened, we should interpret '90' as meaning "(10 − 1) × 10", i.e. "9 × 10."[153]

---

153) Something rather similar seems to have occurred with Lotha **ekha**, which means ONE as an independent numeral, but functions like TEN in some higher round numbers (above 3.522). For more on the Mikir system, see below 5.421.

3.524   Vigesimal systems of round number formation in Kuki-Chin-Naga

Several Kuki-Naga languages express the even round numbers '40', '60', '80' as multiples of TWENTY. As noted above, however, the ordinary independent numeral '20' ($<$ *m-kul) is not used for this purpose; instead we find special forms which appear only in composition. In fact most of these languages do not have a reflex of *m-kul at all, and show different roots for the independent numeral TWENTY. In languages of this type, the odd round numbers '30', '50', '70', '90' are typically expressed additively or subtractively in terms of the next lower or higher multiple of TWENTY. That is, '70' may either be "$(20 \times 3) + 10$" (i.e. $60 + 10$) or "$-10 + (20 \times 4)$" (i.e. $80 - 10$).[154]

In what we might call "super-vigesimal" languages (e.g. Wancho and Chang in the following chart), the word for HUNDRED is also expressed in terms of TWENTY ("20×5") — i.e. HUNDRED is morphologically complex.

	WANCHO	*PHOM*	*KONYAK*	*CHANG*
TEN	ban	an	pen	an
TWENTY (indep.)	tsa/ca[155]	ha	ta	*sauchie*
TWENTY (in comp.)	pu-	pü- ~ bü- pi- ~ bet-	ta- ~ te-	sau-
FORTY	punyi	pinnyi	*teija*	saunyi
SIXTY	puram;hujam	püjam	*telemja*	sausam
EIGHTY	puli	büali	*tepelija*	saulei
THIRTY	ca-ban	ha-püan	ta-pen	*kujih*
FIFTY	punyi-ban	pinyi-püan	teija-pen	anchinsam
SEVENTY	puram-ban	betjam-püan	telemja-pen	anchinlei
NINETY	puli-ban	büali-püan	tepelija-pen	anchiningau
HUNDRED	puga	gho	kho	saungau

The Konyak higher twenties ('40', '60', '80') are each expressed by *three* morphemes instead of just two:

te- '20' $\times$ **-i-** '2'/**-lem-** '3'/**-peli-** '4' $\times$ **-ja** '1'.

That is, '40' $= 20 \times 2 \times 1$; '60' $= 20 \times 3 \times 1$; '80' $= 20 \times 3 \times 1$. This semantically otiose one-factor serves to add a bit of redundancy to these numerals, and shows a certain mathematical sophistication.

In Chang the morpheme **sau-** '20' does not occur independently, and is treated as a classifier; thus **sauchie** means literally "20×1". Chang is super-vigesimal, in that HUNDRED is also expressed in terms of TWENTY ("20×5"). W.T. French derives the form **sau-** from Proto-Northern Naga *ja (above 3.511), though it bears a striking superficial resemblance to the Northern Thai (Kham Myang) word **saaw** '20'. Is it possible that it could be a loan from

---

154) Cf. French **quatre vingts** '80', **quatre-vingt-dix** '90' ("$[4 \times 20] + 10$"), and the Danish system discussed above (3.51).

155) As Das Gupta [1979:28] puts it: "(Wancho) **ca** and **pu** both indicate '20'; **ca** is used when the number is indicated by *adding* the numeral, and **pu** when *multiplying* the numeral."

Ahom?

Chang **kujih** '30' is quite mysterious, with no apparent resemblance to **sam** '3', **an** '10', or **sau-** '20'. It is possible that the first syllable **ku-** is from *m-kul '20', which does not otherwise seem to survive in Chang. If that is true, then **-jih** must mean '10', though its affiliations are uncertain at this time. Chang FIFTY, SEVENTY, NINETY look like *subtractive* forms, based on the next higher multiples of twenty ('60', '80', '100').[156] The final elements **-sam, -lei, -(i)ngau** are the morphemes '3, 4, 5'; the first syllable **an-** is '10'. We can only assume, therefore, that **-chin-** is another bound morpheme meaning '20', in complementary distribution with **sau-** which only appears in the *even* round numbers. Thus **an-chin-sam** '50' would mean "10 (from) 20 times 3", i.e. "60 minus 10"; **an-chin-lei** '70' is "10 (from) 20 times 4", i.e. "80 minus 10", etc.

Finally, consider the complicated vigesimal systems of Meithei (=Manipuri) and Yacham-Tengsa (N. Naga group):

	MEITHEI	*YACHAM-TENGSA*
TEN	**tara**	**thelu**
TWENTY (independent)	**kul**	**machi/tamong**
TWENTY (in composition)	**-phu**	**machi-/tamong-/mesung-**
FORTY	**niphu**	**mesung-anat**
SIXTY	**humphu**	————
EIGHTY	**mariphu**	**tamong-phule**
THIRTY	**kun-thra**	**machi-li-thelu**
FIFTY	*yangkhei*	**tamong-anat-tule-thelo**
SEVENTY	**humphu-tara**	**tamong-asam-tule-thelo**
NINETY	**mariphu-tara**	*tamong-phuicu-le-thelu*
HUNDRED	**cha**	**mesung-phung**

- In Meithei the combining form for '20', *-phu, follows* the UNIT morpheme (**niphu, humphu, mariphu**), unlike the cognate morphemes in Wancho (**pu-**) and Phom (**pü-**), which *precede* the UNIT.

- Meithei **kun-thra** '30', consists of a variant of the independent word for '20' (**kul**) plus a variant of '10' (**tara**). In this form the order of the semantic components is the opposite of that in '40','60','80'. If Meithei '50' were formed like '70' and '90', it would be *niphu-tara.

- We have already noted that Meithei **yangkhei/yaŋkhəy** '50' is probably an intruder from another semantic field, with an original meaning related to 'backbone; ridgepole; midpoint'.

- The Yacham-Tengsa system of round numbers is perhaps the most cumbersome to be found in all of Tibeto-Burman. Marrison [1967:279] gives two different independent forms for '20', **machi** (< *m-kul) and **tamong**. (This latter form looks suspiciously like our special root for FIFTY, above 3.522!)

---

156) See below 4.20.

Yet it is clear that a third morpheme for '20', **mesung-**, is also used in composition: **mesung-anat** '40' (**anat** '2'); **mesung-phung** '100' (**phungu** '5'). In '30' **machi-** is used — this is a slight exception to our generalization that no descendants of *m-kul occur in the higher round numbers (it is still true that they do not occur at all in the higher multiples of 20). In '50', '70', '80', and '90', the morpheme meaning '20' is **tamong-**. Inexplicably the Yacham-Tengsa form for '60' is missing from Marrison [1967:232], so we do not know whether it is **mesung-asam** or (as I suspect) **tamong-asam**. '30', '50', '70', '90' are additive formations based on '20', '40', '60', '80', respectively, with **-li- ∼ -(tu)le-** serving as a linking morpheme. The form **tamong-phuicu-le-thelu** '90' is puzzling (we would expect **tamong-phungu-(tu)le-thelu**) — in fact it is very possible that "*phuicu*" is just a typo for **phungu**.

3.53   Round number formation in Himalayish[157]

In Himalayish languages with vigesimal systems, reflexes of the unitary lexeme *m-kul ⋊ *kal are used to form higher multiples of '20'. This is different from Kuki-Naga, where this etymon is only used for TWENTY itself, with the higher multiples expressed in some other way.

3.531   Systems under heavy influence from Nepali

In languages like Hayu, Thulung Rai, Kham, and Sunwar, the TB numerals are barely preserved for 1-3 or 1-4, let alone anything higher! (See above 2.11.)

It is worth noting that the Nepali higher numerals have a high degree of morphemic opacity. As is characteristic of Indo-Aryan in general, considerable morphophonemic changes are undergone by the simple numerals when they appear in compounds: e.g. Np. **cha, saːṭh, aːṭh, nau** '6-9', but **saːṭhi, sattari, asi, nabbe** '60-90'. Although these were probably not as easy to learn as the original, morphemically transparent TB-derived higher numerals had been, such is the cultural and economic power of Nepali that certain minority peoples seem to have had no alternative.

3.532   Decimal systems of round number formation in Himalayish

A few Himalayish languages have strictly decimal systems, including two of great cultural importance (Tibetan and Newari), and two (but not all three) members of the Gurung-Tamang-Thakali trio:

---

157) The data on Chepang, Gurung, Jirel, Kaike, Khaling, Kham, Magari, Newari, Sherpa, Sunwar, Tamang, and Thakali are from Hale (ed.) 1973, Part IV ("CSDPN"). Hayu is from Michailovsky and LSI III.1 (384-385); Tibetan from several sources. Other sources include Allen 1975 (Thulung Rai), Gvozdanović 1985 (Kiranti languages), Joshi/Rose 1909 (Kanawari), Lu Shaozun 1986 (Cuona Menba), Mainwaring and Grünwedel 1898 (Lepcha), Mazaudon 1985 (Dzongkha), Chhewang Rinzin 1984 (Dzongkha, Sharchop), and Sandberg 1895 (Sikkim Bhutia).

	10	3	30	4	40	5	50
WR.TIB.	bću	gsum	sum-ću	bźi	bźi-bću	lŋa	lŋa-bću
THAKALI	cyu	som	som-cyu	plih	plih-cyu	ngah	ngah-cyu
GURUNG	cyuq	soq	soq-jyu	plihq	plih-jyuq	ngahq	ngah-jyuq

The NEWARI case is less transparent morphophonemically, but identical structurally. In the Newari round numbers, the independent numeral **jhi** '10' appears in attenuated form as **-i** (which then undergoes various further morphophonemic adjustments):

2	**ni-gu:**	20	**ni:-gu:**[158]	/ < ni + i/
3	**swa-gu:**	30	**swi:-gu:**	/ < swa + i/
4	**pe-gu:**	40	**pi:-gu:**	/ < pe + i/
5	**nya:-gu:**	50	**nyae-gu:**	/ < nya: + i/
6	**khu-gu:**	60	**khwi:-gu:**	/ < khu + i/
9	**gu-gu:**	90	**gwi:-gu:**	/ < gu + i/

In all these languages the word for TWENTY itself is expressed in terms of TEN, i.e. "2×10": WT **nyi-śu**, Thakali and Gurung **ngih-syu**, Newari **ni:-** / < ni + i/. Note that the unit morpheme precedes the TEN morpheme in all these systems.

3.533 *Vigesimal systems of round number formation in Himalayish*

Quite a number of Himalayish languages have strict vigesimal systems, where the even round numbers are expressed as multiples of 20, and the odd round numbers are additive or subtractive with respect to a neighboring multiple of 20:

[A] CUONA MENBA (Mama [Southern] dialect: Lu Shaozun 1986:184-185)

ten	tɕi^{53} ~ tɕiŋ^{55} ~ tɕip^{53} ~ tɕik^{53} [159]	
twenty	khA^{55}li^{55}	
forty	chɛʔ^{53}-nAi^{53}	("20×2")
sixty	chɛʔ^{53}-sum^{53}	("20×3")
eighty	chɛʔ^{53}-pli^{53}	("20×4")
hundred	chɛʔ^{53}-le^{31}ŋe^{53}	("20×5")
thirty	khA^{55}li^{55}-tɕi^{53}	("20 + 10")
fifty	chɛʔ^{53}-nAi^{53}-tɕi^{53}	("[20×2] + 10")
seventy	chɛʔ^{53}-sum^{53}-tɕi^{53}	("[20×3] + 10")
ninety	chɛʔ^{53}-pli^{53}-tɕi^{53}	("[20×4] + 10")

In this dialect the word for '20', **khA^{55}li^{55}**, looks like a "dimidiated" (i.e. disyllabified) derivate of ***kal**, with secondary final vowel. In the words for the higher even round numbers (40-100), a different morph for '20' appears, **chɛʔ^{53}-**,

---

158) **-gu:** is a suffix which occurs with all Newari numerals (above 3.13), and is irrelevant to the present discussion.

159) The final of this morpheme assimilates to the initial of the following unit numerals in teen-formation: **tɕi^{53}theʔ^{53}** '11', **tɕiŋ^{55}nAi^{35}** '12', **tɕik^{53}sum^{53}** '13', **tɕip^{53}pli^{53}** '14'.

though this may actually represent a co-allofam of the same etymon, perhaps
*kyal (< ***kal** + **i** ?). (In the other dialect treated by Lu Shaozun, Wenlang
[Northern Cuona], the simple form for '20' is also **khɑ⁵⁵li⁵⁵**, but the variant that
occurs with 40-100 is **khʌi⁵⁵**.) The odd round numbers (30-90) are additive
formations based on the next lower multiple of 20.

[B]   TAMANG and SHARCHOP/TSANGLA

	TAMANG	*SHARCHOP*
ten	ci	she
twenty	kha:l-ki:h	khye-thor
forty	kha:l-nyi:h	khye-nyiktsing
sixty	ha:l-som	khye-sam
eighty	kha:l-pli	khye-pshi
hundred	kha:l-nga:h	khye-nga
thirty	kha:l-ki:h-syi-ci	khye-thor-dang-she
fifty	kha:l-nyi:h-syi-ci	khye-nyiktsing-dang-she
seventy	kha:l-som-syi-ci	khye-sam-dang-she
ninety	kha:l-plih-syi-ci	khye-pshi-dang-she

Tamang and Sharchop have true vigesimal formations from 20 upward,
including 100. (Unlike Cuona Menba, these languages express '20' itself as
"20 × 1"). Tamang apparently goes so far as to maintain the vigesimal system
all the way up to 1000, which is expressed as **kha:l-paca:s** ("20 × 50"), with the
second element from Nepali. (Gurung and Thakali, so closely related to
Tamang, are completely decimal, further demonstrating that decimality vs.
vigesimality is a useless criterion for linguistic subgrouping.)

[C]   KAIKE and JIREL

	JIREL	*KAIKE*
10	cyutambaq	chyu
20	[nye:syu]	[ngi-chyu]
40	khalq-nyiq	nghe-tha:l
60	khalq-sumq	sum-tha:l
80	khalq-syi	li-tha:l
100	[sei-jyiq]	nga:-tha:l
30	khalq-jyik-tangq-cyutambaq	nhi-chyu-chyu
50	khalq-nyiq-tangq-cyutambaq	phera:ng sum-tha:l
70	khalq-sumq-tangq-cyutambaq	phera:ng li-tha:l
90	khalq-syi-tangq-cyutambaq	phera:ng nga:-tha:l

Jirel and Kaike express '20' itself with an obviously innovative decimal
formation, "2 × 10". Jirel uses the inherited monomorphemic TB **khalq** for
everything above 20 (e.g. **khal-jik-tangq-nyiq** '22'), but Kaike uses **nhi-chyu**
throughout the twenties (e.g. **nhichyu-chyu-di** '31' ("20 + 11"), **nhichyu-
chyur-gu** '39' ("20 + 19"), and does not use **tha:l** (< ***kal** ≍ ***m-kul**) until
'40'.

Kaike differs from these other languages in two important respects. While Tamang, Sharchop, and Jirel all form multiples of 20 by putting TWENTY before the UNIT ("20×1,2,3..."), Kaike puts the UNIT before the TWENTY ("1,2,3...×20"). In the first three languages, the odd round numbers are formed *additively* from the next lower multiple of 20. In Kaike, they are formed *subtractively* from the next higher multiple of 20. The morpheme **phera:ng** is used to express "(minus) 10" in these expressions. (It may be related to the forms discussed above, 3.233c.) Jirel uses a Nepali borrowing for '100'.

[D] KANAWARI

Kanawari is a rather well-behaved vigesimal language, though like Kaike and Jirel it has a decimal multiplicative form for '20' itself: **ni-ja** ("2×10"), where **-ja** is apparently a variant of **sai** '10', and **sa- ~ so-** means '-teen'. The even multiples of '20' are expressed in a normal manner, with the UNIT preceding TWENTY: **ni-nija** '40', **shum-nija** '60', **pü-nija** '80'.

The odd round numbers, however, present some interesting peculiarities:

'30'　**ḍe'-nija**
'50'　**ḍai-nija**
'70'　**sa'e-shum-nija**
'90'　**sa'e-pü-nija**

According to Joshi/Rose, the **ḍai-** in '50' is a loan from Hindi meaning 'two and a half': 50 = 2 1/2×20 ! THIRTY seems to be additive, with **ḍe'** '10' evidently also from Indo-Aryan. The morpheme **sa'e-** in '70' and '90' must also mean '10': 70 = 10 + (3×20); 90 = 10 + (4×20). The etymology of this **sa'e-** is still in doubt, however, there being several possibilities, including its being related to the independent Kanawari **sáí** '10' (above 3.22).

[E] MAGARI and KHALING

These are both thoroughgoing vigesimal languages, though they show strong Nepali lexical influence that will undoubtedly lead to future erosion of their original numeral systems.

Magari has borrowed the word '20' itself from Nepali (**bis**), but the structure of its higher round numbers is completely vigesimal. The even ones are expressed as multiples of 20 (**nis-bis** '40', **som-bis** '60', **ca:r-bis** '80'); while the odd ones are additively based on the next lower multiple via the Nepali morphemes **-e-das** (**das** '10'): **bis-e-das** '30', **nis-bis-e-das** '50', **som-bis-e-das** '70', **buli-bis-e-das** '90'. Note that the native TB numeral **buli** '4' survives only in '90', while it has been replaced by Nepali **ca:r** in '80' itself. This demonstrates that the pressure exerted by Nepali on the system is "from the top down", from the higher numerals to the lower. The independent word for HUNDRED in Magari is a Nepali loan (**say**), but in compound numerals an inherited TB etymon (**cha**) appears, e.g. **cha-bis-e-das** '130' (see below 3.545).

In Khaling, TWENTY and its multiples are expressed by the inherited TB

morpheme -(k)haːel (tu-haːel '20', saːh-khaːel '40', suk-khaːel '60', bhaːel-khaːel '80'). The higher odd numbers, however, are not expressed additively or subtractively in terms of these, but have simply been replaced by the Nepali equivalents: pacaːs '50', sattari '70', nabbe '90'. This is actually quite a rational compromise for the language to have made. It retains the advantage of the concise even multiples of 20, but avoids the cumbersome additive structure of the odd round numbers. Khaling thus enjoys the best of both the decimal and vigesimal worlds.

3.534   Hesitation between decimality and vigesimality: vigesi-decimal vacillation

Several Himalayish languages show particularly interesting vacillation between TEN-based and TWENTY-based systems of higher numerals.

[A]   LEPCHA

A particularly schizophrenic case is LEPCHA, which has two different words for '20', one monomorphemic (k'a) and the other a multiplicative form based on '10' (kə-nyət). Two complete sets of higher round numbers coexist (or at least coexisted in Mainwaring's time), one based on k'a '20' and the other on the combining form kə- (with short vowel and unaspirated initial):

	VIGESIMAL		*DECIMAL*	
TWENTY	k'a; k'a-kat	"20 ($\times$ 1)"	kə-nyət	"10 $\times$ 2"
FORTY	k'a-nyət	"20 $\times$ 2"	kə-fəli	"10 $\times$ 4"
SIXTY	k'a-sam	"20 $\times$ 3"	kə-tərək	"10 $\times$ 6"
EIGHTY	k'a-fəli	"20 $\times$ 4"	kə-kəkŭ[160]	"10 $\times$ 8"
HUNDRED	k'a-fəngo	"20 $\times$ 5"		

Non-vigesimal forms for HUNDRED also exist, though they are borrowed from Tibetan: gya; gyo-kat (kat '1').

The odd round numbers present no problem in the decimal system. In the vigesimal system they are expressed (as is usual in systems of this type) additively in terms of the next lower multiple of 20, by means of the morphemes sə kəti "plus ten":

	VIGESIMAL		*DECIMAL*	
THIRTY	k'a-kat sə kəti	"(20 $\times$ 1) + 10"	kə-sam	"10 $\times$ 3"
FIFTY	k'a-nyət sə kəti	"(20 $\times$ 2) + 10"	kə-fəngo[161]	"10 $\times$ 5"
SEVENTY	k'a-sam sə kəti	"(20 $\times$ 3) + 10"	kə-kəkyək	"10 $\times$ 7"
NINETY	k'a-fəli sə kəti	"(20 $\times$ 4) + 10"	kə-kəkyot	"10 $\times$ 9"

We should note that although the independent numeral kəti '10' has the kə- prefix, and though this has been carried over into the decimal higher round numbers, kə- may originally have had no connection with TEN at all. The kə-

---

160) This form is missing from Mainwaring, and is my guess.
161) This form is lacking in Mainwaring; we supply it (perhaps rashly) as a guess; maybe this form was avoided because of its similarity to k'a-fəngo '100'.

in **kəti** is only part of a "prefix run" that extends from 6 to 10 in Lepcha (below 5.34).

[B]　SHERPA

Sherpa has a classically vigesimal form for TWENTY, **khal-jik** ("20×1"). CSDPN does not provide the word for THIRTY, but it does give **khaljik-tang-curkhu** '39' ("20 + 19"), implying that one counts in an unbroken string from 20 to 40, so that '30' must be **khaljik-tang-ci**(**thambaːq**) ("20 + 10").

I am suggesting the term *twenteens* for the numbers between 20 and 40 in a vigesimal system. (Perhaps a whole new stage in the life-cycle needs to be recognized: people from 21 to 39 could be called *twenteenagers*. See above 3.3.)

From '40' on, however, a curious semantic transvaluation of the morpheme **khal-** has taken place. Instead of meaning TWENTY, it now means TEN:

FORTY	**khal-ji**	(**ji** '4')	SEVENTY	**khal-din** (**din** '7')
FIFTY	**khal-ngaq**	(**ngaq** '5')	EIGHTY	**khal-ge** (**ge** '8')
SIXTY	**khal-Tuk**	(**Tuk** '6')	NINETY	**khal-gu** (**gu** '9')

This obviously secondary "decimalized" system continues through and beyond HUNDRED: **khal-citambaq** '100' ("10×10") [etymologically "20×10"], **khal-cupsum** '130' ("10×13") [etymologically "20×13"]. Compare Jirel **khalq-Thuk-cyutambaq** '130' "(20×6) + 10".

In Sherpa we can appreciate the passage from vigesimal to decimal ways of thinking at a transitional stage.

[C]　SIKKIM BHUTIA/DANJONGKA/DZONGKHA

The Dzongkha system recorded by Chhewang Rinzin [1984] is purely decimal. Sandberg's "Sikkim Bhutia" of a century ago was mostly decimal, but also presents a few unmistakably vigesimal features. The word for TWENTY is **ni-shu** ("2×10") in Rinzin, but **khe-chik** in Sandberg (first syllable < *****m-kul** '20', second syllable < *****tyik** '1' [above 3.14]). The higher round numbers are all formed decimally with **chu-** or **chu-tamba** '10', e.g. **zhib-chu** '40', **ngab-chu** '50', **Tuk-chu** '60' (note the **-b** at morpheme boundary in '40' and '50').[162] But Sandberg also cites a vigesimal variant for '60': **khe-sum** ("score-three"). It seems evident that the vigesimal forms are older in the language, relics of a more thoroughgoing 20-based system.

In her article "Dzongkha number systems", the most detailed and insightful study of the numerals of an individual TB language to have appeared to date, Mazaudon demonstrates that "Dzongkha exhibits a coherent vigesimal system equal in complexity and extension to any vigesimal system described in

---

162) These forms are obviously resyllabifications of compounds where the second element began with prefixal **b-** (cf. WT **bću** '10'). Incidentally, this language has developed special "round number combining forms" for almost all of the primary numerals, e.g. **sum** '3', **so-** 'thir-', as in **so-chi** '31', **so-nyi** '32'; **Tuk** '6', **re-** 'six-', **re-chi** '61'.

any part of the world" [1985:150].   Not only does the language retain the general monomorphemic TB root for '20' (**khe** < *__m-kul__), but it also has lexemes for the next three powers of twenty:

20^1	**khe**	20
20^2	**ɲiɕu**[163]	400
20^3	**kheche**[164]	8,000
20^4	**jãche**	160,000

Even so, the language has a normal decimal system of teen-formation (TEN + UNIT).   Above twenty there now coexist two distinct systems of reckoning, one vigesimal and one decimal, with stylistic differentiation: the decimal system is characteristic of formal speech.

	DECIMAL	VIGESIMAL
'20'	**ɲi-ɕu** ("2×10")	**khe-ci** ("20×1")
'21'	**tsa-ci**[165] ("20 + 1")	**khe-ci** (**da**) **ci** ("[20×1] + 1")
'22'	**tsa-ɲi**	**khe-ci** (**da**) **ɲi**
'30'	**sum-cu** ("3×10")	**khe pɹhe da ɲi** ("20×[− 1/2 + 2]")
'31'	**so-ci** ("thir- + 1")[166]	**khe-ci da cu-ci** ("[20×1] + 11")
'35'	**so-ŋa** ("thir- + 5")	**khe ko da ɲi** ("20×[− 1/4 + 2]")
'40'	**ʑi-p-cu** ("4×10")	**khe-ɲi** ("20×2")
'50'	**ŋa-p-cu**	**khe pɹhe da sum** ("20×[− 1/2 + 3]")
'55'	**ŋa-ŋa**	**khe ko da sum** ("20×[− 1/4 + 3]")
'60'	**ɖhuk-cu**	**khe-sum** ("20×3")
'70'	**dyn-cu**	**khe pɹhe da ʑi** ("20×[− 1/2 + 4]")
'80'	**ge-p-cu**	**khe-ʑi** ("20×4")
'90'	**gu-p-cu**	**khe pɹhe da ŋa** ("20×[− 1/2 + 5]")
'100'	**cik-ɹa** ("1 × 100")[167]	**khe-ŋa** ("20×5")
'400'	**ʑ ip-ɹa** ("4×100")	**ɲiɕu-ci** ("400×1")
'500'	**ŋap-ɹa** ("5×100")	**ɲiɕu-ci da khe-ŋa** (["400×1"] + [20×5]")
'600'	**ɖhuk-ɹa** ("6×100")	**ɲiɕu pɹhe da ɲi** ("400×1 1/2")

In the vigesimal system, the even round numbers are expressed straight-

---

163) This form, etymologically "2 × 10", originally meant '20' in the decimal system, but has been transvalued to mean "20 squared" in the vigesimal system! See the following chart, and section 4.02 below.

164) The second syllable is identified by Mazaudon (p.137) with WT **che-ba** 'large', so that the compound means "a large twenty", much as French *une grosse* (> Eng. *gross*) '12 dozen; 144' derives from *une grosse dizaine* "a big dozen".   The etymology of the first syllable of **jãche** remains obscure.   Michailovsky (p.c. 1995) suggests a connection with Tibetan *yaŋ* 'again', i.e. 'even bigger'.

165) The morpheme **tsa** is a fascinating example of a radical but entirely natural semantic slippage.   As Mazaudon shows (p.129), it derives from the WT conjunctive particle **rtsa** used to connect the tens to the units, e.g. **nyi-śu-rtsa-gćig** '21' ("2×10 plus 1").   When the first element was omitted (as often in Tibetan itself), the connective took on the meaning '20'! See the discussion of numerical transvaluation, below 4.02.   The morpheme **da** (< WT **daŋ**) then took over connective function in the Dzongkha vigesimal system (e.g. **khe-ci da ci**).

forwardly as multiples of **khe** '20', but the odd tens and fives are formed by what Mazaudon (following Menninger 1958/1969) calls "back-counting". According to a Tibetan pattern whereby 'one and a half' is expressed as **phyed-daŋ gnyis** (**phyed** 'half', **daŋ** 'with', **gnyis** '2'), i.e. "which with an additional one-half, would be two"), the odd round numbers are formed subtractively in Dzongkha: e.g. '30' **khe pɟhe da ɲi**, i.e. "twenty times one-half-less-than-two", or "twenty times one-and-a-half").[168] Similarly, Dzongkha expresses the odd fives by backcounting in quarters (**ko** 'one fourth; a quarter'), as in '55' **khe ko da sum**, literally "twenty times one-quarter-less-than-three", or "twenty times two-and-three-quarters".

Mazaudon convincingly argues for the ancient status of vigesimal numeral systems in TB, and attributes the relatively good preservation of vigesimality in Dzongkha to the political independence of Bhutan, and Dzongkha's status as a national language, so that it could "resist the spread of the all-powerful decimal system which had the support of both India and China" (p.150). The decimal aspects of the Dzongkha numerals were "borrowed from Tibetan for elegant speech" (p.154).

3.535 Duodecimality in Chepang

One of the strangest numeral systems in TB is to be found in Chepang. A study of the data presented in CSDPN shows it to be *duodecimal* in structure, conceived in terms of TWELVES, not TENS or TWENTIES.

Nowadays only the first 5 Chepang numerals are inherited TB etyma: **yaːt-joʔ, nis-joʔ, sum-joʔ, play-joʔ, ponga-joʔ**.[169] SIX through TEN are expressed by Nepali numerals: **cha-gotaː, saːt-gotaː, ʔaːt-gotaː, naw-gotaː, das-gotaː**. The crucial form for ELEVEN is missing in CSDPN, but presumably it too is simply the unanalyzable Nepali word.

The Chepang word for TWELVE is truly unique: **yaːt-haːle**. The first syllable means ONE (above 3.11), and the second element is clearly the inherited TB word for TWENTY (< **kal ✕ *m-kul*; cf. Khaling (**k**)**haːel**). Though it is theoretically possible that Chepang alone of all the TB languages has preserved an "original" meaning TWELVE, while the rest of the family has changed its meaning to TWENTY, it seems much more likely that it is Chepang which has transvalued the numeral from TWENTY to TWELVE.[170]

---

166) See n. 161 for the first syllable. For an etymological explanation of the special combining forms of the names of the units used for the tens, see Mazaudon 1985:153.

167) One can also say **ɟa-thampa** ("100 full"), with the same "full" morpheme as found, e.g. in Sherpa and Jirel (above 3.22, 3.235, 3.533[C], 3.534[B]).

168) This is not so different from the German and Russian way of expressing the half-hour when telling time, e.g. Germ. *halb vier*, Russ. *polovino cetvërtogo* '3:30', i.e. "half of four".

169) Hodgson [1880:166-167] gives them all the way to 10: **kruk-zho** '6', **chana-zho, prap-zho, takhu-zho, gyib-zho**.

170) Cf. the case of Sherpa, above 3.533, where this same etymon has been transvalued from ↗

The other forms available in Hale (ed.) 1973 support the duodecimal analysis:

ya:t-ha:le ?a:t-gota:	'20' ("[1 × 12] + 8")
ya:t-ha:le das-gota:	'22' ("[1 × 12] + 10")
nis-ha:le	'24' ("2 × 12")
nis-ha:le ponga-jo?	'29' ("[2 × 12] + 5")
nis-ha:le sa:t-gota:	'31' ("[2 × 12] + 7")
sum-ha:le play-jo?	'40' ("[3 × 12] + 4")
play-ha:le nis-jo?	'50' ("[4 × 12] + 2")
ponga-ha:le	'60' ("5 × 12")

With the simple form for SIXTY the decimal and duodecimal systems are reconciled — both 10 and 12 are factors of 60. It is this beautiful fact that lies behind such systems as the Chinese 60-year calendrical cycle of the "10 heavenly stems" and "12 earthly branches." Unfortunately no Chepang forms higher than SIXTY appear in Hale (ed.) 1973 (CSDPN), and probably no Chepang would ever use anything but Nepali numerals in that rarefied range. (CSDPN notes that even the forms listed above are now much rarer than their Nepali equivalents.)[171]

Since the first version of this paper was written (1984), R.C. Caughley, the leading authority on Chepang, has published a short article specifically on the subject of Chepang duodecimality [CAUGHLEY 1989]. Here he adds another form he recorded as **ya:t-ha:le sum-jo?** '15' ("[1 × 12] + 3"), and offers an intriguing possible explanation for the use of twelve as a numeral base: "When counting the tip of the thumb is placed against each interstice in turn, starting from the base of the little finger and ending at the tip of the index finger. Since there are four fingers, each with three interstices, this means a total of twelve for each hand, and makes twelve a natural basis for counting" [1989:197].

3.54   HUNDRED and THOUSAND

3.541   Decimal multiplicative expressions for HUNDRED ("10 × 10")
GALLONG (AMD) **cam-ri** '100' (**cam-** "-TY", as in **cam-nyi** '20' ("-TY × 2"), **cam-um** '30', **cam-ke** '60'; **i-ri** '10') [J.T.SUN 1993:276-277].
SHERPA (Him.) **khal-citambaq** '100' (**citambaq** '10'; **khal** is the inherited etymon for TWENTY (< *m-kul ⋊ *kal), but has become transvalued to TEN in Sherpa, as in **khal-ngaq** '50', **khal-Tuk** '60' (above 3.534[B], below 4.02).
LAKHER (KCN) **sy-hraw** '100' (**sy-** "-TY", as in **sy-pali** '40'; **pa-hraw** '10').

↘ TWENTY to TEN.

171) The Chepang forms for '40', '50', and '60' are inadvertently transposed one column to the left in CSDPN (p. 204), which makes it a maddening task to figure out what is going on! This error has also been noticed by Mazaudon [1985:155].

Lakher has two other ways of expressing '100', either as an independent unanalyzable numeral (**za**), or as "100×1", where the morpheme HUNDRED functions as a classifier (**za-kha** '100', **kha** '1').

3.542　Vigesimal multiplicative expressions for HUNDRED ("20×5" or "5×20")

	HUNDRED	TWENTY	FIVE
Kuki-Naga			
Chang	**sau-ngau**	**sau**	**ngau**
Wancho[172]	**pu-ga**	**pu-**	**ga**
Yacham-Tengsa	**mesung-phung**	**mesung-**	**phung**
Abor-Miri-Dafla			
Monpa Motuo	**k'ai-nga**	**k'ai**	**nga**
Himalayish			
Tamang	**khaːl-ngaːh**	**khaːl**	**ngaːh**
Sharchop	**khye-nga**	**khye**	**nga**
Lepcha	**k'a-fəngo**	**k'a**	**fəngo**
Kaike	**ngaː-thal**	**thal**	**ngaː**

3.543　Where HUNDRED has a special relationship with FIFTY

In MIJU MISHMI [Das Gupta 1977a], '100' is **waie-mo**[173] (**-mo** < **kumo** '1'), while '50' is **wa-ping-mo**, glossed literally as "half of hundred". The numeral system described by Das Gupta is one of the strangest in TB, but certain key forms are lacking ('40', '60') and the morphophonemics of the Miju round number system are still not clear.

3.544　Where HUNDRED functions as a classifier

In languages with this formation, HUNDRED is expressed as "100×1" or "1×100".

(a)　"1×100"

Lolo-Burmese

LAHU **tê ha** '100' (**tê** '1'); **ha** < PLB *hra^1 or *$ʔra^1$

Qiangic

ERSU **$tə^{55}$ $zə^{55}$**

(b)　"100×1"

Himalayish

DZONGKHA　　**ja-ci** (**ci** '1')

Barish

---

172) Another dialect of Wancho [Das Gupta 1979] has hesitation between a vigesimal and a non-vigesimal expression: **hu-ga** (**hu** '20', corresponding to Marrison's **pu-ga**) but also **ho-ta** (**-ta**, **tu-ta** '1'), where **ho** looks like a reflex of the general monomorphemic root *$b-r-gya$ (below).

173) Sun Hongkai et al. [1980] transcribe Deng Geman (Kaman) '100' as **$wə^{55}$ je^{53} mu^{53}**.

GARO   **ritcha-sa** (sa '1'), DIMASA **raja-si**

Kuki-Naga

TANGSA (Yogli) **sha-kha** (kha '1'); KUKI **ja-khat**; KIMSING **shi-shi**
(shi '100', shi '1' — presumably under different tones);[174] Lakher **za-kha**
(alongside two other expressions: see 3.541); Pochury **mza-ke**; GARO **rit-cha-sa**; DIMASA **raja-si**

Of special interest is TANGSA (Moshang): **rok-shi** '10', **rok-ni** '20', etc.
(**rok** = "-TY"), but **rok-sha-shi** '100'. The middle syllable in **rok-sha-shi** is
the root HUNDRED, and the third syllable is ONE, so that here **rok-** does not
have the meaning 'TEN', or indeed any meaning at all — it appears pleonastically, by analogy with the round numbers from 10 to 90.

3.545   *Where HUNDRED is a monomorphemic reflex of PTB *b-r-gya ✕ *b-g-rya[175])*

It seems clear that this etymon has been remodelled by analogy with
EIGHT (below 4.23). For EIGHT, STC recognizes metathesis (*b-r-gyat ✕
*b-g-ryat). We must now invoke metathesis in HUNDRED as well — perhaps
an inevitable occurrence in such a complicated consonant sequence involving a
liquid.

*b-r-gya   *with double prefix*

WRITTEN TIBETAN **brgya**; rGYARONG **perzhe**; AKA (Hruso)
phogwa, purrwa

*b-rya   *with labial prefix*

MIKIR **paro**; NRUANGHMEI **phai** (with preemption of the initial)
GURUNG **praːhq**; THAKALI **prah**

*g-rya   *with velar prefix*

ANGAMI, CHOKRI **kra, krie**; KEZHAMA, MAO **kri**. KHOIRAO **ki**,
SEMA **a-keh** and LIANGMAI **kai** might owe their velar onset to preemption of the liquid initial by the prefix — or they might simply descend
from an unprefixed form where the velar was the root-initial (*gya).
Under this allofam also belong laryngeal-initialled forms like Zeliang-Zeme-Mzieme and Maram **hai**.

*m-rya   *with nasal prefix*

LOTHA **nzoa**; MARING **macha**; MELURI and NTENYI **meza**;
POCHURY **mza-ke**

*r-gya   *with liquid prefix*

JINGPHO **lətsa**[176]; PUIRON **raja**; KOM REM **raza**; PHOM **gho** (pre-

---

174) The apparent homophony of these two syllables reminds one of the Lahu number **hí hí** '8000',
where even the tones are identical. This is pure accident, however: the first syllable is from PLB
*r̀it 'eight', but the second is a loan from Shan **hiŋ** 'thousand'. See Matisoff 1988a:1070.

175) STC #164 *r-gya (pp. 45, 54, 57, 89, 94-95, 109, 131, 137, 151, 161-162).

176) Alongside Jg. **mətsát** 'eight'.

sumably "gh" is a voiced velar fricative < *r); KOKBOROK **racha**; GARO **ritcha-sa**; DIMASA **raja-si**

*-**gya** or *-**rya** (*no unambiguous evidence for any particular prefix*)

SIKKIM BHUTIA **gya**; MAGARI **cha** (above 3.533[E]); KONYAK **kho**; KANAWARI **ra**

YOGLI **ša**; MOSHANG **rok-ša-shi**, NOCTE **cha** (all < PNN *-**khya** [FRENCH 1983:506])

TANGKHUL **sha**; MEITHEI **cha**; RENGMA **tsi**; SANGTAM **thsi**; YIM-CHUNGRU **chhi**

LUSHAI, LAKHER, HMAR, GANGTE, PAITE, TIDDIM **za**; VAIPHEI **ja**

BORO **jou** ; DULUNG **çya**

PUMI **shə^{55}**; QIANG **tshi^{55}**

KAREN (Pa-O) **rja**, (Pwo) **ja**, (Palaychi) **sa**

3.546　A special Abor-Miri-Dafla root for HUNDRED and its relationship to TEN

*m-li(ŋ) (? < *m-l-ŋya)

Abor-Miri	**li-ko ~ ling-ko**
Minyong	**ling-ko**
Dafla	**leng-go**
Nishi	**lunkh**
Apatani	**lange, laŋ-e**
Monpa Cuona	**c'ɛʔ^{53} le^{21} nge^{53}**
Lhopa	**luɯng**
Darang Deng	**məluɯm^{55}**
Chulikata	**malu:**
Mishmi	**malo**[177]
Gallong	**hamyi** (< *s-mliŋ ?)

It will be observed that these forms bear more than a slight resemblance to a set *riŋ ✕ *yiŋ 'TEN' discussed above (3.233a,b): e.g. Abor-Miri **eying-ko** '10', etc. However, my present view is that these two roots are distinct (cf. pairs of reflexes like Lhopa **uɪjuɯng** '10', **luɯng** '100'),[178] though they may have "contaminated" each other.[179] It is possible that these forms for HUNDRED are ultimately to be derived from *m-lŋya (ult. < *b-rgya) via apocope of the root vowel.

---

177) Cf. perhaps Mishmi [DUBEY] **muou** '10' (above 3.237, 3.42).

178) J.T. Sun [1993:121] sets up a Proto-Tani root *luɯŋ 'hundred', distinct from PTani *rjuɯŋ 'ten' (144), citing forms like Bengni and Bokar **luɯŋ**, Bangru **ləŋ^{53}**, Dhammai **buɪ-loŋ**, and Hruso **phu-yu**.

179) A case of confusion of '10' and '100' through borrowing is pointed out for Kanauri by Joshi [1909:108], where Tibetan **nyi-gyá** '200' has been borrowed as Kanauri **ni-já** '20'.

3.547   THOUSAND

(a)   *s-toŋ

STC #32 (pp. 21, 94) sets up the root *s-toŋ on the basis of forms from two languages, WT **stoŋ** and WB **thoŋ**.   To these we may add:

Khoirao	**tang**
Maram	**tang**
Sema	**khe-thon-he** (**khe** '1')
Mao	**thu**
Newari	**dwa:**
Sikkim Bhutia	**tong-ta**
Naxi	**dtv**[1]
Pumi	**stĩ**[55]
Qiang	**xto**[55]
Dulung	**tu**[55]
Karen	(Pwo) **thon**, (Sgaw) **kəʔtho**

(b)   *ś-riŋ ⋊ *s-raŋ

More problematic are forms meaning THOUSAND in a number of languages with sibilant/affricate, dental, or liquid initials plus the rhyme *-iŋ*, as well as similar forms in other languages with the same kinds of initials but the rhyme *-aŋ*.   It is my feeling that all the following words are related to one another somehow, and that we should leave open the possibility that they are ultimately connected to either of the phonologically similar etyma meaning TEN or HUNDRED already discussed (above 3.233, 3.546):[180]

(1)   with the rhyme *-iŋ*

GANGTE **sing**; TANGSA (Yogli) **hing**; KIMSING **hing-shi** (**shi** '1'); MEITHEI **lising**; MARING **lising**; PUIRON **lising**; KOM REM **lising**; DIMASA **rijing-si** (**se** '1'); TANGKHUL **thing-kha** (**kha** '1'); JINGPHO **ching**; KAREN (Pa-O) **təʔ-rèng** (**təʔ** '1'), (Palaychi) **hrɛng**

(2)   with the rhyme *-aŋ*

TARAON **reja:ng**; AO (Chungli) **meirijang**; MZIEME **tsang**; ZELIANG and ZEME **chang**; LOTHA **tsanga**; LIANGMAI **shang**; KUKI **sang**; PAITE **sang** (alongside **sing** '10,000'); AO (Mongsen) **miyarsang**; VAIPHEI **sang-khat** (**khat** '1'); LUSHAI **sang**; LAKHER **sa** (low tone)

(c)   A few languages have forms with palatal nasal or semivowel:

ANGAMI **n**(**y**)**ie**; MELURI **anye**; KHEJA **nie kele**; RENGMA **ye**

(d)   There remain a few miscellaneous forms whose affiliations are much in doubt:

• ERSU **hpu**[55] '1000' looks like the root for TWENTY found in a few languages (MEITHEI **phu**, WANCHO **pu-**, PHOM **pü**).   See above 3.511.

---

180) Note that the Germanic word for THOUSAND is historically a compound meaning "swollen hundred" (PGmc *θus-hundi, Old Norse θushundrad).

- YIMCHUNGRU **amükhepin** is totally mysterious, though the element **mükhe** looks much like a reflex of the root *m-kul '20' (but the Yimchungru word for '20' is given as **muku** in GEM 279).
- TIDDIM **tul** and MIKIR **suri** look as if they are related to each other, but not obviously to anything else.

3.5471 Multiplicative forms for THOUSAND

A number of languages have multiplicative formations for THOUSAND, with a wide variety of possibilities (since there are so many factors of such a big number):

(a) "100 × 10"

NOCTE	**cha-ichi**	(**cha** '100', **ichi** '10')
NTENYI	**meza-ta'a**	(**meza** '100', **ta'a** '10')
DAFLA	**leng-rengcheng**	(**leng** '100', **rengcheng** '10')
APATANI	**lã-lyã**	(**lã** '100', **lya** '10')
ABOR-MIRI	**li-ying-ko**	(**li** ~ **ling** '100', **eying** '10')
ADI[181]	**ling-ko-iying**	(**ling-ko** '100', **iying** '10')

(b) "10 × 10 × 10"

GALLONG	**cam-ri-iri**	(**cam-** '10; -TY', **iri** '10')

The morpheme **cam-** is used only in round numbers (e.g. **cam-nyi** '20'); **iri** is the independent numeral.

(c) "20 × 50"

TAMANG	**kha:l paca:s**	(**kha:l** '20', **paca:s** '50' [< Nepali])

(d) "20 × 10 × 5"

CHANG	**sau-an-ngauni**	(**sau** '20', **an** '10', **ngau** '5')

(e) WANCHO has the paradoxical form **puban: pu** means '20' (cf. **puga** '100', **ga** '5'), and **ban** means '10', so we would expect this word to mean '200', not '1000'![182]

3.5472 Borrowings of Indo-Aryan reflexes of PIE *gheslo- 'thousand' in TB languages

Finally, we should mention several forms for THOUSAND in Himalayish and Kamarupan languages that are borrowings from Indo-Aryan, ultimately from the PIE root *gheslo- (cf. Greek *xīlo-* < *xeilo <*xeslo, Sanskrit s*ahasra* '1000' (sa- 'one'), Armenian *hazar*, Persian *hazar*; Avestan *hazarva*, Nepali [SCHMIDT 1993] *hajaar*):

ADI **ejar** ~ **hajar** ("a borrowed word of Indo-Aryan origin") [MEGU 1985:77]

GURUNG **hãqjãr**

---

181) This Adi form is from Megu 1985. Note the different order in which the morpheme **ko** 'one' appears in Abor-Miri and Adi.

182) Could **puban** be a typo in Marrison [1967:267] for **puga-ban** ("100 × 10")?

KANAWARI **hanzár** (apparently with rhinoglottophilia in the first syllable; see Matisoff 1975.)
KONYAK, PHOM **haja**
SANGTAM **hajar**

4. THE PRIMARY NUMERALS: TWO TO NINE

Tibeto-Burman languages mostly reflect one and only one etymon per numeral, especially with the *lower numerals* 2-5; the *higher numerals* 6-9 show more variation, with occasional additive, subtractive, and multiplicative complications, and a number of isolates and roots of limited distribution (not mentioned in STC).

For numerals above NINE, a language will occasionally use different words according to the thing being counted, e.g. Apatani **lya** '10' (of nonhumans) ~ **alyã** '10' (of humans) [above 3.233]; Dzongkha **tsa** '20' (for counting objects from 21-29, e.g. **tsa-ɲi** '22') vs. **ɲer** '20' (for dates, e.g. **ɲer-ɲi** 'the 22nd of the month'; Mazaudon 1985:129).

4.01 Mutual influence of numerals (phonological): convergence and contamination

Since the numerals are such a uniquely structured semantic field, where the members typically occur one after the other in a fixed, rapid sequence (counting), it is no wonder that they are subject to all sorts of assimilatory phenomena. Examples may be found affecting all parts of the TB syllable:

(*a*) *Influence on prefix*

The most striking of these effects involves the prefixes that are so characteristic of numerals in the non-Sinospheric branches of TB. In the case of consecutive numerals, we speak of "prefix runs", treated below in detail (5.2 et seq.), e.g. Jingpho **məsūm** '3' (< **g-sum*) ← **məlī** '4' (< **b-ləy*).[183] In rare cases the prefix of a non-consecutive numeral may be a "contaminating agent", the best example being WT **brgya** '100', apparently modelled after **brgyad** '8' (above 1.12).

(*b*) *Influence on root-initial consonant*

Consecutive numerals may influence each others' initial consonants, with several famous examples in Indo-European.[184] Cf. e.g. PHOM **shüt** '8' → **shü** '9'; SERDUKPEN **khu** '5' ← **khit** '6'; KHALING **saːhpu** '2' ← **suhpu** '3';

---

183) *In this discussion we use the symbols → and ← to indicate the direction of influence.*

184) E.g. Germanic FIVE influenced FOUR (we would expect English **whour*); Russian **desjat'** '10' influenced **devjat'** '9' (instead of the expected **nevjat'*). See below 5.2. At first glance it looks as if Latin **quattuor** '4' influenced the initial of **quinque** '5' (instead of the expected **pinque*); but the labiovelar in five was a regular assimilatory development in etyma of the form **p...kw* (other examples are 'oak' (PIE **perkwo-* > Lat. **quercus**, not **percus*) and 'cook' (PIE **pekw-* > Latin **coquo**, not **poquo*). See Jasanoff 1994.

KHOIRAO **kati** '2' ($<$ *kani) ← **kathum** '3'.

Milang **rangal** '7' has perhaps been influenced by **pangu** '5' ($<$ *b-ŋa), since the velar nasal is present "by right" in FIVE, but not in SEVEN. The liquid prefix in **rangal** also has ancient status with FIVE (STC sets up the proto-allofam *l-ŋa; cf. WT lŋa, Kom Rem **ranga** '5'), but not in SEVEN, which is reconstructed with *s-. (The final lateral -l in **rangal** remains a complete mystery — as does the aberrant Milang language in general!)

(c) *Influence on rhyme*

Consecutive numerals may influence each others' rhymes, e.g. THULUNG RAI **yet** '7' [$<$ *-is] ↔ **let** '8' [$<$ *-yat] (here the influence seems mutual); Sunwar **tsəni** '7' → **tsəsi** '8' (Gvozdanović 1985:143; here both the prefix and rhyme of '8' have been affected); SERDUKPEN **khit** '6' (originally with final velar, *d-k-ruk) ← **sit** '7'.[185)]

(d) *Generalization of final consonant to suffixal status*

We have mentioned how some Karen dialects, especially Pa-O (Taungthu) have generalized the final dentals that occur "by right" in SEVEN and EIGHT ($<$ PTB *-s and *-t respectively) to other numerals where they do not belong etymologically (**līt** '4', **ngāt** '5', **kūt** '9'). (See above 1.21, and Benedict 1979:18-20.)

(e) *Influence on tone*

Consecutive numerals may be affected by junctural phenomena that make themselves felt in counting. Lahu **šê̂ʔ** '3', with non-etymological high-stopped tone (written with -ˆʔ), doubtless developed the glottal stop as a demarcational feature to set it off from the next higher numeral **ɔ̂** '4', that begins with a vowel.[186)]

Tones of successive numerals have undergone widespread convergence in Loloish, where ONE, TWO, THREE, FOUR, FIVE all show reflexes of PLB Tone *2, along with SEVEN and NINE. Only SIX and EIGHT, deriving from *stopped syllables, escaped this generalizing tendency.

(f) *Additive or subtractive copying of an adjacent numeral*

In the most extreme cases, an entire numeral is expressed in terms of the next higher or lower one, e.g. Mikir **throk** '6' → **throk-si** '7' ("6 + 1") [4.229]; Meithei **nipan** '8' ("2 from 10"), **mapan** '9' ("1 from 10") [above 3.232; below 4.23, 4.24]. See below 4.20.

4.02 Transvaluation of numerals (semantic)

Throughout this paper we are concerned with pointing out indisputable or possible cases of "numerical transvaluation", i.e. a shift in the referent of a numeral from its etymological meaning. In this section the most interesting

---

185) Since Serdukpen '6' also influenced the *initial* of '5' (above), we can establish a three-link "push-chain": 7 → 6 → 5.

186) See my note 413 in STC (p. 152), and below 4.122.

examples are listed together for convenience's sake.[187]

- Interchanges and confusions between 'ONE' and 'TEN' (above 3.4), and between 'TEN' and 'HUNDRED' (above 3.547b).
- YACHAM-TENGSA **tamong** '20' looks like the same etymon that means 'FIFTY' in most other languages where it occurs (above 3.522).
- The MIKIR reinterpretation of **throk** 'SIX' to mean 'TEN' in the higher round numbers (above 3.523).
- The change in the value of *__m-kul__ 'TWENTY' to 'TEN' in SHERPA (above 3.534[B]), and to 'TWELVE' in CHEPANG (i.e. from "2 × 10" to "2 + 10"; above 3.535).
- The DZONGKHA numeral **ɲiçu**, which sometimes means the etymologically correct 'TWENTY', but is often used to mean '20 × 20' or 'FOUR HUN-DRED' (above 3.534[C]).[188]
- In LEPCHA an etymological flipflop between 'EIGHT' and 'NINE' seems to have occurred: **kǎkǔ** '8' (but cf. PTB *__d-kəw__ 'nine', below 4.24), **kǎkyót** '9' (but cf. PTB *-__gyat__ '8', below 4.23).
- The mysterious word **zon** means 'TWO' in BUMTHANG but 'SEVEN' in SHARCHOP. The interchange between TWO (PTB *__g-nis__) and SEVEN (PTB *__s-nis__) is in fact the most obvious link between separate elements in the TB system of primary numerals, and certainly seems to bespeak a very early QUINARY or 5-based principle in the system (below 4.11, 4.14, 4.20, 4.22).
- The moribund numeral systems of the Kiranti languages of E. Nepal are rife with examples of "downward shifts" to a lower numeral, or occasionally "upward shifts" to a higher one [GVOZDANOVIĆ 1985]:
  - •• In LIMBU the numeral **iboong** '9' has evidently been transvalued from its original meaning '10' (compare Kulung **ik-poŋ**, Yakkhaba **ip-poŋ** '10') [above 3.236; GVOZDANOVIĆ 1985:162].
  - •• The Moli dialect of BAHING has etymologically correct forms for '6' (**rukhu**) and '7' (**cūni**), but in the Bigutar dialect the cognate forms **ruka** and **cani** mean '5' and '6', respectively [GVOZDANOVIĆ 1985:135].
  - •• In KULUNG the word **tupci**, etymologically '6', has come to mean '5' in the Pawoi dialect, while the word **retci**, originally '8', now means '7' in the Bung dialect [GVOZDANOVIĆ 1985: 135].
  - •• The Saptesor dialect of DUMI has **sukpo** '3' and **bhalukpo** '4', but the cognate forms in the Kubhinde dialect, **sɔkpu** and **bhlɔkpu**, mean '2' and '3', respectively. On the other hand, Saptesor **dumpo** '6' looks like it originally meant '7' (a downward shift; cf. WT **bdun** '7').

---

187) In Appendix I we shall offer a semantic diagram or "flowchart" that schematizes all these shifts in meaning.

188) As we have seen (above 3.534 [C]), Dzongkha also furnishes a beautiful example of a "trans-field transvaluation", i.e. a case where a word from a non-numerical semantic field has acquired a numerical meaning: **rtsa** 'and' > Dz. **tsa** '20'.

•• Saptesor **rekpo** '9' looks like it comes from *__*b-rgyat__ '8' (an upward shift);
in Kubhinde the cognate form **rɔkpu** has come to mean '4' (half of eight),
just as in the Raniṭar dialect of BANTAWA, where the cognate **reṭkapok**
now also means '4' instead of EIGHT [GVOZDANOVIĆ 1985:135-136].

•• In SUNWAR, the word **gow** means '10', but certainly seems to descend
from *__*d-kəw__ '9' [GVOZDANOVIĆ 1985:143]. Similarly, Sunwar **yaan** '9'
looks like it derives from *__*g-ryan__ '8' (below 4.236).

4.1 The Lower Numerals: TWO to FIVE

4.11 Profile of number TWO

<div align="center">TWO   *g-ni-s/k</div>

Like THREE, TWO is one of the most phonologically and lexemically
stable numerals in TB. Again like THREE, the only consonantal prefix that
can be reconstructed for TWO at the PTB level is the velar *__*g-__/*__*k-__. The
relatively rare final consonant *__*-s__ is reconstructed for this root on the testimony
of WRITTEN TIBETAN (WT) **gnyis**, rGYARONG **kenes**, and KANAURI **nis**
[STC p. 4]. Forms in other languages reflect *__*-ik__ instead of *__*-is__, notably
WRITTEN BURMESE (WB) **hnac**. Still others seem to derive from open
syllables, e.g. LAHU **nî** < PLB Tone *2. (Cf. similar open-syllable derived
forms in Loloish words for SEVEN, below 4.224.) The STC, without much
discussion (see n. 486, p. 185), considers the final *__*-s__ to be suffixal at the PST
level, and the etymon is so reconstructed in the Indexes: "*__*g-nis__ = *__*g-ni-s__".[189]
For Proto-Lolo-Burmese (PLB) I have reconstructed *__*(ʔ)ni-t__ [TSR #160], but
there is no hard evidence to enable us to distinguish between *__*-t__ and *__*-k__ here,
and perhaps *__*(ʔ)ni-ʔ__ would be a better reflection of this indeterminacy.

Scattered around here and there are forms with secondary final -t:

LEPCHA	**nyi, nyət**[190]
AO (Mongsen)	**anet**
YACHAM-TENGSA	**anat**

All these forms, however, go back to *__*-s__, since *__*-s__ > -t seems to be the normal
development in these languages, as demonstrated by the fate of *__*s-rus__ 'bone' >
LEPCHA **əhrət**; AO **teret, terat**; Yacham-Tengsa **telet** (STC #6; GEM 34).

The unmistakable similarity even at the PST/PTB levels between TWO
and the word for SEVEN (PTB *__*s-nis__) makes it obvious that some semantic
connection was involved. Since 7 − 2 = 5, it is reasonable to guess that there
might have been a QUINARY basis for the PST numeral system (STC p.
16).[191] There is, however, no evidence to link SIX with ONE, EIGHT with

---

189) See STC #4, and pp. 16, 75, 94, 130, 131, 147, 162, 168, 169, 185, 186.
190) LEPCHA here exhibits the variational pattern **-i-** ~ **-ya-**, that is so well documented for TB as
a whole (see STC n. 251, p. 84; VSTB pp. 40-43).
191) KHMER has such a system to the present day, where SIX is expressed as "5 + 1", SEVEN as
"5 + 2", EIGHT as "5 + 3", and NINE as "5 + 4". See below 4.15.

THREE, or NINE with FOUR.

It is not surprising that there should be hints of morphological accretions to this root, since words for TWO (perhaps the most important of all the numerals, with the possible exception of ONE) tend to combine with spatial and temporal morphemes in idiosyncratic ways. (Cf. English *two, twelve* (< **twa-li:f** "two left [beyond ten]"), *twenty, twain, twice, between, betwixt, twin,* etc.)

Cases of lexical replacement of the basic PTB etymon for TWO are excessively rare, the most striking example being JINGPHO **ləkhôŋ**, which has never been successfully related to anything else. (Curiously, JINGPHO also has an isolated form **ləŋâi** 'ONE', with the same prefix and under the same [rare and secondary] falling tone.[192])

4.111   *Forms with velar prefix*

ANGAMI (Khonoma) **kena**, (Kohima) **kenie**; CHAKHESANG and CHOKRI **küna**; KEZHAMA **kenhi**; KHOIRAO **kati** (with denasalization of root-initial, maybe under the influence of **kathum** '3'); MAO **kahei**; MARING **khani**; MELURI **keni**; MZIEME **kena**; NRUANGHMEI **kanei**, **künei**; NTENYI **kenyi**; POCHURY **küni**; PUIRON **kani**; RENGMA **khohüng**; SEMA **kini**; TANGKHUL **khani**; ZELIANG and ZEME **kena** DIMASA **gini**; GARO **gini, gni** (Momin [n.d.] also gives a reprefixed form **gegni**, alongside **gesa** '1', **gedok** '6', **gesni** '7')

CHULIKATA **ka:ni**; DENG DARANG **kə²¹n⁵⁵** (with syllabic nasal; this form seems well on the way to "preemption via apocope of the root vowel"; see e.g. FIVE, below 4.1411); DENG GEMAN **kɯ²¹jin⁵³** ("j" is the palatal semivowel); DIGARU MISHMI **ka:-ying**; IDU **ka-nyi**; MIJU **kinin**; MISHMI (Dubey) **kani**; TARAON **ka:ing**; AKA (= HRUSO) **kshi**

These Abor-Miri-Dafla forms require some comment. Several languages show a final nasal (DENG GEMAN, DIGARU MISHMI, MIJU, TARAON), which appears suffixal. (This is especially clear in Miju, which has two nasals in the syllable.) However, forms like DENG GEMAN **-jin**, DIGARU **-ying**, and TARAON **-ing** are more equivocal. Two hypotheses seem equally likely in the present state of our knowledge: either these final nasals are also suffixal, and the root-initial **n-** has become weakened to **y-** under the palatalizing influence of the following **-i**; or else the root-initial **n-** and the palatal vowel switched places by *metathesis*, so that the nasal is not an original suffix, but rather the original root-initial. This latter alternative is not so far-fetched, since for DAFLA, Das Gupta [1969] records two variants for TWO, **anyi** and *ain*. The strange sibilant in AKA **kshi** could represent a fricativization of the palatal semivowel: *n-* >

---

192) For an explanation of this form as an allofam of the Jg. first-person pronoun ŋai, see Matisoff 1995a "Watch out for number one", LTBA 17.1 (to appear).

ny > y > sh.[193]

The distinctive initials in KAREN (Palaychi) **chi** and (Sgaw) **khi** [JONES 1961] might well be preemptive survivals of the velar prefix. Other Karen dialects (Pa-O, Pwo **ni**) simply reflect the prefixless root.

rGYARONG forms for TWO (e.g. **kenes**) have a velar prefix, but so do all its numerals from 2 to 10 (except for 8).

4.112　Forms which show no trace of a consonantal prefix

DULUNG a²¹ ni⁵⁵; RAWANG əni

APATANI nǐ; ABOR-MIRI, GALLONG, LHOPA, MINYONG, TAGIN **a-nyi**; PADAM **a-ni**; NISHI **anni**; DAFLA [DAS GUPTA] **anyi** ～ **ain**; MILANG **ne**; MONPA (Cuona) **nAi²³** [194]

BORO **nôi**; KOKBOROK **nɯy**

CHEPANG **nis-joʔ**; GURUNG **ngihq**; JIREL **nyiq**; KAIKE **nghyi**; KANAWARI **nish**; KHAM **neh-plo**; MAGARI **nis**; NEWARI *ni*-**gu**(-**li**); SHERPA **ngyi**; SIKKIM BHUTIA **nyi**; TAMANG **nyi:h**; THAKALI **ngih**; THULUNG RAI **nə**; TIBETAN (Lhasa) **nyii**

ERSU **nɛ⁵⁵**; BAI **nɛ²**; PUMI **ni²³**; QIANG **nyi⁵⁵**

AO (Chungli) **ana**; KIMSING **anai**; CHANG **nyi**; KONYAK **i** (with palatalization of the initial); LIANGMAI **nia**; LOTHA **eni**, **oni**; MEITHEI **ani**; PHOM **nyi**; SANGTAM **anyü**; TANGSA (Moshang) **ani**, (Yogli) **anei**; WANCHO **an**(**y**)**i**

GANGTE **nih**; KOM REM **hni**; KUKI **ni**; LAKHER **no**; PAITE **nih**; THADO **ni**; TIDDIM **nih**

We must include here JINGPHO **nī**, a bound form which is never used in isolation, but only in certain set expressions like **nī náʔ** 'two nights' and round numbers like **nī tsā** '200'. It has also been grammaticalized into a plural or collective suffix, e.g. **gwì nī** 'the dogs'. The independent Jingpho numeral for TWO is the mysterious **ləkhôŋ** (below 4.115).

4.113　Forms with other than velar prefixes

YIMCHUNGRU has a curious form **manie**, its only numeral to carry a **ma**- prefix.

MIKIR has **hini**, its only numeral with a **hi**- prefix.

MARAM and NOCTE have prefixes of the type CVC- with the lower numerals 1-3: '2' MARAM **hang-na**, NOCTE **va-nyi**, **wan-ni**.

A few Chin languages have a general prefix **pa**- used with all numerals:

---

193) Something very similar happens in LAHU, where /y/ acquires local friction before the high front vowels /i,e/, becoming a voiced slit spirant. See Matisoff 1973a:5-6.

194) J.T. Sun [1993:319,340,463] reconstructs PTani *ñi. W.T. French [1983:572] reconstructs PNorthern Naga *ʔ-ni, with the *glottal prefix apparently motivated by the vocalic prefix in the Yogli, Moshang, and Wancho forms.

HMAR **pahni**, LUSHAI **pahnih**, VAIPHEI **pani** (below 5.43).

4.114   Forms with velar suffix
*ni-k

Several languages have forms pointing to the rhyme *-ik. We have already mentioned WB **hnac**, which could descend either from PLB *-ik or *-it (cf. WB **hrac** '8' < PLB *?rit or *?ryat, but also WB **chac** 'joint' < PLB *tsik).

Forms from Abor-Miri-Dafla and Himalayish also attest to the antiquity of a velar suffix with this numeral:[195]

HAYU **nak-pu** [for humans], **na?ung** [for non-humans] [MICHAILOVSKY 1981:167]; SUNWAR **nik-syi**; BAHING **nik-si**; SHARCHOP [CHHEWANG RINZIN] **nyik-tsing**; MONPA (Motuo) **nyik-tsing**; SERDUKPEN **n(y)ik**; THULUNG **nək**

The second syllable of the SUNWAR and BAHING forms certainly seem related to the -**tsing** in SHARCHOP and MONPA, though their wider affiliations are still unknown.

4.115   Unusual forms

There remain a few strange forms that we here assemble for simultaneous delectation:

JINGPHO	ləkhôŋ	[no known etymology]
KHALING	saːh-pu	[apparently contaminated by **suh-pu** '3']
BUMTHANG	zon	

This Bumthang form is virtually identical in appearance to SHARCHOP **zon** and CENTRAL MONPA **zum** "SEVEN". The Khaling forms are cited as **saakpu** '2' and **sukpu** '3' in Toba and Toba 1975, and are clearly cognate to DUMI **sak** '2' and **sukli** '3'.

4.12   Profile of number THREE
THREE   *g-sum

This is perhaps the most stable of all TB numerals, with the fewest forms that do not conform to a relatively simple prototype. No doubt this stability is largely due to the relative conservatism of consonants like s and m, along with the conceptual saliency of the number THREE itself.

Prefixally speaking, this is also a very consistent numeral. The only prefix of obviously long standing with this etymon is the velar, *g-/*k-. Secondarily, a number of AMD and Kuki-Naga languages have developed vocalic prefixes which run through most or all of the lower numerals, and of course THREE is affected like the others (below 4.122). Most TB languages, however, now show

---

195) STC (notes 60 and 61, p. 16) fudges on whether suffixal *-k is to be set up for this root, deciding finally that it is not; but the AMD and Himalayish forms settle the matter in the affirmative as far as I am concerned.

no overt trace of any prefix at all with this root.

STC sets up the proto-vowel as *-u-, and this is the reconstruction of choice.[196] Many daughter languages (including Chinese) have -a- vocalism, however, and perhaps this cannot be dismissed as secondary in all cases — i.e. we may ultimately be forced to recognize *-u- ✕ *-a- variation in this etymon, a pattern already noticed in a number of another etymologies (e.g. 'fragrance' PTB *b-suŋ ✕ *b-saŋ [STC #405]).[197]

4.121　Forms with velar prefix

WRITTEN TIBETAN gsum

CHULIKATA kaːsh (with preemption via apocope of the rhyme); DENG DARANG kə²¹suɯng⁴⁵; DENG GEMAN kɯ²¹səm⁵³; IDU [SUN 1983] ka³¹ song³⁵; MIJU MISHMI ksam; TARAON kaːsaːng

KEZHAMA katsü; KHOIRAO kathum; MAO kosü; MARING khiyum; MELURI keche; MIKIR kethom; MZIEME ketsum; NRUANGHMEI kathum; NTENYI keching, keshang; POCHURY küche; RENGMA keshan; SEMA küthu; TANGKHUL kathum; ZELIANG and ZEME kechum

DIMASA gatham; GARO git(t)am

rGYARONG kesom ~ kesam is not criterial, since all its primary numerals (except 1 and 8) have acquired a secondary velar prefix (below 5.2).

4.122　Forms with no overt trace of a consonantal prefix

MIRI (=MISING), GALLONG, MINYONG, and TAGIN a-um; ABOR (=PADAM) a-ŋum; AKA zu; APATANI hĩ; DAFLA (=NISHI) (a-) om, um; BOKAR a-hum; LHOPA aɦium[198]; MILANG ham; MONPA (Motuo) sam, (Dubey and Cuona) sum; SERDUKPEN ung

ANGAMI se; CHAKHESANG süh; CHANG sam; CHOKRI sü; LIANGMAI shum; PHOM jam; WANCHO a-jam, a-zam; AO (Chungli) asem, (Mongsen) asam; KIMSING acam; LOTHA etham; MEITHEI ahum; SANGTAM asang; TANGSA (Moshang) atum, (Yogli) adim; YACHAM-TENGSA and YIMCHUNGRU asam; NOCTE van-ram, wan-rʌm; KONYAK lem;

LAKHER thô; GANGTE, KOM REM, KUKI, THADO, PAITE,

---

196) See STC #409 (pp. 28, 75, 81, 94, 131, 136, 142, 143, 152, 153, 162, 169, 170, 181, 182, 186-187, 196).

197) See STC n. 486 (last 3 lines on p. 186 and continuing on p. 187). The Middle Chinese vocalism in this root is generally considered to be "irregular" (cf. STC n. 436, p. 162). See also, e.g., *s-riŋ ✕ *s-raŋ 'thousand', above 3.547(b).

198) The voicedness of the ɦ in this Lhopa form is perhaps an indirect reflection of the influence of the *g- prefix on the voiceless sibilant root-initial. Alternatively, the younger a- prefix (see below 5.512) might have caused the voicing. J.T. Sun [1993:110] sets up Proto-Tani *ɦium for this root.

PUIRON, and TIDDIM **thum**
RAWANG [BERNARD 1934] **ətsum**; DULUNG **a²¹ sɯɯm⁵³**
BORO and KOKBOROK **tham**
ERSU **si⁵⁵**; PUMI **sãu²³**; QIANG **tshi⁵⁵**; BAI **sa¹**
CHEPANG **sum-joʔ**; GURUNG **soq**; HAYU **tshuk-pu** [for humans],
tshuʔ-ung [for non-humans]; JIREL **sumq**; KAIKE **sum**; KANAWARI
shum; KHALING **suh-pu**; KHAM **sohm-lo**; LEPCHA **sam**; SHARCHOP
sam; MAGARI **som**; NEWARI **swa-gu:**; SHERPA **sumq**; SIKKIM
BHUTIA/DANJONGKA **sum, sung**; SUNWAR **sã:**; TAMANG and
THAKALI **som**; THULUNG RAI **sium**; TIBETAN (Lhasa) **sum**
KAREN (Pa-O) **som**, (Pwo) **ɔ-ən**, (Palaychi) **tyq**, (Sgaw) **ɔ-ə**
PROTO-LOLO-BURMESE ***sum** (Tone *2) > WB **sûm**, LAHU **šɛ̂ʔ** ~
šɛ̂, etc.

The glottal stop in LAHU is secondary, a junctural feature that arose to
separate THREE from FOUR in counting, since Lahu **ɔ̂** '4' has a vocalic
onset.[199] For a similarly secondary final laryngeal in this etymon, cf. the
HAYU form **tshuʔ-ung** just cited.

4.123 Forms with miscellaneous secondary prefixes

Several Chin languages have developed a general numeral prefix, **pa-**, used
with all the primary numerals: HMAR, LUSHAI, VAIPHEI **pathum** (below
5.43).

A couple of Naga languages have developed a fully syllabic CVC- prefix
with the lower numerals (below 5.512): MARAM **hang-tum**; NOCTE **van-ram**.

Most strikingly, JINGPHO has developed a nasal prefix with this numeral
— **məsum**. This is part of a celebrated JINGPHO run of the labial nasal prefix
from THREE to FIVE (**məsūm, məlī, məŋā**). Clearly it is not THREE that is
influencing the two higher numerals, but vice versa, since both FOUR and
FIVE have been associated with labial prefixes from PTB times. See below
4.13, 4.14.

4.13 Profile of number FOUR[200]

FOUR   ***b-liy** or ***b-ləy**
The following forms are cited in STC #410:[201]
WRITTEN TIBETAN **bźi**; THULUNG **bli**; KANAURI **pö**; MAGARI
buli; DIGARO **kəprei**; MIRI **pi**; NUNG **əbyi, əbəli**; JINGPHO **məli**;
WRITTEN BURMESE **le:**; MARU **byit** (with secondary final stop);

---

199) See above 4.01(e).
200) This numeral plays interesting conceptual roles in some TB languages. We have seen how the
    Boro system is basically quaternary (above 3.32[C]); many AMD languages express EIGHT as
    a multiplicative formation "4×2" (below 4.237).
201) See STC pp. 33, 61, 88, 91, 94, 104, 111-112, 131, 152, 158, 171-172, 180, 196.

MIKIR **phli**.

The weak root-initial lateral in this etymon offered little resistance to the preemptive propensities of the prefix (below 4.131).

The *b- that goes with FOUR is one of the best-attested of all numeral prefixes in TB. To the forms given in STC, add:

APATANI **pilye, pɯ-lje** (also **pe**, with preemption); MONPA (Cuona) **pli**[53], [DUBEY] **blee**, (Central) **b**(i)**ci** ∼ **p**(i)**ci**; SERDUKPEN **bi:si**[202]
AO (Chungli) **pezü**, (Mongsen) **phüli**; KEZHAMA **pedi**; KIMSING **balai**; KONYAK **peli**[203]; LAKHER **pali**; MAO **padei**; MARING **phili**; MIKIR **phli** (also **phir** [MARRISON], with metathesis and/or apocope); MUKLOM TANGSA **balee**; NOCTE **beli**; NRUANGHMEI **padei**; RENGMA **pezi**; SEMA **bidhi**; TANGSA (Moshang) **bali**, (Yogli) **bəlai**; YACHAM-TENGSA **phale**; YIMCHUNGRU **phiyi**
DULUNG **a**[21] **bli**[53] (with secondary vocalic prefix)
BORO **brôi**; DIMASA **biri**; GARO **bri**; KOKBOROK **brɯy**
CHEPANG **play.joʔ**; HAYU **b**(l)**iʔung**; GURUNG **plihq**; TAMANG and THAKALI **plih**; THULUNG RAI **blə**; SHARCHOP **pshi**; DUMI **balikpi**
In several Chin languages (e.g. HMAR, LUSHAI, VAIPHEI **pali**) the **pa-** prefix is secondary, used as a general prefix with all the numerals. See below 5.43.

4.131   *With preemption of the initial by the labial prefix:*

ABOR-MIRI **a-pi**; APATANI **pe** (also **pilye** [for counting humans]); DAFLA **api**; GALLONG **appi**; LHOPA **api:**; MILANG **pe**; MONPA (Motuo) **p'i**; NISHI and PADAM **appi**; TAGIN **epi**[204]
KHALING **bha:el** (alongside **bho:m** '5', also with preemption); KANAWARI **pü**; NEWARI *pe*-**gu:**

A rare "survival via preemption" of the *b- prefix in Burmish is MARU **bit**. (The development *-iy (= *əy) > MARU -it is regular, as is the parallel *-uw (= *əw) > MARU -uk.)[205] The same development has taken place in the moribund UGONG language, where a more conservative form **pli** now varies with **pi**.[206]

4.132   *With nasal (not stop) prefix: *m-ləy < *b-ləy*

As with FIVE (below 4.14, 4.142), many languages have a labial nasal

---

202) J.T. Sun [1993:124] reconstructs PTani *pri.
203) This is the only Konyak numeral that carries a prefix. W.T.French [1983:492] reconstructs Proto-Northern Naga *bələy.
204) Many of these AMD forms have acquired a secondary vocalic prefix, after the preemption.
205) This point has been much discussed in the literature, sometimes with acrimony. See STC, p. 60.
206) This Burmish language, formerly known by the misnomer "Kanburi Lawa", has been rediscovered by D. Bradley. See, e.g. Bradley 1978.

instead of a labial stop prefix with FOUR. STC regards this nasalization of the prefix as a secondary development, and does not push the stop ~ nasal variation back to the PTB stage.

> JINGPHO **məli**; KHOIRAO **malhi**; KOM REM **manli**; LIANGMAI **madai**; LOTHA **mezü**; MARAM **madai**; MEITHEI **mari**; MELURI **mezu**; MZIEME **m(a)dai**; NTENYI **mez(h)ü**; POCHURY **mzü**; PUIRON **mali**; SANGTAM **müzyü**; TANGKHUL **mati**; ZELIANG **mdai**; ZEME **medai**

4.133   With labial fricative prefix:
AKA (= HRUSO) **fi-ri**; LEPCHA **fəli**

4.134   Forms that do not overtly reflect any consonantal prefix

KAREN:	Pa-O **lit** (with suffix; above 4.01[d]); Pwo **li, liʔ**; Palaychi and Sgaw **lwi**
HIMALAYISH:	KAIKE **li**, BAHING **le**, KULUNG **li-chi**, LIMBU **li-si**
CHIN:	GANGTE, KUKI, PAITE, THADO, TIDDIM **li**
N.NAGA:	CHANG **lei**; PHOM **a-li**; WANCHO **a-li**

4.1341   Forms that indirectly reflect a consonantal prefix
PLB *hləy² > WB **lê**; LISU **li⁵⁵**; YI (Dafang) **łi³³**; NAXI (Lijiang) **lu³³**; MPI **li⁶**; LAHU **ɔ̂(n)** [207]; AKHA **ɔ̀**; BISU **ha**; PHUNOI **hàn** (with rhinoglottophilia), etc.

4.1342   With replacement of the root-initial lateral by a stop [208]
ANGAMI (Khonoma) **da**, (Kohima) **die**; CHAKHESANG **daa**; CHOKRI **da**

4.135   With fusion of prefix and initial to a spirant/affricate
NAXI (Moso) **ʐv³³**
JIREL **syi**; SHERPA **ji**; SIKKIM BHUTIA/DZONGKHA **zyi ~ syi**
ERSU **ʐo³³**; PUMI (Taoba) **ʐʅ⁵⁵**, (Qinghua) **ʒɛ⁵⁵**; QIANG (Taoping) **dʒʅ⁵⁵**, (Mawo) **gʐə**; ERGONG **wʑɛ**; MUYA **ʐʅ³⁵**; QUEYU (Yajiang) **ʑi³⁵ tɕã⁵³**; GUIQIONG **tsʅ⁵⁵**; NAMUYI **zɿ³³**; SHIXING **ʒuʙ³³**
BAI [DELL 1981] **sɯ²**, (Dali, Jianchuan) **ɕi⁴⁴**, (Bijiang) **si⁴⁴**

---

207) The unusual Lahu vowel reflex is regular, with several parallel examples that have been much discussed in the literature. See my note 195 and notes 263-264 in STC, pp. 61,91. The most extensive recent treatment is in Matisoff 1994a:46-50. The optional nasalization of the vowel is due to "rhinoglottophilia" after the zero-initial [see MATISOFF 1975].

208) For a treatment of the relatively widespread phenomenon of d ≈ l interchange in ST/TB, see Matisoff 1990b, "The dinguist's dilemma."

4.136　With velar prefix:

*g-ləy >	MINYONG	aki (with preemption)
*g-b-ləy >	rGYARONG (Zida)	kewdzyi
		(with lenition of the *-b- to -w-)
	IDU MISHMI	kapri
	TARAON	kaːpraːi
	DENG-DARANG	kə²¹ pʁəi⁵⁵
	CHULIKATA	kaːppi
	DIGARO	kəprei

With this last allofam also belong MIJU MISHMI **kambran**, DENG GEMAN **kɯ²¹bʁɯn⁵³**, evidently with a fully syllabicized prefix and secondary nasalization in both syllables: ***g(N)-b-ləy-(N)**.[209]

4.14　Profile of number FIVE

FIVE　***l-ŋa ⋊ *b-ŋa**

In STC #78[210] the following forms are cited in support of the reconstruction ***l-ŋa ⋊ *b-ŋa**:

WT **lŋa**; JINGPHO **məŋa**; WB **ŋâ**; GARO **boŋa**; LUSHAI **ŋa ~ pəŋa**.[211]

Many TB languages have forms for FIVE with labial *nasal* prefix, **mV-**, rather than with a labial *stop* prefix, as the proto-prefix is supposed to have been. STC claims that the nasal developed secondarily from the stop (e.g. Jg. **məŋa** < ***b-ŋa**). Yet an ***m-** prefix is independently required for PTB beyond any question, both with nominal and verbal roots. It seems to me quite likely that stop ~ nasal prefixal variation existed in this etymon *already at the proto-level* — a situation one might expect, given that the root-initial itself is a nasal.[212]

Since FOUR is also reconstructed with ***b-**, 4-5 constitutes a *proto-prefix run* (above 1.1; below 5.51).

The lateral prefix is much less well attested in TB as a whole ("***b-** is much more generally represented": STC p. 94) and one feels that STC relied overmuch on the testimony of Written Tibetan **lŋa** in reconstructing ***l-** for the proto-language. One possible origin for the **l-** would be the widely distributed root ***lak** [STC #86] 'HAND', given the well-known fact that hands have five fingers.[213]

---

209) Cf. for example the syllabic prefix **gùm-** that JINGPHO developed in its word for HORSE: **gùmrà(ŋ)** < ***m-raŋ**.

210) See STC, pp. 31, 54, 58, 94, 112, 131, 137, 152, 162, 186, 187, 196.

211) It should be noted that the LUSHAI form with **pə-** is of no weight in reconstructing ***b-**, since it is a late, secondary prefix used with *all* the numerals (below 5.43).

212) The case is somewhat less strong for positing ***b- ⋊ *m-** variation at the proto-level for FOUR (above 4.132).

213) Cf. Proto-Austronesian ***ka-lima** 'five; hand', as well as evidence for a "hand-based" quinary numeral system in some Himalayish languages (below 4.15).

Other TB forms that point to a liquid prefix for FIVE are:
"OLD KUKI" *r-ŋa (e.g. Rangkhol **ringa**) [STC p. 94]
KOM REM **ranga**
MONPA (Cuona) **le²¹nge⁵⁴**, (Northern) **lyange**
The MILANG form **rangal** '7' looks like it has been heavily contaminated by a lateral-prefixed version of FIVE, though the final -**l** is a problem (above 4.01b). Cf. also the inner lateral prefix in PADAM **pilngo** (below 4.144), and the MIJU and KAMAN forms in **k-l-** (below 4.148).

4.141   Forms with labial stop prefix

DIMASA and GARO **bonga**
CHEPANG **ponga:-jo?**
MILANG **pangu**
DULUNG **pɯ²¹nga⁵³**; RAWANG [BARNARD 1934] **hpung-nga**
The following Kuki-Naga forms all participate in FOUR/FIVE prefix runs (see above 4.13; below 5.2, 5.4).

AO (Chungli) **pungu**, (Mongsen) **phanga**; KEZHAMA **pangu**; KIMSING **bangi**; LAKHER **pangaw**; MAO **pongo**; MARING **phanga**; MIKIR **phongo**; MUKLOM TANGSA and NOCTE **banga**; NRUANGHMEI **pangu**; SEMA **pongu**; TANGKHUL **phanga**; TANGSA **banga**; YACHAM-TENGSA **phungu**; YIMCHUNGRU **phüngü**

In the following "Angamoid" languages, FOUR is an unprefixed monosyllabic form, so that these words for FIVE are isolated with their labial prefix:

ANGAMI (Khonoma) **pengu**, (Kohima) **pengou**; CHAKHESANG **puhngu**; CHOKRI **püngu**

4.1411   With preemption of the root-initial by the labial prefix:

*b-[ŋ]a >	BORO	ba[214]
	KOKBOROK	ba[215]
	RENGMA	pfü

Several forms have been uncovered where the preemption served to compensate for the apocope of the root vowel (see below 5.131):

	KHALING (Him.)	bho:m
	PUIRON (KCN)	pang
	AKA (AMD)	phum

4.1412   With secondary labial prefix

A few Chin languages have developed **pa-** as a universal numeral prefix,

---

214) Cf. the other Boro preemptive form **do** '6' < *d-ruk.
215) Kokborok, like Boro, also shows preemption in its form for '6', dɔk.

used with all the numerals even if they already bear a prefix inherited from PTB (below 5.43). Naturally these languages cannot be used as evidence for the ancient status of *b- with this numeral:

HMAR, LUSHAI, VAIPHEI **panga**.

4.142 Forms with labial nasal prefix

IDU [Sun 1983] **ma³¹nga³⁵**; MISHMI [Dubey] **manga**; TARAON **maːngaː**; DENG DARANG **ma²¹nga⁴⁵**; CHULIKATA **maːngaː**; rGYARONG (Zida) **kemnga** (with superadded ke-)

All the **m**-prefixed Kuki-Naga forms for FIVE occur in languages which also have forms for FOUR with the **m**- prefix; i.e. all these words for FIVE participate in 4-5 prefix runs (above 4.132):

KHOIRAO **manga**; LIANGMAI **mangiu**; LOTHA **mungo**; MARAM **mingu**; MEITHEI, MELURI **manga**; MZIEME **mengei**; NTENYI **münga**; POCHURY **mnga**; SANGTAM **münga**; ZELIANG **mengei**; ZEME **mengeu**

4.1421 With preemption of the root-initial by m-

BAI (Dali) **mu³**

4.143 Forms with labial spirant prefix:

LEPCHA **fəngo** (alongside **fəli** '4')

4.144 Doubly prefixed forms:

*b-l-ŋga > PADAM/ABOR **pilngo** (see LSI III.1, p. 622); SHIMONG ADI **pi-ri-ŋo**

4.145 Forms that show no overt trace of a consonantal prefix:

DZONGKHA/SIKKIM BHUTIA **nga**; GURUNG **ngaːhq**; JIREL **ngaːq**; KAIKE **ngaː**; KANAWARI **nga**; SHARCHOP **nga**; NEWARI **nyaː-guː**; SHERPA, TAMANG and THAKALI **ngaːq**; THULUNG RAI **ngo**; DUMI **ŋo**; KULUNG **ŋa-chi**; LIMBU **ŋa-si**

CHANG **ngau**; KONYAK and PHOM **nga**; WANCHO **aga** (with denasalized initial)

GANGTE, KUKI, PAITE, TIDDIM **nga**

MONPA (Motuo) **nga**; APATANI **ngo**; in this category we may include other AMD forms with vocalic prefix, e.g. ABOR-MIRI, DAFLA, GALLONG, MINYONG, TAGIN **ango**; NISHI **aːngo**; LHOPA **ongo**; BANGNI **u-ŋu**

ERSU **nguar³³**; Muya **ŋɑ³⁵**; Queyu **ɲua-tɕã⁵³**; Guiqiong **ŋẽ³⁵**; Namuyi **ŋa³³** PROTO-LOLO-BURMESE *ŋa² > WB **ŋâ**; LAHU **ŋâ**; LISU **ŋwa**; MPI **ŋo²**; BISU **ngà** ~ **hà**; PHUNOI **ʔàn** (with rhinoglottophilia), etc.

BAI (Jianchuan, Bijiang) **ŋv³³**
KAREN (Pa-O) **ngat** [with suffix]; other Karen dialects have lenited the nasal to a palatal semivowel: Pwo **jɛ, jaiʔ**; Palaychi and Sgaw **jɛ**

4.146   With spirantized initials:

In several QIANGIC languages the velar nasal has become a velar or "uvular" fricative:
PUMI **ɣuã**; QIANG **ʁua³³**; Shixing **fiɑ̃**

4.147   With velar prefix and apocopated root-vowel (?)

Two weird forms from AMD can perhaps be accounted for if we assume a variant secondary velar prefix and apocope of the root vowel, something like *g-l-ŋ[a]:
KAMAN   **kɯ²¹len⁵⁵**; MIJU **klin**

4.148   A contaminated form in SERDUKPEN

SERDUKPEN **khu** '5' seems to have undergone contamination of its initial by **khit** '6'. Curiously enough, this word for '6' has itself undergone the influence of the next higher numeral **sit** '7' (above 4.01c, below 4.218).

4.15   Traces of quinary numeral systems in Himalayish

Several Kiranti languages of E. Nepal have numeral systems where the number FIVE has an explicit morphological relationship with their word for HAND or FINGER:[216]
BANTAWA (Raniṭar dialect): **chuk** 'HAND', **ükchuk** 'FIVE' ("1 × 5"),
   hüachuk 'TEN' ("2 × 5")
MEWAHANG: **huk** 'HAND', **ihuk** "FIVE"
CHOURASE: **brem** 'FINGER', **kollabremci** 'FIVE' (**kolo** 'one')
YAKKHA: **mukta** 'HAND', **muktapi** 'FIVE'
Yakkha in fact has the most thoroughgoing quinary system reported so far for any TB language, with interesting additive and subtractive features (discussed below 4.20) which make it look quite a bit like Khmer.

We have noted (above 4.11) the virtual identity of the PTB forms for TWO (*g-nis) and SEVEN (*s-nis). Although these quinary formations in Kiranti appear to be of quite recent origin,[217] they at least demonstrate that the idea of counting by fives still occupies a niche in TB conceptual space.

---

216) See the charts in Gvozdanović 1985:135-136. As noted above (n. 213), such a constellation
   of ideas is also found in Austronesian (PAN *ka-lima 'hand; five').
217) Such is the opinion of Gvozdanović [1985:137].

4.2 The Higher Numerals: SIX to NINE

4.20 Additive, subtractive, and multiplicative formations

There may once have been a certain mnemonic advantage in forming the more "remote" higher numerals additively, subtractively, or multiplicatively in terms of other, more "familiar" numerals. Conceivably it was easier for early French speakers to call '70' *soixante-dix* ("60 + 10") instead of *septante* — the speaker struggling to keep count at a numerical level far surpassing the number of his fingers and toes need only have run through the teens over again, keeping the TENS place constant, from *soixante-onze* '71' ("60 + 11") through *soixante-dix-neuf* '79' ("60 + 19"); similarly for *quatre-vingt-dix* '90' ("80 + 10", literally "[4×20] + 10") through *quatre-vingt-dix-neuf* '99' ("[4×20] + 19").

TB languages also provide many examples of these phenomena, some of which have already been mentioned:

4.201 Multiplicative formations

[A] In the quaternary system of BORO (above 3.32c), EIGHT is expressed as "4×2" (**zokkay-nəy**), with the special morpheme **zokkay** 'group of four' (rather than with the unrelated cardinal numeral **brə** FOUR). The other numerals between five and ten are formed additively (below 4.203).

[B] Many AMD languages have similar multiplicative expressions for EIGHT, e.g. Apatani **a-pi** '4', **nyi** '2', **pih-nyi** ∼ **pryih-nyi** '8' (see below 4.237).

[C] There are cases where a confusion between FOUR and EIGHT is evident from a comparison of closely related dialects. In the Annapurna subdialect of the Bhojpur dialect of BANTAWA, **retkapok** means '4'; but in Chhinamakhu subdialect of Bhojpur it means '8' [GVOZDANOVIĆ 1985:136].

[D] As we have seen, in Eastern KAYAH (=KARENNI=RED KAREN) the numerals SIX and EIGHT are expressed as doubles of THREE and FOUR (**sō** 'three', **sō swá** 'six'; **lwī** 'four', **lwīswá** 'eight'), while SEVEN and NINE are in turn additively formed from SIX and EIGHT (**sō swá tə-** 'seven' "[3×2] + 1"; **lwīswá tə-** 'nine' "[4×2] + 1").

[E] Perhaps MIJU MISHMI **katam** '6' is multiplicatively based on **ksam** '3'. The aberrant MILANG language has **ham** '3' and **sap** '6'.

[F] According to Hodgson's data (reproduced in LSI III.1:384), HAYU once had a numeral **chhu-ning** '6' that was a multiplicative formation based on **ning** '2' ("3×2"). At this same period there was also a numeral **u:-ning** '5', that looks like it was influenced by '6'.

4.203 Subtractive formations

Subtractive numeral expressions typically involve the two highest units EIGHT and NINE, and/or higher numbers ending in them (18, 19; 28, 29; 38, 39, etc.), and/or the odd round numbers (30, 50, 70, 90). Examples may readi-

ly be found in Indo-European languages. LATIN has subtractive forms for '18' and '19' (**duodēvīginti** "2 from 20" and **undēvīginti** "1 from 20").[218] In SANSKRIT, pairs of alternative forms exist for '19' and the other higher numbers ending in nine; either an ordinary additive form with respect to the next lower round number, or a subtractive form in terms of the next higher round number:

SANSKRIT

'17'   **saptadaśa**

'18'   **aṣṭādaśa**

'19'   **navadaśa** ("9 + 10") or **ūnaviṃśati** < **ekonaviṃśati** ("one diminished 20" [p.c. Robert P. Goldman 1994])

'20'   **viṃśati**

'40'   **catvāriṃśat**

'49'   **navacatvāriṃśat** ("9 + 40") or **ūnapañcāśat** ("one-diminished 50")

'50'   **pañcāśat**

We have already discussed subtractive formations for the higher teens in TB languages (above 3.316), as well as subtractive ways of expressing the odd round numbers in TB vigesimal systems, e.g. in CHANG (above 3.524) and in DZONGKHA (above 3.534[C]). As far as the basic unit numerals themselves are concerned,[219] the best examples of subtractivity in TB are to be found in MIKIR (above 3.523; below 5.42[C]) and MEITHEI (above 3.232; below 5.445):

	MIKIR	*MEITHEI*
ONE	isi	ama
TWO	hini	ani
EIGHT	nirkep	nipal ∼ nipan
NINE	sirkep	mapan
TEN	kep	(tara)

4.204   Additive formations of the basic unit numerals [220]

KHMER is a striking example of a language where all the higher numerals from SIX to NINE are formed additively on the basis of FIVE:

---

218) Roman numerals in their written form make extensive use of the subtractive principle, e.g. FOUR "IV" (1 from 5), NINE "IX" (1 from 10), FORTY "XL" (10 from 50), etc. These must have been a lot easier to carve in stone than non-subtractive alternatives like "IIII", "VIIII", or "XXXX".

219) Again looking beyond Sino-Tibetan, cases of subtractive EIGHT and NINE, while not exactly frequent, are attested in language families around the world, e.g. Finnish **üksi/ühde** 'one', **ühdeksan** 'nine'; **kaksi, kahde** 'two', **kahdeksan** 'eight' (p.c., Adam Jacobs 1992); Indonesian **sembilan** 'nine' ("taking one [from ten]' < **ambil** 'take away', **se-** 'one'); **delapan** 'eight' (prob. "[taking] two [from ten]" < **dua** 'two').

220) We have already mentioned additive formations involving the round numbers of vigesimal ↗

KHMER[221)]

ONE	muəy	SIX	prammuəy
TWO	pii	SEVEN	prampii
THREE	bəy	EIGHT	prambəy
FOUR	buən	NINE	prambuən
FIVE	pram	TEN	dɑp

As we have seen (e.g. above 4.11), there is some evidence for a similar relationship at the PTB level between TWO (*g-nis) and SEVEN (*s-nis), though not between THREE/EIGHT or FOUR/NINE. In the peculiar quaternary system of BORO (see above 3.32[C]), '6' and '7' are expressed as "[4×1] + 2" and "[4×1] + 3", while '8' is "4×2" and '9' is "[4×2] + 1":

BORO

ONE	se	SIX	*zokkay-se kanəy*
TWO	nəy	SEVEN	*zokkay-se katam*
THREE	tam	EIGHT	*zokkay-nəy*
FOUR	brə	NINE	*zokkay-nəy kase*
FIVE	ba	TEN	*dos* [< Indo-Aryan]

An isolated case of a basic numeral being derived additively from the next lower one is to be found in MIKIR, where **throk-si** '7' is formed from **throk** '6' plus **isi** '1' (below 4.229).[222)] However, it is in the "endangered" numeral systems of certain Kiranti languages that the most elaborate additive sets of basic numerals have been reported, e.g. in the Raniṭar dialect of BANTAWA:

BANTAWA (Raniṭar dialect)

ONE	ũk-ṭa	SIX	bhan-ka-chuk
TWO	hũa-ṭa	SEVEN	bhan-hũ-chuk
THREE	sum-kaṭ	EIGHT	bhan-sum-chuk
FOUR	reṭ-kaṭaṭ	NINE	bhan-reṭ-chuk
FIVE	ũk-chuk	TEN	hũa-chuk

As we have seen (above 4.15), FIVE is here expressed as "1 × 5/HAND" and TEN is "2 × 5/HAND". SIX through NINE are additive formations based on FIVE. (Presumably **bhan-** means something like 'add to', and the second syllable in SIX, **-ka-** is a "suppletive allomorph" of **ũk** 'ONE'.)

Even more spectacular is the YAKKHA system, where the use of non-numerical morphemes like HAND and FOOT, along with a variety of multiplicative, subtractive, and additive strategies, enable the language to express the entire basic set of numerals *with only three native numerical TB roots (1-3)*, without having recourse to any Nepali loans:

\ systems (above 3.524); and of course in connection with the teens and other higher numbers ending in 1-9, where additivity between the ten and the unit is the norm (above 3.3).

221) Huffman 1970:25.

222) This is quite comparable to Russian **semj** '7', **vosemj** '8'.

YAKKHA [GVOZDANOVIĆ 1985:137]

ONE	**kolok**	SIX	**muktapi usongbi kolok**
TWO	**hitci**	SEVEN	**muktapi usongbi hitci**
THREE	**sumji**	EIGHT	**muktapi usongbi sumci**
FOUR	**sumcibi usongbi kolok**	NINE	**mukcurukbi kolok hongbi**
FIVE	**muktapi**	TEN	**muktapi hita**
	TWENTY	**langcurukmukcuruk**	

Here FOUR is expressed as "3 + 1" (**usongbi** 'plus'), while FIVE is the root for "HAND" (**mukta**), and TEN means "HAND×2". SIX through EIGHT are additive formations based on FIVE ("5 + 1", "5 + 2", "5 + 3"), and NINE is a subtractive expression "HANDS minus ONE", based on TEN (**-curuk-** 'plural', **muk-curuk** 'hands' [i.e. the number of fingers on both hands], **hongbi** 'minus'). Logically enough, TWENTY is simply the compound "HANDS and FEET" (**lang** 'foot')!

4.21 Profile of number SIX

SIX   *d-ruk* (STC)/*d-k-rok* (JAM)

In TSR #35 I reconstruct this etymon as Proto-Lolo-Burmese *C-krok*, where "C-" stands for a voiced prefix for which there is evidence on tonal grounds, and which I interpret as pointing to PTB *d-krok*.

The problem is that many TB languages reflect a VELAR prefix (or at any rate a velar component in the prevocalic part of the syllable), instead of — or in addition to — a dental one. STC (note 321, p. 115) attempts to account for this in terms of regular sound change of an initial *dental group to a velar one, i.e. "prefixal *d-r-" > Written Burmese **khr-** (vs. "cluster *dr-" > WB **khy-**), but this is not very convincing.[223] In my view, both a velar and a dental element must be recognized at the PTB level.

4.211 Forms that reflect a dental (but no velar) before the -r

STC #411 cites the following 6 forms:

WT **drug**; KANAURI **ṭug**; LEPCHA **tərək**; DIGARO **thərɔ**; GARO **dok** (with preemption of the -r-); MIKIR **therok**.

To these we may add:

BORO **do** (with preemption; cf. Boro **ba** '5' < *b-[ŋ]a); DIMASA **do**; KOKBOROK **dok**

DZONGKHA **druʔ**; JIREL **ṭhuk**; SHERPA **ṭuk**; SIKKIM BHUTIA **ṭuk**; GURUNG **ṭuhq**; TAMANG **ṭuːh**; THAKALI **ṭuh**; KULUNG **tuk-chi**; LIMBU **tuk-si**; CHAMLING **tukara**

---

223) STC #411, and pp. 41, 45, 75, 76, 82, 83, 88, 94-95, 114, 115, 116, 141, 146, 154, 161, 162, 171, 182.

AO t(e)rok; KIMSING tarok; LOTHA tirok; MARING tharuk; MEITHEI taruk; MELURI taro; MIKIR throk; NTENYI togho, tüo; POCHURY toro; SANGTAM thüro; TANGKHUL tharuk; TANGSA (Moshang) taruk, (Yogli) türuk; YACHAM-TENGSA thelok; YIMCHUNGRU thruruk

The affricate-initialled prefixes in the following KCN forms are all parts of prefix runs affecting the higher numerals of these languages as a whole (see below 5.44):

LAKHER charu; LIANGMAI charuk; MAO choro; NRUANGHMEI cünei; RENGMA tsaro; SEMA tsogho

4.212　Forms that reflect a velar (but no dental) before the -r:

*kruk > WRITTEN BURMESE khrok

ACHANG	xẓoʔ⁵⁵
LANGSU	khjauk⁵⁵
ZAIWA	khjuʔ⁵⁵
NUSU	khʲu⁵³
NAXI (Yongning)	khəˀ¹³
JINGPHO	kruʔ⁵⁵
TRUNG	k'lu⁴⁴
NEWARI	*khu*-guː
MONPA (Cuona)	kroʔ⁵³, (Dubey) gro
KOM REM	karuk
PUIRON	keruk[224]

J.T. Sun [1993:132] reconstructs Proto-Tani *krə, on the basis of APATANI xrjɯ, BENGNI a-kjɯ, BOKAR (=ADI=LHOBA) a-kɯ, PADAM (=ABOR) a-ke, GALLONG ak-kə, NISHI (=DAFLA) ax, NYISU a-kr (with the latter two forms showing monosyllabification via apocope of the final vowel).

Several other AMD forms, however, have unexplained final -ŋ: MISING (=MIRI) a-kəŋ, MINYONG ak(k)eng. These are paralleled by a few other forms from languages in adjacent areas of Tibet and Bhutan: MONPA (Motuo) k'ung, (Central) khung; TSANGLA/SHARCHOP khuwoong/khong. It is conceivable that all of these are ultimately to be derived from a nasal-finalled allofam *krəŋ.

4.213　Forms that reflect both a dental and a velar element

Just as with the doubly-prefixed EIGHT (*b-r-gyat ≍ *b-g-ryat), where there is evidence for both orderings of the prefixes in different branches of the

---

224) Kom Rem and Puiron have a velar prefix with other high numerals as well: Kom Rem and Puiron karet '8'; Puiron kakwa '9' (see below 5.442, 5.443).

TB family, so is there with SIX:

(a)   *d-k-ruk >   DENG DARANG $tə^{41}x^{1}o^{54}$
                     TARAON   ta:hro
                     IDU   tarho
                     MISHMI   tiaro

In this category also belong LOLOISH forms with velar initials that reflect Proto-Loloish *LOW-stopped tone (e.g. LAHU khɔʔ). This proto-tone implies a voiced prefix (which in this case we assume to be *d-) at an even earlier stage. See the discussion in Matisoff 1972a:14-15 ("TSR"), and the LB forms cited in TSR #35:

WB khrok; LAHU khɔʔ; AKHA ko˰; AHI $tṣhuʔ^{44}$; SANI $khuʔ^{22}$; HANI [GAO HUANIAN 1955] khu^{21}; HANI [HU and DAI 1964] ku^{21}; LISU [FRASER] $hchaw^{6}$; LUQUAN [MA XUELIANG 1949] $tš'u^{55}$; NASU [GAO HUA-NIAN 1958] $tšuʔ^{44}$; MOSO $tṣ'wa^{55}$

(b)   *k-d-ruk

The rGYARONG dialects reflect a double prefixation in the reverse order, *k-d-ruk, with the velar being of demonstrably more recent origin. (All the rGYARONG numerals from 2 to 9 have the velar prefix ke-: below 5.2.) rGYARONG forms for SIX include: keʈa (Zida dialect); katruk, truk, keʈo, ki-trog, kə-tshuo, koco, ktru, ku-tok (cited in Nagano 1984); kətṣok (ZMYYC).

From the limited data available, it looks as if some languages of the QIANGIC group also reflect doubly prefixed prototypes. Most languages of the group have non-committal affricates (ERSU $tṣhu^{55}$, PUMI (Taoba) $tṣhu^{35}$, Muya $tçhyi^{35}$, Queyu tṣhō, Shixing $tçho^{55}$) but Qiang (Mawo) xtṣə, (Taoping) $xtṣu^{33}$ [ZMYYC #916] seem clearly to point to a complex proto-consonant group where the first element was a velar, *k-d-ruk.[225]

GARO [MOMIN] gedok is another reprefixed form, paralleled by gesa '1', gegni '2', gesni '7'.

4.214   Forms with initial resonant, with no overt sign of a prefix

In this category belongs Chinese itself, with the Old Chinese form reconstructed as *li̯ôk in GSR #1032 (see above 1.26).

TB languages that also reflect the bare root *ruk or *rok are scattered around the family:

(Himalayish)   KAIKE ru; KHALING ra:; THULUNG RAI ru
(Northern Naga)   CHANG lak; KONYAK wok; PHOM vok
(Abor-Miri-Dafla)   AKA [LSI] rieh; CHULIKATA ahe;
                    SULONG [ZMYYC #916] $yək^{33}$

---

225) A couple of other Qiangic languages have simple velar stop initials (Namuyi qhu^{33}, Guiqiong $khɔ^{33}$).

(unclassified)　　　TUJIA **wo²¹**

We may here include Northern Naga forms with a vocalic prefix: NOCTE **irok** (part of an **i-** run from 6 to 10); WANCHO **arok** (part of an **a-** run from 5 to 9).

In several Chin languages the word for SIX begins with a voiced velar stop:

PAITE, TIDDIM, VAIPHEI　　**guk**

GANGTE, KUKI, THADO　　　**gup**²²⁶⁾

At first glance it might look as if these are forms where the velar prefix has preempted the root-initial (*k-[r]uk), but in fact **g-** is the regular reflex of *r- in these languages,²²⁷⁾ so that these forms may also be referred back to the simple unprefixed allofam *ruk.

4.215　Naga forms with sibilant prefix

Several Naga languages reflect secondary prefixal *s-:

*s-ruk > ANGAMI **suru, sorou**; CHAKHESANG **shührüh**; CHOKRI **shwürü**; KEZHAMA **sarü**; KHOIRAO and MARAM **saruk**; ZEME **seruk**

Also, with *s- > h-: MZIEME and ZELIANG **heruk**.

4.216　Miscellaneous forms with labial initials

ERGONG (Qiangic group) **wtchau** shows an unusual labial prefix with this root. This is not to be compared with Chin forms like HMAR and LUSHAI **paruk**, languages where the **pa-** prefix has been generalized with all the numerals (below 5.43).

A few other languages have forms with initial f-, e.g. YI (Xide) **fu⁵⁵**, BAI (ZMYYC) **fv⁴⁴**, but, as is usually the case in TB, this labiodental consonant is a secondary development from an earlier sequence of consonant + resonant.²²⁸⁾

4.217　Isolates

MILANG **sap** is quite enigmatic. To it we may perhaps compare MIJU **katam**/DENG GEMAN **kɯ²¹təm⁵³**. The stop/nasal interchange is paralleled in EIGHT (MIJU **grin**; see below 4.236). This would imply that the final nasal is secondary, as it is in EIGHT. On the other hand, maybe these forms are somehow multiplicatively related to THREE *-sum ≍ *-sam ("3 × 2"), like a well-established group of forms for EIGHT that derive from "4 × 2" (below 4.237). In this case the final nasal would be primary, and the stop in Milang secondary.

The obscure SERDUKPEN language of northern Arunachal Pradesh has a strange sequence of numerals from FIVE to SEVEN:

khu '5' / **khit** '6' / **sit** '7'.

---

226) Apparently with assimilation of the final stop to the roundedness of the vowel.
227) As convincingly demonstrated in Solnit 1979. See also below 4.2212.
228) LAHU f-, for example, comes from PLB *hw- or *ʔw-. See, e.g. Matisoff 1979.

It looks as if the rhyme of SIX has been influenced by SEVEN, while the initial of FIVE has been influenced by the initial of SIX — i.e. the rhyme of '7' is -it "by right", just as the initial of SIX is a velar stop "by right". It is as if SIX, having given up some of its autonomy to SEVEN, then turned around and proceeded to take revenge on the next lower numeral in the pecking order, FIVE. (See above 4.01.)

4.22   Profile of number SEVEN

$$\text{SEVEN PTB *s-nis}^{229)} \text{ / PLB *s(n)i-t [TSR #128]}$$

As we have already pointed out several times,[230] the similarity of this reconstruction to that for TWO, *g-nis, has led to the deduction that the TB numeral system must once have had a quinary basis (STC, pp. 16, 93), so that SEVEN was expressed as "5 + 2". This seems very reasonable — but nobody has been able to identify any part of the proto-form *s-nis as meaning FIVE. The prefix *s- is of no help in this connection, since the TB root for FIVE either took labial (b-, m-) or lateral (l-) prefixes, not s-.

WB has hnac '2', khu'-hnac '7', with the morpheme khu' meaning something like 'unit; individual thing.' I have suggested elsewhere that it may be related to a TB root *k(r)ut 'HAND', the connection being via the five fingers used in counting.[231]

4.221   Forms overtly reflecting the *s- prefix, with retention of the root-initial

STC lists KANAWARI stis[232] (with denasalization of initial), rGYARONG kĕnĕs ~ kĕsnĕs, GARO sni, and JINGPHO sənit. To these we may add:

DULUNG (= TRUNG) suɯ²¹ nyit⁵⁵

BORO shni; DIMASA sini; KOKBOROK chini

NEWARI nhae-gu:

WRITTEN BURMESE khu'-hnac (alongside hnac '2'); ATSI nʔyit; MARU nʔat (the glottalized vowels in Atsi and Maru reflect *s-)

KEZHAMA sinyi; KHOIRAO sini; MARAM sina; MUKLOM TANGSA sanat; TANGKHUL shini; ZELIANG sinna; ZEME sena; MZIEME hena (part of a run of he- from 6-9)

QIANG (Mawo) stə (with denasalization of the root-initial); PUMI (Taoba) n̄ī³⁵; ERGONG snie/sɲie; QUEYU n̄ã⁵⁵

---

229) See STC #5, and pp. 16, 79, 93-94, 130, 131, 147, 162, 168, 169, 185, 186.
230) E.g., above 4.02, 4.11, 4.14, 4.20.
231) See Matisoff 1985a: 432.
232) Given as stish in Joshi 1909:2-3.

4.2211 With development of prefixal *s- to a dental stop or affricate

ANGAMI **thena, thenie**; AO (Chungli) **tenet**, (Mongsen) **teni**; CHAKHESANG **thena**; CHOKRI **thüna**; YIMCHUNGRU **thünie**; YACHAM-TENGSA **thanyet**; SANGTAM **thünye**

LIANGMAI **chania**; MAO **chani**; NRUANGHMEI **cünei**; RENGMA **tsanü**; SEMA **tsini**

4.2212 With rhotacism of the nasal root-initial

Many Kuki-Chin-Naga languages have forms with prefixal s- (or one of its reflexes, t- or th-), but instead of a nasal root-initial they have r-, gh- (i.e. a voiced velar fricative) or g-:

KUKI-CHIN

GANGTE **sagih**; HMAR **pasari**; KOM REM **sari**; KUKI **sagi**; LAKHER **sari**; LUSHAI **pasarih**; PAITE **sagih**; PUIRON **sari**; THADO **sAgi**; TIDDIM **səgiʔ**; VAIPHEI **sagi**

MANIPUR/NAGA

MEITHEI **taret** (cf. **taruk** '6'); MELURI **terü** (cf. **taro** '6'); NTENYI **tüghü** (cf. **togho, tüo** '6'); POCHURY **türü** (cf. **toro** '6').

The STC regards these forms as reflecting a quite separate root from *s-nis,[233] but in my opinion they cannot be ostracized from this etymon. I believe these forms merely show "rhotacism" — i.e. a "liquefaction" of the nasal. We have already noted the regular development of *r to g in many KUKI-CHIN languages (above 4.214). It seems clear that after some of these languages underwent rhotacism of the intervocalic nasal, the resulting liquid was then hardened to a voiced velar fricative or stop: *-n- > -r- > -g-.

In the case of MEITHEI and some NAGA languages, the rhotacism was undoubtedly favored by the next lower numeral SIX, which has an -r- "by right". Note that the same dental prefix occurs in '6' and '7' in these languages, forming what we might call "prefix-cum-root-initial runs" of SIX and SEVEN.

4.223 Forms with no overt trace of a prefix: *nis

MONPA (Cuona) **nis**[55]; LIMBU **nuu-si**[234]

GURUNG **ngiq**; TAMANG **nyis**; THAKALI **ngis**[235]; KAIKE **ne**

KAREN[236] (Pa-O) **nə́t**, (Pwo) **nwè**, (Palaychi) **nwìq**, (Sgaw) **nwí**

---

233) See STC p. 94, lines 1-2. The only KCN form cited there is Lushai sari, but Benedict's claim would have to apply to all the cognates that we offer here. Solnit [1979:114] follows Benedict in calling all the KCN forms "apparently unrelated to TB *s-nis".

234) Limbu -si is suffixal, occurring with most of the other numerals ('3' sum-si, '4' lii-si, '5' n(g)aa-si, '6' tuk-si, '8' phang-si). See Gvozdanović 1985:162.

235) Cited as "nigs" in CSDPN, an obvious typo.

236) Benedict [1979:13] sets up Proto-Karen *hnəs, and seems to be claiming that the final *-s (reflected by Pa-O -t) is to be considered part of the root — even though with the numerals ↗

KONYAK **nyit**; PHOM and CHANG **nyet**

GUIQIONG **ɲĩ⁵⁵**; MUYA **ɲyi³⁵**

In this category we may also include forms where the sibilant prefix has been replaced by a vocalic one:

MARING **ani**; NOCTE **ingit** (part of a run of **i-** from 6-10);

WANCHO **anat** (part of a run of **a-** from 5-9)

There is a strong tendency to palatalize the nasal before the following high vowel in this root (cf. the KONYAK, PHOM, CHANG, GUIQIONG, and MUYA forms just cited), and this development was carried to an extreme in THULUNG RAI **yet**, where the nasal feature of the initial has disappeared entirely after palatalization.[237]

4.224   Preemption of the nasal initial by the sibilant prefix: *s-[n]i-s/t

(a) *LOLOISH*

Preemption of the initial by the prefix is the rule for this root in Loloish (see TSR #128). A couple of languages reflect a stopped prototype *šit:

AKHA **sjiq/shi˯**; Hani [Gao Huanian 1955] **sɿ²¹**.

Most Loloish languages, however, have forms pointing to an open syllable under PLB Tone *2, *si²:

LAHU **šī**; LISU [Fraser] **shi⁵**, [Jui 1948] **sɿ¹¹**; LUQUAN **ši⁵⁵**,[238] etc.

These latter forms are strikingly parallel to the Loloish words for TWO that also descend from open syllables under Tone *2 (above 4.11), providing still another bit of evidence that the etyma for TWO and SEVEN are historically related. The major difference in treatment of these numerals in Loloish is that preemption of the initial never occurred with TWO, but always did with SEVEN.

(b) *SERDUKPEN*

In this little-known and lexically aberrant language of Arunachal Pradesh, which so far has not been shown to be closely related to any other TB group, the word for SEVEN is **sit** (< *s-[n]it), a classic case of prefixal preemption. As we have noted, this numeral influenced the final of the next lower numeral, **khit** '6' (above 4.01c).

(c) *QIANGIC*

Qiangic has treated the initial consonant sequence *s-n- of this etymon in a

---

∖ FOUR, FIVE, and NINE it is a *suffix* (p. 19; see above 1.21). This apparent contradiction is resolved by assuming that in 'SEVEN' the suffix had already become "welded" to the root at the PST level (p. 20). For Loloish forms which support the suffixal nature of the *-s in SEVEN, see below 4.225.

237) This form also shows convergence in rhyme with the next higher numeral, THULUNG **let** 'eight'.

238) Contra TSR #128, the LUQUAN form does not come from a stopped syllable; if it did, the ⁵⁵ tone would have constriction. The same is true of the LUQUAN form for TWO (contra TSR #160), cited above 4.11.

variety of ways (summarized below 5.0), with several languages showing preemption of the nasal root-initial by the prefix:

NAMUYI ʂɿ³³; SHIXING ʂɛ̃⁵⁵; PUMI (Qinghua) xiɛ̃¹³.

In SHIXING and PUMI (Qinghua), the nasal initial has been preserved by becoming "prosodized" or "suprasegmentalized" in the shape of vowel nasalization. The ERSU form ʃɿ⁵⁵ n̩⁵⁵ shows an alternative evolutionary strategy: disyllabization via vocalization of the prefix and preservation of the former root initial in the shape of a syllabic nasal. A further step was taken in QIANG (Taoping), where the word for SEVEN is a new monosyllable, ɕiŋ³³, such that the original prefix is now the root-initial, and the original root-initial is now the syllable-final consonant![239]

(d) *SHARCHOP and MONPA*

These two Himalayish languages have peculiar-looking forms for SEVEN that certainly appear related to each other:

SHARCHOP (=TSANGLA) **zon**;[240] MONPA (Motuo) **zum**

There are at least two etymological possibilities here. Either these forms are related to WT **bdun** (below 4.228); or else they are similar to the QIANG forms just discussed, so that the **z-** reflects the old sibilant prefix and the final nasal **-n/-m** represents the old root-initial.

(e) *With metathesis of the sibilant prefix and the nasal initial?*

It is not clear how to interpret a group of Northern Naga forms with labial nasal prefixes and sibilant root-initial:

KIMSING **mishi**; TANGSA (Moshang) **mashi**, (Yogli) **mishi**

Could these forms have arisen from a metathesis of the sibilant prefix with the old root-initial nasal? The obvious objection to this analysis is that there is no reason for the root-initial *n- to have changed its point of articulation to m- after the metathesis. It is thus probably better to consider these forms as resulting from reprefixation after preemption, i.e. *s-n- > *s- > *m-s-.

4.225 *AMD forms for SEVEN with a velar prefix*

A number of AMD languages have developed a velar prefix with this etymon:[241]

ABOR-MIRI **ki-nit**; PADAM-MISING [Tabu Taid] **kɯnɯt**; LHOPA (=BOKAR) **kɯnɯ**; APATANI **kanu**; BENGNI **ka-ni**; MINYONG **kenit**; DAFLA [Das Gupta], GALLONG, PADAM, TAGIN **kane**; NISHI [Dubey] **ken**[242]

---

239) Monosyllabization of disyllabic compounds is a strong tendency in Qiangic. See Benedict 1983 and Matisoff 1991:493. Cf. similarly preempted and apocopated forms for FIVE (above 4.1411), the Nishi forms for '7' and '8' (below 4.225), and the general discussion (below 5.131).
240) See also BUMTHANG **zon** '2', and "Transvaluation of numerals", above 4.02.
241) J.T. Sun reconstructs Proto-Tani *kV-nɯt [1993:213].
242) This monosyllabic NISHI form is another illustration of the process of "prefix preemption ↗

One would expect some problems with appending a velar prefix to the root for SEVEN, since that would make it look even closer to the general etymon for TWO, *g-nis. Yet all these AMD languages manage to keep the two numerals distinct, by one means or another: either by introducing a new vocalic prefix for TWO (e.g. ABOR-MIRI, DAFLA, GALLONG, MINYONG, TAGIN a-nyi), or by innovating a new root for SEVEN (e.g. CHULIKATA joh, TARAON wẽ, GEMAN DENG nɯɯn⁵³, IDU ĩũ, MIJI myah, MILANG rangal; see below).

4.226   Forms with a double prefix: velar plus sibilant: *g-s-ni-s
A couple of languages have innovated a secondary velar prefix while retaining the older sibilant one:

rGYARONG (Zida dialect) keʂnyit ~ keʂnyis ~ keshnyes.

All numerals in this dialect (except EIGHT) have developed the ke- prefix as well. The retention of the inner sibilant prefix in SEVEN serves to distinguish it from kenes '2'. See below 5.2.

GARO [MOMIN n.d.] has gesni alongside the simpler sni cited above; this optional secondary prefix also appears in gesa '1', gegni '2', gedok '6'.

4.227   Forms with nasal final consonant as well as nasal initial
Two AMD languages have forms for SEVEN with nasals in both initial and final position:

MIJU nin (alongside kinin '2')

DENG GEMAN(= KAMAN) nɯɯn⁵⁴ (no parallelism with kɯ²¹jin⁵³ '2') The most plausible explanation of these forms is that the final consonant has assimilated to the root-initial nasal. Both of these languages show convergence of the rhyme in SEVEN and EIGHT (MIJU grin '8', DENG GEMAN grɯɯn⁵³ '8'), even though the latter is to be reconstructed with PTB *-t.[243]

A more far-fetched hypothesis would be to try to relate these forms to WT bdun (next section).

4.228   Written Tibetan bdun and its possible congeners
WT bdun '7' has always been something of a mystery, hitherto thought to be an isolate in TB, with cognates to be found only within "Bodish": i.e. Tibetan dialects and a few other closely related Himalayish languages:

LHASA tüün; SHERPA din; JIREL duin; SIKKIM BHUTIA (=DZONGKHA) dŭin (the vowel in Chhewang Rinzin's speech is ɪ);

---

with apocope of the root-vowel", so that the former root-initial has become the new final consonant. See also NISHI pin '8' (alongside, e.g. MINYONG pini, below 4.237); also such forms for FIVE as phung (above 4.14), and some Qiangic words for SEVEN, above 4.224(c). See below 5.131.

243)   As we have seen (above 4.224), THULUNG has also analogically levelled the rhymes of these two numerals, but in favor of a final stop rather than a nasal: yet '7', let '8'.

BAIMA **de**[13] (p.c. Sun Hongkai 1991); also probably KHALING **taːer**. While most dialects of rGYARONG have doubly-prefixed versions of the normal root *s-nis (above 4.226), the Hanniu dialect has **daen**,[244] which goes well with these Bodish forms (especially with KHALING), and seems to be an obvious loan from Bodish into rGyarong.

MONPA (Motuo) **zum** and SHARCHOP **zon** have some resemblance to these forms, but can equally well be considered monosyllabicized versions of the normal root *s-nis (above 4.224d).

Likewise, Kaman **nɯn**[53] and Miju **nin**, despite their own superficial similarity to **bdun**, are best considered to have arisen from the "normal" root by assimilation of the final consonant to the nasal initial (above 4.227).

That leaves as the only conceivable candidates for cognacy with WT **bdun** the following AMD forms:

Taraon	**wẽ**
Deng Darang	**weng**[54]
Idu [TALUKDAR 1962]	**ĩũ** (alongside **inyũ** '8')
Idu [SUN 1983]	**i**[55]**ɦiong**[55] (alongside **i**[55]**lioŋ**[35] '8')
Mishmi	**iuo**
Chulikata [LSI]	**joh**

All of these are perhaps to be referred back to PTB *b-dun, the initial *b- being interpreted as a prefix. The etymon could somehow have developed a secondary palatalization to *b-d(y)un (cf. the Lhasa form with front rounded vowel), which could have weakened the preceding -d- to the point where it was preemptible by the prefix (> *b-[d]yun). The b- could then itself have weakened to w-, or dropped entirely, leaving y- or a high front vowel as the syllable-initial.[245]

4.229 An additive form in MIKIR

In MIKIR, '7' is formed additively on the basis of '6':

throk '6' + **isi** '1' = **throk-si** '7'.

As far as I know, this is the only such case in Tibeto-Burman.[246]

Since MIKIR also expresses EIGHT and NINE subtractively on the basis of TEN (below 4.23, 4.24), it has no monomorphemic inherited numerals between SIX and TEN.

4.22.10 Isolates?

There remain a few forms for SEVEN which cannot be related to anything else in the current state of our knowledge:

---

244) Data from Nagano 1984.
245) For a somewhat analogous sound-change, cf. WT **dbus** 'head; central' > Lhasa **üü**.
246) See above 4.203.

(A)   MILANG **rangal**.
      Could there by some contamination here from FIVE ***l-ŋa** ?
(B)   AKA (=HRUSO) [LSI] **mulh**.
      This slightly resembles the last syllable of the MILANG form.
(C)   DHAMMAI **mjaʔ**; MIJI [SIMON n.d.] **myah**
      These closely resemble the Aka form.
(D)   LEPCHA **kə-kyək**.
      This is part of a 7-10 prefix run, also including **kə-ku** '8', **kə-kyot** '9', **kə-ti** 10'.
(E)   LOTHA **ti-ing** is especially interesting. The first syllable is part of a secondary run of the **tV-** prefix (**ti-rok** '6', **ti-ing** '7', **ti-za** '8', **to-ku** '9', **ta-ro** '10'). The second syllable looks like a prefixed and apocopated form of ***-nis** (cf. NOCTE **ingit**), but cf. also Lotha **ti-ingya** '50' (above 3.522).

4.23   Profile of number EIGHT

EIGHT ***b-r-gyat** ≍ ***b-g-ryat**[247]

Key forms for justifying this doubly prefixed reconstruction are WT **brgyad** and JINGPHO **mətsát**.[248] To these we may add rGYARONG (Zida) **warzhe(t)**, with **wa-** also < ***b-**.[249] Naturally enough, the complex consonant sequence involving a liquid was particularly prone to metathesis and preemption. The CHINESE cognate shows preemption of the rest of the initial cluster by the labial prefix: OLD CHINESE ***pwăt** [GSR #281]. It seems clear that the initial consonant group in HUNDRED (***b-r-gya** ≍ ***b-g-rya**) has been influenced by that of EIGHT, despite the fact that these two numerals are not neighbors in linear order.[250]

The busy numeral EIGHT is involved in subtractive, multiplicative, and transvaluational phenomena. In MIKIR and MEITHEI, '8' (as well as '9') is expressed subtractively in terms of TEN (above 3.232, 4.203); and in many Abor-Miri-Dafla languages it is expressed multiplicatively as "4×2" (below 4.237). In LEPCHA, '8' and '9' have undergone an etymological flipflop (above 4.02; below 4.24).

The prefixal behavior of this numeral is predictably complicated. Some languages merely reflect a ***g-** or an ***r-**, either one of which could function as the "root-initial" because of this etymon's metathetic propensities. Others

---

247) See STC #163, and pp. 35, 45, 54, 57, 74, 88, 95, 96, 131, 141, 144, 161-162, 179, 191. I have reconstructed this etymon for Proto-Lolo-Burmese as ***ʔritᴸ** (TSR #171).

248) Other examples of Jingpho **mə-** < ***b-** include **məlī** 'four' <***b-ləy**, **məŋā** '5' < ***b-ŋa**.

249) EIGHT is the only rGyarong numeral from 2-9 not to have the prefix **ke-** (see below 5.2). Undoubtedly its double prefix was weighty enough to allow it to escape the steamroller of the velar prefix run.

250) See above 3.54, and STC n. 148 (p. 45).

have a labial (or other) prefix before the velar or the **-r-**. Very often the reflexes in the daughter languages begin with a voiceless spirant or affricate, such that it is difficult to decide which elements of the complex proto-cluster might immediately underlie them. STC notes (n. 148) that "Kuki-Naga has replaced the (labial) prefix [by a dental]: PKN ***d-ryat** < ***g-ryat**, apparently under the influence of TB ***d-ruk** '6' and ***d-kəw** '9'." The STC is here groping toward the notion of "prefix run" (see below 5.2 ff.).

4.231　Kuki-Chin-Naga forms with dental or affricate prefix
LAKHER **chari**; MAO **chacha**; TANGKHUL **chishat**.

ANGAMI and CHAKHESANG **thetha**; CHOKRI **tütha**; KEZHAMA **tiche**; KIMSING **tecat**; LIANGMAI **tachat**; LOTHA **tiza**; MELURI and POCHURY **tüze**; NRUANGHMEI **tacüt**; NTENYI **tüza**; RENGMA **tükhü**; SEMA **thache**; TANGSA (MOSHANG) **tachat**, (Yogli) **tüchat**; YACHAM-TENGSA **thesep, teset**; YIMCHUNGRU **tizha**; ZELIANG **tesat**; ZEME **desat**.

The Liangmai and Nruanghmei forms break up runs of numerals with affricate prefixes:

	LIANGMAI	NRUANGHMEI
SIX	charuk	cüruk
SEVEN	chania	cünei
EIGHT	tachat	tacüt
NINE	chakiuh	cükiu

4.2311　With preemption of the root-initial by the dental prefix
KONYAK **tet** (< ***d-ryat**), alongside **tu** '9', also a preemptive form < ***d-kəw**); AO (Chungli) **ti** (but AO Mongsen has **tsit**, apparently a fusional rather than a preemptive form; cf. Jg. **mətsát**)

4.232　Forms with other prefixes
A variety of secondary prefixes have been attached to this etymon in one language or another, including **vowels, p-, k-,** and **s-,** usually as part of a prefix run affecting most or all of the higher numerals:

(a)　Vocalic prefix > NOCTE **isat**; WANCHO **achat** (below 5.512)

(b)　**p-** > HMAR **pariet**; LUSHAI **pariat**

This **pa-** is an innovative prefix that goes with all the numerals in a few Chin languages (below 5.43), and has nothing to do with the PST/PTB ***b-** posited for EIGHT in particular (which is reflected, e.g. in the labial initial of the preemptive Chinese cognate).[251]

---

251) Forms from certain BAI dialects, e.g. Jianchuan and Dali **piɑ⁴⁴**, seem clearly to be loans from Chinese. The Bijiang dialect of BAI has a form with initial affricate (below 4.235), which looks like an independent reflex of the PST/PTB etymon.

(c) k- > KHOIRAO **kachat** (not part of a prefix run); KOM REM **karet** (alongside **karuk** '6'); PUIRON **karet** (alongside **kakwa** '9').

LEPCHA has a 7-10 run of a velar prefix: **kăkyăk** '7', **kăkū** '8', **kăkyót** '9', **kăti** '10', wherein '8' and '9' seem to be reversed etymologically (above 4.02, below 4.24).

(d) s- > MARAM **sachat** (part of a 6-9 run of sV-); MZIEME **heset** (part of a 6-9 run of he-).

SERDUKPEN has a doubly-prefixed form **sargiat** (< *s-r-gyat), where the young sibilant prefix has been preposed to the older liquid one. Perhaps quite akin to this Serdukpen form is the strange AKA word given in LSI, **sikzi** (< *s-g-ryat?).

4.233   Forms with velar initials

DZONGKHA/SIKKIM BHUTIA **gye**; MONPA [DUBEY] **giet**; JIREL **gyet**; KAIKE **kye**; SHERPA **ge**
GANGTE **giet**; KUKI **get**; PAITE, TIDDIM, VAIPHEI **giat**
SANGTAM **ke** is a lone monosyllabic form that breaks up a 6-10 run of dental prefixes (**thüro** '6', **thünye** '7', **tüku** '9', **thüre** '10').
QIANG (Mawo) **kha**ʲ [ZMYYC 1291] looks like an apocopated and metathesized form, and resembles several other reflexes with final -r or a rhotacized vowel (e.g. NUSU **ṣɑ**ʲ⁵³ and KOKBOROK **char**, below 4.235).

4.234   Forms reflecting initial r- or a cluster of C + r

The Gurung-Tamang-Thakali group reflect *b-ryat, preserving both a labial and an r in this word: GURUNG **prehq**; TAMANG **preht**; THAKALI **preh**.

THULUNG RAI [AGAM SINGH RAI 1944; ALLEN 1975] **let** (prob. < *ryat); KHALING **ri**; KANAWARI **rai**; KULUNG **ret-chi**; LIMBU **jEt-tshi**
PROTO-LOLO-BURMESE *ʔrit ✕ *ʔryat > WB **hrac**; LAHU **hí**; AKHA **yeh**˰; AHI **xi**⁴⁴; SANI **he**²²; HANI [GAO HUANIAN 1955] **xae**²¹; HANI [HU and DAI 1964] **še**²¹; LISU [FRASER] **h'i**⁶; LUQUAN **ʔhən**⁵⁵; NASU [GAO HUANIAN 1958] **xɛn**³⁴; ACHANG **çet**⁵⁵; ZAIWA **ʃit**⁵⁵; MARU **ʃɛʔ**⁵⁵; ANONG **çen**⁵⁵; NAXI (Lijiang) **xo**⁵⁵, (Yongning) **xu**¹³ [see TSR #171]²⁵²⁾ JINUO **xe**⁴⁴; TUJIA **jie**²¹
The aberrant and isolated SULONG language of Arunachal Pradesh has a form with liquid initial, **la**³³ (ZMYYC 1291).

4.235   Forms with affricate/sibilant initials that could reflect either *gy- or *ry-

AO (Mongsen) **tsit**; CHANG **sat**; MARING **chot**; PHOM **šət** (alongside

---

252) Note the rhinoglottophiliac nasalization in LUQUAN, NASU, and ANONG.

pa-šet '18', shə '9'); WANCHO achat; NOCTE i-sat; TANGSA (Yogli) təchat, (Moshang) tachat [with dental prefix: above 4.231]; KONYAK tet (with preemption by the dental prefix).[253]

NEWARI *cya.-gu:*

Most Qiangic languages have sibilant spirants or affricates: ERGONG ʒyiɛ (< *r-gy-); ERSU ʒɿ⁵⁵; MUYA ɕyʁ⁵³; PUMI (Qinghua) ʂuɛ²³, (Taoba) ɕyʁ³⁵; QIANG (Taoping) tʂʰe³³; QUEYU ɕyɛ⁵⁵; SHIXING ɕyi⁵⁵. A couple of Qiangic languages have forms with initial semivowel or h-: NAMUYI hĩ³³ (with rhinoglottophilia); GUIQIONG je⁵⁵.

BAI (Bijiang) tɕuɑ⁴⁴. Other Bai dialects have apparent loans from Chinese; cf. Jianchuan and Dali piɑ⁴⁴, above 4.232.

DULUNG ɕăt⁵⁵; TRUNG šiat⁴⁴; NUSU ʂɑ˩⁵³

GARO chet; DIMASA jai; KOKBOROK char (with the final -r apparently due to metathesis: *g-ryat > *gyar > char); cf. the NUSU form just cited, as well as QIANG (Mawo) khaᴶ, above 4.233.

KARENIC: PA-O sɔ́t; PWO xɔ̀ʔ; PALAYCHI xó; SGAW xɔ̀ʔ. KAYAH (=KARENNI) has a multiplicative form for EIGHT (above 1.21, below 4.238).

4.236 A new allofam with final nasal: *g-ryan

Several AMD and geographically contiguous Himalayish languages have forms with final nasals that do not appear to be caused by rhinoglottophilia, but seem to be genuine reflections of an allofam like *g-ryan:

MIJU MISHMI grin; KAMAN (=DENG GEMAN) gᴶɯn⁵³; MILANG rayeng; MONPA (Cuona) cen²³, (Motuo) jen ["j" is palatal semivowel]; TSANGLA jen; SHARCHOP yin [CHHEWANG RINZIN 1984].

Cf. also SUNWAR yaan 'NINE', which is perhaps a transvalued ("upstepped") reflex of this allofam for EIGHT.[254]

4.237 A new AMD etymon *lyoŋ ?

Some other AMD languages have forms with lateral initials and nasal or open finals, that are apparently independent of the group in 4.236, and that we tentatively refer to a new root like *lyoŋ:

DENG DARANG liɯm³⁵; IDU [SUN 1983] i⁵⁵liong³⁵ (alongside i⁵⁵ɦioŋ⁵³ '7'); IDU [TALUKDAR 1962] inyŭ (-ny- apparently < earlier -ly-; cf. also IDU iŭ '7'); MISHMI ili; CHULIKATA [LSI] ilu:

---

253) W.T. French reconstructs PROTO-NORTHERN NAGA *C/V-gyat, with unspecified vocalic or consonantal prefix [1983:482].

254) This is all the more probable since SUNWAR gow means 'TEN', but looks like an "upstepped" reflex of NINE (*d-kəw). For another possibility, see below 4.246.

4.238   Multiplicative forms in Abor-Miri-Dafla and elsewhere: 8 = 4×2

Many AMD languages have multiplicative compound forms for EIGHT of the structure "4×2":[255]

ABOR-MIRI   a-pi '4', a-nyi '2' > pi-nyi ~ pui-nyi '8'
APATANI   a-pi '4', nyi '2' > pɯʔ-nyi ~ pryɯʔ-nyi '8'
LHOPA   apiː '4', anyi '2' > piː-nyi '8'

Similarly: DAFLA, GALLONG, PADAM, TAGIN pine, YANO plə-ne; MINYONG pini[256]; NYISU plin; NISHI pin, piin.[257]

Other multiplicative formations for EIGHT in TB include BORO zokkay nəy (above 3.32C, 4.201), and KAYAH (=KARENNI = RED KAREN) lwīswáʔ (above 1.21, 4.201).

4.239   Isolates

BORO thai-*dang*-nia 'eighth' [BIBLE SOCIETY OF INDIA 1972b] is a totally mysterious form. (The prefix thai- and suffix -nia occur with all the Boro ordinal numerals in the *Book of Revelation*. See below 4.245.)
LIMBU phang-si [GVOZDANOVIĆ 1985:162] is also a puzzlement.

4.24   Profile of number NINE
NINE   *d-kəw (=*d-kuw) ⤨ *s-gəw ⤨ *d-gaw

4.241   With dental or sibilant prefix

STC reconstructs only the prefix *d- for the PTB level, relying especially on WT dgu and Nung tegö.[258] This *d- also receives considerable support in Kuki-Naga, but this evidence is more equivocal, since in most of these languages the dental prefix in NINE is a part of a larger "prefix run",[259] involving EIGHT and often SIX, SEVEN, and/or TEN as well.

Kuki-Naga languages showing a prefix of the shape tV- or thV- for NINE include: ANGAMI theku, thepfü; AO tuku, tüku; CHAKHESANG thechi; CHOKRI thüchi; KEZHAMA tepfü; KIMSING tak(a)u; KONYAK tu (with preemption of the initial); LOTHA and SEMA toku; MARING tako; MELURI tokhu; MOSHANG takru (-r- < ?); NTENYI tükhu; POCHURY

---

255) J.T. Sun reconstructs a Proto-Tani multiplicative compound, *pri-ñi [1993:125].

256) Cf. MINYONG a-nyi '2', but a-ki '4' (< ?). Note that the inherited PTB root for FOUR, *b-ləy, survives in Minyong only in its multiplicative derivative EIGHT.

257) NISHI and NYISU show apocope of the final vowel, as also in Nishi ken '7' (cf. Padam kane).

258) See STC #13, and pp. 19, 23, 45, 61, 94-95,116, 131, 134, 154, 162, 185, 188, 196. "*d-gew" in STC's Appendix I (p. 202) is a typo for the poorly attested variant *d-gaw (see note 9). STC also recognizes a Kuki-Naga variant *d-kwa, probably reflecting a secondary suffix (< *d-kuw-a). Similar to the Nung form cited in STC are ANONG dɯ³¹gɯ³¹ and DULONG dɯ³¹gɯ⁵³ [ZMYYC].

259) See below 5.44, "Innovative runs in the higher numerals".

toku; SANGTAM **tüku**; YACHAM-TENGSA **thaku**; YIMCHUNGRU **tuku**; YOGLI **tükau**. To these we may add extra-KCN forms like MONPA [Cuona: SUN et al. 1980] **tu²¹ku⁵⁴**, [DUBEY 1983] **dugu**; and SERDUKPEN **dikhi**.

Not enough is known about the history of prefixes in KCN to be sure that all of the above reflect *d- rather than, e.g. *s-. The same uncertainty attaches to the origin of the *affricate* prefixes in KCN languages: e.g. KHOIRAO **chaku**, LAKHER **chaki**, LIANGMAI **chakiuh**, MAO **choku**, NRUANGHMEI **cükiu**, TANGKHUL **chiko**. These affricates also typically occur in "runs" in KCN, but are to be found elsewhere as well, e.g. KOKBOROK (Barish) **chuku**, JINGPHO **jəkhû**. I have already observed in a previous analysis of this etymon[260] that Jingpho **jə-** here may well come from *s-, given the fact that the Jingpho causative prefix **šə-** (< PTB *s-) undergoes a predictable morphophonemic change to **jə-** before verb roots beginning with an aspirated consonant (as in NINE) or a sibilant. Perhaps there was a tendency for Jg. *s- to become an affricate in non-causative contexts as well.

There is in fact considerable evidence for according *s- just as ancient an association with NINE as that enjoyed by *d-. Two of the forms cited in STC #9 as evidence for *d- point more straightforwardly to a sibilant prefix: KANAURI (= KANAWARI) **zgui** and GARO **sku** (also Dimasa **sugu**). To these we may add forms from the Qiangic group: PUMI [LU SHAOZUN 1983] **sgiɯ⁵⁵** and QIANG (Taoping) **xguə³³**.[261] None of these sibilant-prefixed words for NINE is participating in a "prefix run" — the neighboring numerals lack such a prefix. We should also mention AKA (= HRUSO: data from LSI) **stheu, sthö '9'**. This puzzling form may reflect a doubly-prefixed prototype *s-d-[k]əw which underwent preemption of the root-initial velar.

These cases are to be sharply distinguished from those where NINE has a sibilant prefix shared by the neighboring numerals, as in MARAM **soki '9'** (but also **saruk '6', sina '7', sachat '8'**); or ZEME **sekui '9'**, but also **seruk '6', sena '7'** ('8' is **desat**). (The closely related MZIEME has **he-** from SIX to NINE.)

Some languages have forms with sibilant or affricate initials which require explanation:

> PHOM **šə** (French [1983:527] refers this to PROTO-NORTHERN NAGA *C/V-gə:w, along with other forms with dental, vocalic, or zero-prefix.)
> HANI (Caiyuan; Biyue) **tsi³¹** (This is the only form with an affricate initial in Loloish proper, but cf. also JINUO **tçy³³**.)
> BAI (Dali and Jianchuan) **tçɯ³³**, Bijiang **tçi³³**

4.242　With velar prefix

A secondary velar prefix is occasionally found with NINE: PUIRON

---

260) Matisoff 1980 ("Stars, moon, and spirits..."), pp. 15-17.
261) QIANG (Mawo; ZMYYC 1292) has an unusual prefixal **r-** (**rguə**), which is probably velar in articulation like Taoping **x-**, and could descend from *s- as well.

kakwa (alongside **karet** '8'); LEPCHA **kəkyót** (part of a velar "run" from SEVEN to TEN; the Lepcha words for EIGHT and NINE seem to have undergone an etymological flipflop; above 4.02); and rGYARONG **kəngu** (all rGYARONG numerals from 2-7 also have a velar prefix).

4.243   With prenasalized initial

Several Qiangic languages have forms with prenasalized initials (ERGONG **ngiɛ**; MUYA **ŋguɯ³⁵**; ERSU **ngɛ³³**; NAMUYI **ŋgu³³**), as does the Lijiang dialect of NAXI (**ŋgv³³**).

4.244   With no overt trace of a consonantal prefix

NORTHERN NAGA

A few languages in this group either have no prefix (CHANG **guh**), or a vocalic one (NOCTE **i-khu**, WANCHO **a-ku**).

QIANGIC and HIMALAYISH

Several Qiangic and Himalayish languages show no trace of a prefix with this etymon:

GUIQIONG **gui³³**; QUEYU **gɯ⁵⁵**; SHIXING **guɐ³³**; THULUNG **gu**; KHALING **gɦiu**

Neither Lolo-Burmese nor Karenic show any evidence of a prefix:

LOLO-BURMESE

PLB **gəw²* > WB **kûi**; LAHU **qɔ̂**; AKHA **γɔ̀**; HANI (Shuikui) **yu³¹**; LISU **ku⁵⁵**; NAXI (Yongning) **gv³³**; ACHANG **kau³¹**; ZAIWA **kau²¹**; LANGSU (=MARU) **kuk³¹** (the secondary **-k** is regular for the rhyme **-əw*); NUSU **gɯ³⁵**; TUJIA **kɯe⁵⁵**

KARENIC

PHO (Moulmein), PALAYCHI, SGAW **khwí**; Pa-O **kút** (with suffixal **-k**: see above 1.21)

4.245   A new root for NINE in Abor-Miri-Dafla and elsewhere:
**k-n (y/w) a-ŋ*

There is a newly discovered root for NINE in AMD, with possible Barish and Karen cognates. It seems to have a velar prefix, a nasal root-initial, a semivowel (**y** or **w**), and sometimes a final nasal as well. We may reconstruct it roughly as **k-n (y) a-N* or **k-n (y/w) a-ŋ*. The vocalism of the prefix fluctuates greatly, which we can symbolize by setting up a dummy vowel (**kV-n (y) a-ng*).[262] Reconstructing a prefix consisting only of a consonant is tantamount to saying that any vowel that intervenes between that prefix and the root-initial is not distinctive — being unstressed, it is too prone to influence from the vowel

---

262) J.T. Sun has just independently reconstructed this etymon as Proto-Tani **kV-(n)aŋ* [1993:186].

of the root or from anything else.

(a)　With **-o-** vocalism in the minor syllable:

LHOPA	**konong**	(? < ***k-nwaŋ**)
MINYONG	**konang**	
ABOR-MIRI	*ko-nang*-**ko**	
GALLONG	**kona**	
PADAM	**kona**	

(b)　With **-i-** vocalism in the minor syllable:

TARAON　　　　　**kinyaːng**
IDU　　　　　　　**kinyi**
CHULIKATA　　　**khili** (with lateral rather than nasal root-initial)
APATANI　　　　　**kíwa** (< ***k-[n]wa-N**, with loss of nasal root-initial)

(c)　With **-a-** or **-ə-** vocalism in the minor syllable:

DENG DARANG　**kə²¹ɲɯng⁵⁵**
MILANG　　　　　**kanyem** (with labial rather than velar final)

(d)　With **-e-** vocalism in the minor syllable:

NISHI　　　　　　**keya** (with loss of nasal element)
TAGIN　　　　　　**kéya** (ditto)
DAFLA　　　　　　**kéya** (ditto)
GALLONG　　　　　**kenga** (with velar rather than palatal nasal)

To this group of forms also belong NYISU **kja:** and BENGNI **kju-a:**.

(e)　With no velar-initialled minor syllable:

MISHMI　　　　　**a-niu-ma**
KAMAN/　　　　　**nən⁵⁵ mu⁵³** (with dental rather than velar final)
　DENG GEMAN
MIJU　　　　　　　**nat-mo** (with final stop homorganic to the Kaman nasal)

As a long shot, we may perhaps relate this new root to a couple of isolated forms elsewhere in TB:

BORO [BIBLE SOCIETY OF INDIA 1972b] **thai-*ne*-nia** 'ninth'.[263]

W. KAYAH (Karenic) **nɯə'** [264]

SUNWAR **yaan** '9' does not seem to belong with this etymon, despite a certain phonological similarity to some of its reflexes.　It is more likely to be a transvalued reflex of ***g-ryan** '8' (above 4.236).

---

263)　Cf. **thai-*daŋ*-nya** 'eighth', above 4.238.　Unlike 'eighth' and 'ninth', the Boro ordinal numerals 'sixth' (**thai-*do*-nia**) and 'seventh' (**thai-*shni*-nia**) faithfully preserve the general TB roots for SIX and SEVEN that have been lost in the usual quaternary system of cardinal numbers (above 1.24).

264)　The apostrophe marks a Kayah high tone which is the reflex of Proto-Karen *D-1, a tone occurring in syllables with former final stop.　This seems directly cognate to the Miju form in -t.　This dialect of Kayah has another form for '9', **da** (Tone *B-1), whose etymology remains obscure (data from D. Solnit).

4.246   Subtractive forms

In two important languages the word for NINE is formed subtractively on the basis of TEN: MEITHEI **ma-pan** and MIKIR **sirkep** (above 4.20).

5.   PREFIXAL BEHAVIOR WITH NUMERALS

5.1   Prefixal Variability and Replaceability

We have seen many examples of completely different prefixes being attached to the same etymon in one language or another (e.g. NINE *d-gəw ⋊ *s-kəw). Yet for a given etymon, there is high variability even in the treatment of the *same* prefix between closely related languages, or dialects of the same language. As a random example we may take some QIANGIC forms for SEVEN, all descending from the general root *s-nis (see above 4.22). Some languages preserve an overt trace of the prefix (Qiang [Mawo] **stə**, Pumi [Taoba] **ɲĩ³⁵**, Ergong **snie/sɲie**, Queyu **ŋã⁵⁵**); others have lost the prefix entirely (Guixiong **n̠ĩ⁵⁵**; Muya **n̠yi³⁵**); while still others show preemption by the prefix of the root-initial (Namuyi **ʂɿ³³**, Shixing **ʂɛ̃⁵⁵**, Qiang [Taoping] **ɕiŋ³³**, Pumi [Qinghua] **xiɛ̃¹³**).

Does it make sense to speak of the "repertoire of prefixes" that a given numeral has been observed to develop somewhere or other in ST? It is actually not too useful simply to list all the prefixes that have been attested for a given numeral, because of the phenomenon of "prefix runs", whereby consecutive numerals (and sometimes even *all* the numerals) acquire the same prefix by a kind of assimilation.

We must factor out obviously late assimilatory developments, but there is no way we can claim that prefixal variation was absent even at the PTB level. The *Conspectus* has done a good job of identifying the most widespread prefixes attested for each numeral, but it does not go far enough in acknowledging that more than one prefix may be of ancient standing with any given numeral, or that in some cases it is impossible to say which of several prefixes is "older".

Can we establish the relative age of the ensemble of prefixes used with any given numeral? There are certain arbitrary aspects of the treatment in STC. For some numerals prefixal variation is posited at the proto-level (FIVE, EIGHT); the implication is that all other prefixes that pop up in daughter languages are secondary. Sometimes this *is* clearly the case, the limiting situation being that of e.g. LUSHAI, where a single prefix has been generalized for all the numerals. Similarly, in cases of reprefixation, the more outer prefix is clearly younger than the more inner one (below 5.2). In other cases (SIX, NINE, maybe FOUR), however, there seems no reason not to posit prefixal variation as far back as one can go. The intrinsic variability of prefixes

militates against a too rigid view of setting some up as "proto-" and all others as secondary.

5.11 Voicing and vocalization of prefixes

There is no evidence for a voicing contrast in stop prefixes at the PTB level. The STC conventionally reconstructs *b-, *d-, *g-, but "archiphonemic" symbols like *B, *D, *G would do just as well. We cannot usually put so fine a phonetic point upon proto-prefixal matters.

In PROTO-LOLO-BURMESE, tonal reflexes force us to distinguish between *g- (a putative subtype of the "C-" prefix) and *k-, when they occurred before resonantal initials.[265] Yet this distinction is shaky, since there is no direct evidence for a distinctively *velar* voiced stop prefix in PLB. "C-" is a very vague proto-entity! JINGPHO does have such a voicing contrast synchronically, e.g. kə- vs. gə-, but there is much variability here. LaRaw Maran [in prep.] and others [DAI et al. 1981; LON DIEHL p.c.] claim a tonal difference in minor syllables of this type, according to the voicing of the prefix. In some of our sources, synchronic variation in the voicing of a prefix is explicitly reported, e.g. Central MONPA b(i)ci ∼ p(i)ci '4' [DAS GUPTA 1968].

5.111 The vowels of prefixal syllables

Some languages show considerable fluctuation in the vocalization and/or aspiration of the same prefix from numeral to numeral. Thus, SEMA kini '2', but küthu '3'; tsogho '6', but tsini '7'; thache '8', but toku '9'. Cross-linguistically, the same prefix may be differently vocalized when attached to the same numeral (cf. the fluctuation of the vowel from language to language in the new AMD root for NINE, above 4.245).

5.12 Contamination by prefixes of non-contiguous numerals

As STC observes (n. 148, p. 45), the prefix of HUNDRED has been altered in many TB languages to bring it into line with that of a more basic numeral with which it already shared a high degree of phonetic resemblance, i.e. EIGHT. (See above 3.54; 4.23.)

5.13 Prefix preemption of the root-initial of a numeral

Particularly apt to preempt are the *b- in FOUR, the velar (*g- or *k-) or dental (*d-) in SIX, and the sibilant *s- in SEVEN (above 4.13, 4.21, 4.224). This preemptability is due to the "weakness" of the root-initial consonant (*-l- in FOUR, *-r- in SIX, *-n- in SEVEN). Preemption can be the road to

---

265) In TSR [MATISOFF 1972a], "C" is used as a cover symbol to stand for a *voiced prefix that caused its syllable to belong to the LOW category of stopped syllables, even if the following root initial was *voiceless. Conversely, the *k- prefix had the power to shift a syllable with a voiced resonantal root-initial into the HIGH stopped class.

survival for a prefix, most strikingly in isolated forms in subgroups where prefixes do not generally persist, e.g. the labial prefix in FOUR in the Burmish language known as MARU or LANGSU (MARU **bìt**, LANGSU **pjik**[31] [ZMYYC 1287]). With respect to EIGHT, STC recognizes metathesis even for the proto-level: ***b-r-gyat** ⋊ ***b-g-ryat**. Either prefix could (and often did) preempt the root initial. With respect to SIX, what one calls "preemption" depends of course on what one takes the proto-form to be. Which prefix is "more inner"? Is it ***d-k-rok** or ***k-d-rok**? How can STC be more sure of this than of the order of the double prefix in EIGHT?

5.131   *Preemption via apocope of the root vowel*

In forms like AKA (AMD) **phum** '5' ($<$ ***b-ŋa**), the root-vowel **-a** had disappeared, so that the former root-initial consonant **ŋ-** now appears in syllable-final position, becoming a labial (**-m**) by assimilation to the original prefix. This erstwhile prefix must now step in to discharge the duties of the root-initial, while the unstressed vowel of the old prefix (which had presumably been just schwa) gets restressed, assuming a rounded quality due to its doubly labial environment. This can only be described as a radical reorganization of the functional parts of the proto-syllable:

	PREFIX	PREFIXAL VOWEL	ROOT INITIAL	ROOT VOWEL	FINAL CONSONANT
PTB	b	ə	ŋ	a	zero
AKA	ph	u	m		
	ROOT INITIAL	ROOT VOWEL	FINAL CONSONANT		

We have found a number of monosyllabic forms of this apocopated/preemptive type, including: PUIRON **pang** '5', KHALING **bho:m** '5', BORO and KOKBOROK **ba** '5' (all $<$ ***b-ŋa**); BORO **do** '6' ($<$ ***d-ruk**; all cited above 4.1411); CHULIKATA **kāsh** '3' ($<$ ***g-sum**; below 5.511); NISHI **ken** '7' ($<$ ***k-nit**; above 4.225); NISHI **ax** and NYISU **a-kr** '6' ($<$ PTani ***a-krə**; above 4.212); QIANG (Taoping) **çiŋ³³** and Ersu **ʃi⁵⁵ ṇ⁵⁵** '7' ($<$ ***s-nis**; above 4.224c).

This process applies equally well to dissyllabic compounds, where each syllable once had an independent meaning. Here the initial consonant of the first element in the compound becomes the initial of the new monosyllable:

   QIANG (Sanlong) **han** '12' ($<$ **ha** '10' + **ni:ʔ** '2'; p.c. J.P. Evans); NYISU **plin** and NISHI **pin** '8' ($<$ ***pri-ñi** "4 × 2"; above 4.237); HILL MIRI **čem-piŋ** '80' ($<$ ***čam-p(r)i-ñi**).

5.2   "Prefix Runs" and Reprefixation

Prefix runs are a special kind of secondary prefixation, whereby adjacent numerals come to have identical (or very similar) prefixes. This is basically an assimilatory phenomenon — an analogical interinfluence between the prefixes

on consecutive numerals.[266] Such prefix runs were already a feature of PTB, as can be deduced even if one strictly adheres to the reconstructions in STC:

$1 \leftrightarrow 2 \leftrightarrow 3$　　　　　　*g-tyik, *g-nis, *g-sum
$4 \leftrightarrow 5$　　　　　　　　*b-ləy, *b-ŋa (⋇ *l-ŋa)

Consider the numerals of rGYARONG (= JIARONG):[267]

ONE	tšek
TWO	kenes
THREE	kesom; kesam
FOUR	kewdži
FIVE	kemŋa
SIX	keṭa
SEVEN	kesñit; kesñis; kesñes
EIGHT	warže (t)
NINE	keŋgu
TEN	ṣtṣi
ELEVEN	ṣətšek
TWELVE	ṣənes
TWENTY	kenes-tṣi

Thus with reference to PTB, rGyarong has not opted to save the prefix in ONE, but has retained it in TWO and THREE; not only that, it has generalized its use all the way up to NINE (broken only by EIGHT). The case of rGyarong 10-12 is somewhat different: Here the prefix ṣ- (usually vocalized with shwa as ṣə-) has come to mean TEN, and as such is present in all the teens (including 11 and 12). See the discussion of the interinfluence of ONE and TEN (above 3.43).

Some striking examples of other secondary prefix runs include: **pa-** (LUSHAI, other Chin); **ka-** KAMAN (Miju Mishmi); **a-** (ABOR-MIRI); JINGPHO **lə-** in '1-2' (ləŋâi, ləkhôŋ) and **mə-** in '3-5' (məsūm, məlī, məŋā), etc.

We may now introduce a further terminological distinction: *perfect runs* vs. *broken runs*. A *perfect run* is an unbroken sequence of numerals with the same prefix. It may be long (e.g. the Lushai **pa-** run) or short (e.g. the PTB *b- run in '4-5'). A language may have a series of perfect runs, which among them exhaust the primary numerals (rather like a good gin rummy hand with three melds!), e.g.:

---

266) See the discussion of "Mutual influence of numerals", above 4.01. Similar phenomena are readily found in other language families, including Indo-European (above, *loc. cit.*) and such branches of Austroasiatic as Aslian (see the discussion of made-up rhyming numerals in Semai in Knowlton 1976) and Katuic (p.c., Gérard Diffloth). Ives Goddard observes that the Proto-Algonkian numerals from 1-5 all have initial **ny-**, pointing to an original quinary system (p.c. 1994).

267) My thanks to Professor Kun Chang for these data from the Zida (= Tzuta) dialect. The additional forms offered for comparison are from Nagano 1984.

1-3:   **a-** / 4-5:   **ma-** / 6-9:   **ta-**

A *broken run* is interrupted at some point by a numeral with a different prefix, or no prefix at all (e.g. the long rGYARONG **ke-** run above is broken by EIGHT which has a different prefix (**wa-**). If EIGHT were someday to succumb to the analogical pressure of its neighbors, the enlarged **ke-** run would extend all the way from 2-9. Sometimes TEN is included in a prefix run, but often a language's topmost run will end with NINE, since TEN is frequently an unprefixed monosyllable in TB (above 3.2).

Changing from a synchronic to a diachronic perspective, we may speak of *secondary prefixation* or *reprefixation*. To return to our rGYARONG example, the numerals from FOUR to SEVEN have had a secondary velar prefix superadded to their "original" ones inherited from PTB:

	PTB	*Proto-rGYARONG*	*ZIDA*
FOUR	*b-ləy	*k-b-liy	ke-w-dʑi
FIVE	*b-ŋa	*k-m-ŋa	ke-m-nga
SIX	*d-ruk	*k-d-ruk	ke-ʈa (ʈ < *-d-r-)
SEVEN	*s-nis	*k-s-nis	ke-ʂnyes, etc.

It goes without saying that the inner prefix (i.e. the one closer to the root) is historically older. (The TB languages are not much given to infixation!) The analogical pressure to create a prefix run causes a new system to be overlaid atop the old. Note that by superadding a velar to SEVEN, it is brought that much closer into line with TWO (PTB *g-nis).

Similarly, in Bodo-Garo: GARO **gni** '2' reflects inherited PTB *g-nis, but a reprefixed form **gegni** (< *g-g-nis) also occurs. This now forms a run with **gesa** 'ONE', from a root not mentioned in STC (see above 3.15).

For Old Chinese, Baxter [1985] has suggested that the initial *s- in *siəd '4' arose through the influence of the sibilant in '3' *(t)səm. (See above 1.26.)

5.21 *Mnemonic and rhythmic considerations*

Although we cannot go into this psycholinguistic topic seriously here, it seems clear that prefix runs serve an important mnemonic function, e.g. in teaching children to count. (It is even a help to English-speaking children that 'six' and 'seven' both begin with the same consonant!)

Several languages whose numerals almost all have prefixes (i.e. are "sesquisyllabic" in structure, in the sense of Matisoff 1973b) have a break or two in this rhythm at certain points. Thus, in JINGPHO:

<div align="center">

ləŋâi, ləkhôŋ — məsūm, məlī, məŋā

krúʔ

sənìt, mətsát, jəkhû

sǐ

</div>

The monosyllabic intruders serve to demarcate the string of numerals into manageable units, without making it necessary to homogenize the separate

prefixes of SEVEN, EIGHT, and NINE. The numerals can thus be recited in easy mouthfuls, giving the counter a chance to take a breath between groups:

1,2 ... 3,4,5 ... 6 ... 7,8,9 ...10.

We thus introduce the concept of the **monosyllabic breather** into prefix-run theory. (See below 5.445.)

English speakers sometimes break up a string of numbers when counting rapidly by pronouncing certain key ones *implosively*, with a sort of gasp, as they gather breath for the next sequence. The numbers of many languages have a curious tendency to fall into a natural rhythm, e.g. the Sino-Japanese set:

ichi, ni
san, shi, go
roku, shichi, hachi
kuu, juu

5.3   Numeral Prefixes in Himalayish

The languages of the Himalayish group are relatively poor in numeral prefixes, tending merely to preserve one or two of those set up for the PTB stage, and refraining by and large from introducing new ones.[268]/[269]

5.31   Bodish languages with distinctive dental-initialled SEVEN

Tibetan and its dialects, as well as other closely related Himalayish languages, have a distinctive word for SEVEN, typified by WT **bdun** (see above 4.228):

	WRITTEN TIBETAN	LHASA TIBETAN	SHERPA	JIREL	KAIKE	DZONGKHA[270]
ONE	gćig	ciːq	cikq	dokpei	ʈi	chi
TWO	gnyis	nyii	ngyi	nyiq	nghyi	nyi
THREE	gsum	sum	sumq	sumq	sum	sum
FOUR	bźi	shi	ji	syi	li	zhi
FIVE	lŋa	nga	ngaːq	ngaːq	ngaː	nga
SIX	drug	ʈhuu	ʈuk	ʈhuk	ru	ʈuk
SEVEN	bdun	tüün	din	duin	ne	duin
EIGHT	brgyad	kEE	ge	gyet	kye	gye
NINE	dgu	qu	gu	gu	gu	gu
TEN	bću	cu	citham-baːq	cyutaːm-baːq	chyu-tamba	chu-

---

268) rGYARONG, with its rich and complex prefix combinations, including double prefixes for most numerals, is probably best regarded as not belonging to the Himalayish branch of TB, but rather to the newly articulated Qiangic group. See above 1.23.

269) Many TB languages of Nepal have lost their higher numerals, replacing them with Indo-European ones from Nepali. These are discussed above 2.1.

WT preserves the PTB velar prefix run in 1-3, reminiscent of what we find in many AMD languages (below 5.51), as well as a curious pattern of prefixation in the higher numerals 6-10, which all have either **b-** (7,8,10) or **d-** (6,9) — a sort of "interdigitated" or "discontinuous" run. None of these modern Himalayish languages directly preserves any numeral prefixes, though the ***dr-** combination in SIX is reflected by retroflex initials (except in Kaike, which shows total prefix loss in SIX). Kaike is also peculiar in not sharing the special root for SEVEN with dental stop: **ne** looks as if it comes from the ordinary root ***s-nis**.

5.32   KHALING

The KHALING language of Nepal has several interesting features in its numeral system:

ONE	**tu**	SIX	**raː**
TWO	**saːhpu**	SEVEN	**taːer**
THREE	**suhpu**	EIGHT	**ri**
FOUR	**bhaːel**	NINE	**ghu**
FIVE	**bhoːm**	TEN	**tadam**

The form for TWO is of obscure origin; its initial may have been influenced by THREE. Both TWO and THREE show the suffix **-pu**, which we have already observed in KHAM and HAYU (above 2.1), perhaps originally a marker of masculine gender. **suhpu** '3' closely resembles HAYU **tshukpu**, where the final **-m** of the root has also been replaced by a velar/laryngeal element. In FOUR and FIVE the original labial prefixes have become the root initials via apocope of the root vowels, a phenomenon which occurs sporadically elsewhere in TB (above 5.131). The form for SEVEN with dental stop initial and liquid final seems related to the Bodish forms represented by WT **bdun**, though in the absence of detailed knowledge of Khaling phonology it is hard to be sure.

5.33   GURUNG-TAMANG-THAKALI and NEWARI

	GURUNG	TAMANG	THAKALI	NEWARI
ONE	**grihq**	**kiːh**	**ṭih**	**cha-gu (-li)**
TWO	**ngȋhq**	**nyiːh**	**ngih**	**ni-gu (-li)**
THREE	**sõq**	**som**	**som**	**swã-guː**
FOUR	**plihq**	**plih**	**plih**	**pẽ-guː**
FIVE	**ngaːhq**	**ngaːh**	**ngaːh**	**nyaː-gu**
SIX	**ṭuhq**	**ṭuːh**	**ṭuh**	**khu-guː**
SEVEN	**ngiq**	**nyis**	**ngis**	**nhae-guː**

---

270) These forms are from Sandberg's "Sikkim Bhutia" [1895], now the national language of Bhutan under the name Dzongkha (or Danjongka). They are closely confirmed by the tape of Chhewang Rinzin [1984]; on this tape the vowel of SEVEN sounds like barred-i: **dɨn**.

EIGHT	prehq	preht	preh	cyaː-guː
NINE	kuq	ku	ku	gũ-guː
TEN	cyuq	ci	cyu	jhi-guː

The closely-knit GURUNG-TAMANG-THAKALI group have virtually identical systems, with preservation of the labial prefix in FOUR and EIGHT (< *b-ryat), and an indirect reflection of a (dental or velar) prefix in the retroflex ʈ of SIX; no trace of a prefix appears in 2,3,5,7,9, or 10. ONE reflects a velar prefix overtly in GURUNG; in TAMANG this prefix has preempted the liquid root initial r- (presumably from PTB *ty- in *g-tyik); in THAKALI the velar prefix has fused with the r- of the root to yield a retroflex.

NEWARI maintains prefixes via preemption in FOUR (labial) and SIX (velar); the palatal initial in EIGHT points to a prototype *gyat, with neither a labial nor a liquid prefix (above 4.235).

All of these languages reflect the "normal" root for SEVEN, *(s-)nis, realized as virtually identical to TWO in Gurung-Tamang-Thakali.

5.34 Kanawari (= Kanauri) and Lepcha (= Rong)

	KANAWARI	LEPCHA
ONE	id	kat
TWO	nish	nyăt; nyi
THREE	shum	sam
FOUR	pü	fəli
FIVE	nga	fəngo
SIX	ʈuk	tərək
SEVEN	stish	kə-kyək
EIGHT	rai	kə-kŭ
NINE	zgui	kə-kyót
TEN	sai	kə-ti

The KANAWARI numerals, characterized as "merely corruptions of the Tibetan numerals" in Joshi/Rose [1909:2-3], are of course quite independent of the latter, featuring such non-Tibetan traits as id for ONE (ultimately cognate, I believe, to Lepcha kat (< PTB *k-yat; above 3.11, 3.121); preemption of the root-initial by the prefix in FOUR (< *b-[l]əy); stish for SEVEN (from the "normal" root *s-nis); no trace of a labial or velar in EIGHT; a sibilant prefix in NINE, and the form sai for TEN!

Besides preserving the labial prefix in FOUR and FIVE as fə-, and the dental prefix tə- in SIX, LEPCHA has innovated a striking velar prefix run in 6-10, unparalleled elsewhere in TB to my knowledge. Further testifying to the close interinfluence of these numerals, Lepcha seems actually to have reversed the etyma for EIGHT and NINE, with -kyót '9' apparently < *gyat EIGHT, and ku '8' apparently < *d-kəw NINE (above 4.02). The form kyək for SEVEN remains a complete mystery.

5.35 Monpa dialects

	MONPA [Dubey 1983]	M.CUONA [Sun et al. 1980]	CENTRAL MONPA [Das Gupta 1968]	M.MOTUO [Sun et al. 1980]
ONE	thee	t'eʔ^{53}	thur	t'or
TWO	nai	nʌi^{23}	n(y)itsing	ñiktsing
THREE	sum	sum^{53}	sam	sam
FOUR	blee	pli^{53}	b(i)si/p(i)si	p'i
FIVE	lenga	le^{21}nge^{53}	nga	nga
SIX	gro	kroʔ^{23}/^{54}	khung	khung
SEVEN	nis	nis^{55}	zum	zum
EIGHT	giet	cen^{13}	yen	jen
NINE	dugu	tu^{21}ku^{54}	gu	gu
TEN	chi	tɕi^{54}	se	se

- Among the Monpa dialects must be included the language known as SHAR-CHOP or TSANGLA (E. Bhutan), which seems virtually identical to Das Gupta's Central Monpa and Sun et al's Monpa Motuo. The numerals '1-10' in Sharchop, as best I could transcribe them from a tape-recording (see above, n. 1) are: **thur, nyiktsing, sam, pshi, nga, khoŋ, zon, yin, gu, še.**

- In FOUR all dialects preserve the labial prefix (with preemption of the root-initial in Motuo). In FIVE only Dubey's dialect and Cuona preserve the *l-prefix; these are also the only two dialects that preserve the dental prefix in NINE. All dialects have lost the velar prefix in TWO and THREE.

- The higher numerals show considerable interdialectal variation. Dubey's dialect and Cuona reflect the velar prefix in SIX, but the other dialects have forms with simple velar initials and aberrant nasal finals (above 4.212). Central and Motuo Monpa have an idiosyncratic word for SEVEN (**zum**), which clearly belongs with Sharchop **zon** (above 4.02), but whose further affiliations are very much in doubt (above 4.224d, 4.228). Dubey's Monpa reflects a simple velar initial in EIGHT (above 4.233), but the other dialects have forms (as in SIX) that may reflect a nasal-finalled allofam of the general root (above 4.236).

5.4 Numeral Prefixes in Kuki-Chin-Naga

5.41 Retention of the proto-system (as conceived in STC)

The only runs recognized in STC for the PTB level are 1 ↔ 2 ↔ 3 (*g-tyik, *g-nis, *g-sum) and 4 ↔ 5 (*b-ləy, *b-ŋa).

Of all the Kuki-Naga languages, only Maring (an obscure language "in the extreme south of the Naga region") presents a system more or less exactly like the one conceived of in STC for the proto-language. (In fact in all of TB only Maring and Written Tibetan have such systems!)

MARING

ONE	khat
TWO	khani
THREE	khiyum
FOUR	phili
FIVE	phanga
SIX	tharuk
SEVEN	ani
EIGHT	chot
NINE	tako
TEN	chip

In fact, among living languages, Maring may well be the winner of the Miss Proto-Tibeto-Burman Numeral Look-alike Contest. It has the 2-3 velar run,[271] and the 4-5 labial run. It presents a dental prefix in SIX and NINE (STC has *d-ruk and *d-gəw). The only innovation is the vocalic prefix in SEVEN (STC has *s-nis). The affricate in EIGHT is from some cluster that includes *-gy-, but we cannot tell exactly what combination of prefix plus initial consonant underlies this initial.

5.42 Degeneration of the proto-system: prefix loss

Some KCN languages have few or no prefixes with numerals, and thus, *a fortiori*, no prefix runs:

[A] Chang-Phom-Konyak

	CHANG	PHOM	KONYAK
ONE	chie	hük	ja
TWO	nyi	nyi	i
THREE	sam	jam	lem
FOUR	lei	ali	peli
FIVE	ngau	nga	nga
SIX	lak	vok	wok
SEVEN	nyet	nyet	nyit
EIGHT	sat	shüt	tet
NINE	guh	shü	tu
TEN	an	an	pen
TWENTY	sau-chie 20 × 1	ha	ta
HUNDRED	sau-ngau 20 × 5	gho	kho

- Note the diverse roots for ONE, and the interesting initial correspondences in THREE and SIX.
- Chang has no numeral prefixes. In Phom and Konyak, only FOUR has prefixes (and they are not the same).

---

271) Maring **khat** '1' is non-prefixed, a different etymon from *g-tyik. See above 3.12, 3.14.

- The initials in EIGHT and NINE are distinct in Chang, but have converged in the other two languages, merging to **sh-** in Phom and to **t-**in Konyak.   In Phom EIGHT has influenced NINE; in Konyak, apparently the preempting prefix in NINE has influenced EIGHT.
- Note the unusual root for TEN; is the Konyak form in **p-** related to the other two?
- The root forms of the WANCHO numerals are closely related to those of the languages in this group (below 5.43).

[B]   Kuki-Chin

	PAITE	TIDDIM	GANGTE	KUKI	THADO
ONE	khat	—	khat	—	xʌ́t
TWO	nih	nih	nih	ni; ba	ní ∼ nì
THREE	thum	thum	thum	thum	thúm;thíng
FOUR	li	li	li	li	lì
FIVE	nga	nga	ngâ	nga	—
SIX	guk	guk	gûp	gup	gúp
SEVEN	sagih	səgiʔ	sagih	sagi	sʌ́gì
EIGHT	giat	giat	giet	get	—
NINE	kua	kua	kuo	ko	—
TEN	sawm	sawm	sâwm	som	sɔ̀m

- The Paite, and most of the Tiddim, Gangte, and Kuki forms are from the *Book of Revelation* (especially XXI.19-20), where they appear as ordinals. In Paite, Tiddim, and Kuki (but not in Gangte), a cardinal numeral is turned into an ordinal by prefixing **a-** and suffixing **-na**: **a-khat-na** 'first', **a-sagih-na** 'seventh', etc.
- Note the distinctive ***kat** for ONE (above 3.12) and ***som** for TEN (above 3.231).
- Final **-k** in '6' becomes labial **-p** in Gangte, Kuki, Thado, probably via assimilation to the rounded vowel.
- The only numeral to preserve a prefix is SEVEN.[272]
- EIGHT is from ***gyat** (above 4.233).
- NINE has peculiar vocalism, pointing to an ***-a** suffix.   (See also LUSHAI, 5.43 below.)
- There are distinctive roots for THOUSAND and MYRIAD (fortunately available from contexts in the *Book of Revelation*).   See above 3.547(b).

5.421   Additive and subtractive degeneration of the proto-system
MIKIR

ONE	isi
TWO	hini

---

272) By coincidence, SEVEN is also the only dissyllabic English numeral!

THREE	kethom
FOUR	phir '4'　(phli-kep '40') [GEM];
	phli '4'　(phli-kep '40') [GRÜSSNER]
FIVE	phongo, pho [GEM]
	pho, phonho [GRÜSSNER]
SIX	throk
SEVEN	throksi
EIGHT	nirkep
NINE	sirkep
TEN	kep
ELEVEN	kre-isi

- Tones are provided in Grüssner 1979, but have been omitted here.
- This is an idiosyncratic system, well in keeping with the isolated genetic status of Mikir as a whole.　(STC hesitates to assign Mikir to the core of Kuki-Naga.)
- There is an innovative prefix in TWO (above 4.113), but prefix preservation from THREE to SIX.　The prefix in FOUR has a tendency to preempt the root-initial liquid via metathesis with the following vowel.
- Grüssner correctly calls the disyllabic form of FIVE 'older'.　There is thus a tendency to preemption here also, with the monosyllabic form (pho) having lost the root-initial nasal.
- SEVEN is an additive formation based on SIX: 7 = 6 + 1.　See above 4.204.
- EIGHT and NINE are subtractive formations based on TEN: 8 = "two from ten", 9 = "one from ten" (Grüssner: *'zwei bis zehn, eins bis zehn'*).　See above 4.203.　"*Das Element /r/* [in nirkep, sirkep] *ist zweifelsohne mit dem Affix /ra/ verwandt, das bei der Bildung zusammengesetzter Zahlen erscheint.*"
- A distinctive allomorph /kre-/ for TEN appears in the compound numeral ELEVEN (presumably 10 + 1).

5.43　Secondary general numeral prefixes

Some languages have innovated the same prefix throughout, producing a set of numerals with a single long prefix run.　In such systems, prefixes have minimum diversificatory power.　A single universal prefix is like no prefix at all:

	LUSHAI	HMAR	VAIPHEI	WANCHO
ONE	pakhat	pakhat	pakhat	tuta
TWO	pahnih	pahni	pani	ani, anyi
THREE	pathum	pathum	pathum	ajam
FOUR	pali	pali	pali	ali [DAS GUPTA],
				li [GEM]
FIVE	panga	panga	panga	aga

222

SIX	paruk	paruk	guk	arok
SEVEN	pasarih	pasari	sagi	anat
EIGHT	pariat	pariet	giat	acet [Das Gupta], achât [GEM]
NINE	pakua	pakuo	kua	aku

- Lushai and Hmar have generalized a prefix **pa-** to all the numerals from 1-9. That this is a secondary development with respect to PTB is obvious, since the **pa-** is superadded to SEVEN, which (alone of all the numerals in Kuki-Chin) always preserves its 'inner' **sa-** prefix which goes back to PTB.
- The process of generalization of a **pa-** prefix has not been carried so far in Vaiphei, so far only affecting 1-5.
- As far as the root forms of the numerals go, and in every other respect, Wancho certainly does not belong here, but rather with Phom-Chang-Konyak (above 5.42a). Like Lushai, however, it has generalized a prefix (this time **a-**) for all the numerals 1-9.[273]

5.44   Innovative runs in the higher numerals (6-9 or 6-10)

These runs involve analogical levelling or redistribution of inherited prefixes, and/or the introduction of totally new ones. These innovative prefixal systems are classifiable in several ways, especially according to their *continuity* or *discontinuity* ; i.e. whether they completely or only partially segment the numerals into consecutive sets. As always, however, some systems are idiosyncratic and resist classification (e.g. KOM REM, below 5.443).

A.   Non-exhaustive segmentation

5.441   With loss of one of the two lower runs:

	TANGKHUL	*LIANGMAI*	*YIMCHUNGRU*	*LAKHER*
ONE	akha/khatkha	khad	khülang	-kha;sa-
TWO	khani	nia	manie	-no
THREE	kathum	shum	asam	-thô
FOUR	mati	madai	phiyi	-pali
FIVE	phanga	mangiu	phüngü	-pangaw
SIX	tharuk	charuk	thruruk	-charu
SEVEN	shini	chania	thünie	-sari
EIGHT	chishat	(tachat)	tizha	-chari
NINE	chiko	chakiuh	tuku	-chaki
TEN	thara	kariu	thürü	-hraw; sy-

---

273) This may merely be an artifact of the data in GEM. It is quite possible that all the numerals in Phom-Chang-Konyak-Wancho *can* optionally take the **a-** prefix (cf. Wancho '4', given as **li** in GEM, but as **a-li** in Das Gupta 1979). This brings out the important point that having a single prefix usable with all numerals is like having no prefixes at all; in neither case are prefixes exploited for distinctive purposes.

- Tangkhul retains the 2-3 run; the pair 4-5 both show a labial prefix, but with repartition into stop vs. nasal, so the run is lost; the high run includes only 8 and 9 (**chi-**).
- Liangmai loses the 2-3 run, but retains 4-5 as **ma-**; the high run includes 6, 7, and 9 (**cha-**), but is broken by 8 (with innovative dental prefix **ta-**).
- Yimchungru loses the 2-3 run, but retains 4-5 as **phV-**; the high run extends all the way from 6-10, but is divided into two interdigitating "sub-runs": 6, 7, and 10 have aspirated **thV-**, while 8-9 have unaspirated **tV-**.
- Lakher (= Mara) loses the 2-3 run, but retains 4-5 as **pa-**; the high run includes 6, 8, and 9 (**cha-**), but is broken by 7 (which reflects original PTB *s-). Convergence has also occurred among the *rhymes* of 7, 8, 9. All these Lakher numerals may be preceded by the secondary prefixes **mia-** or **sa-** (the latter meaning ONE). This is only superficially analogous to the languages of the LUSHAI group (above 5.43) which have generalized a single prefix for all the numerals: in the latter the original prefixes have been *replaced* (except in SEVEN), while in Lakher the new generalized prefixes are *superadded* to the "inner" prefix (e.g. **sa-pangaw, sa-charu, sa-sari, sa-chari**).

5.442   With loss of both lower runs:

	ANGAMI		CHOKRI	PUIRON
	Kohima	Khonoma		
ONE	puo	po	pü	khat
TWO	kenie	kena	küna	kani
THREE	se	se	sü	thum
FOUR	die	da	da	mali
FIVE	pengou	pengu	püngu	pang
SIX	sorou	suru	shwürü	keruk
SEVEN	thenie	thena	thüna	sari
EIGHT	thetha	thetha	tütha	karet
NINE	thepfü	theku	thüchi	kakwa
TEN	kerü	kerü	küri	som

- All these languages lose the prefix for THREE, which breaks up the 2-3 run; Angami and Chokri retain the prefix in FIVE, but lose it in FOUR; on the other hand Puiron retains the prefix in FOUR, but apocopates the root-final vowel in FIVE, causing the former labial prefix to be reanalyzed as the initial consonant of the resulting monosyllable (see above 5.31).
- Angami and Chokri generalize a dental prefix for 7-9; Puiron retains the old *s- in SEVEN, but develops a velar run for 8-9.

5.443   KOM REM

ONE	inkhat	SIX	karuk
TWO	inhni	SEVEN	sari

THREE	inthum	EIGHT	karet
FOUR	manli	NINE	ko:
FIVE	ranga	TEN	som

This language has a secondary 1-3 run with in-, similar to the syllabic prefixes of the languages in 5.446, below.[274] Unlike the latter, however, the 4-5 run is absent in Kom Rem, since FIVE has a liquid (not a labial) prefix, as in Written Tibetan lṅa.[275] With the higher numerals, SIX and EIGHT form a discontinuous run in ka-, interrupted by the conservative sa- in SEVEN. The -n- in FOUR may have arisen as a 'nasal prosody' through the influence of the prefix ma-.

B. Exhaustive segmentation

In systems of this type, one or two innovative run(s) in the higher numerals directly follow two runs in the lower numerals, yielding a threeway (ternary, tripartite) or fourway (quaternary, quadripartite) grouping. In a pure system of this type, with no discontinuities, each numeral from 1 or 2 to 9 is flanked by at least one other numeral with the same prefix. This is rather similar to a winning hand in gin rummy: if each similarly prefixed sequence represents a "meld", the "hand" of numerals is exhaustively subdivided into discrete configurations.

5.444 Where both lower runs are preserved with their original prefixes, velar and labial respectively

All the languages in this group have a form for TEN with a root-initial liquid (above 3.233, 3.234).

[A]   MELURI-POCHURY-NTENYI

	MELURI	*POCHURY*	*NTENYI*
ONE	ke; kesü	khe	kesü
TWO	keni	küni	kenyi
THREE	keche	küche	keching; kechang
FOUR	mezu	mzü	mezhü; mezü
FIVE	manga	mnga	münga
SIX	taro	toro	togho; tüo
SEVEN	terü	türü	tüghü
EIGHT	tüze	tüze	tüza
NINE	tokhu	toku	tükhu
TEN	tera	türa	dagha; ta?a

• SIX has the t- prefix.

---

274) This prefix is reminiscent of the favorite Mikir prefix ing-, which occurs with many dozen common nouns (but only with one numeral, ingkoi TWENTY). See above 5.421.

275) It will be remembered that STC sets up *l-ṇa as a PTB allofam of *b-ṇa (above 4.14).

• The highest run includes 6-10.

[B] MAO-NRUANGHMEI

	MAO	*NRUANGHMEI*
ONE	kali	khüt
TWO	kahei	kanei; künei
THREE	kosü	kathum
FOUR	padei	padei
FIVE	pongo	pangu
SIX	choro	cüruk
SEVEN	chani	cünei
EIGHT	chacha	tacüt
NINE	choku	cükiu
TEN	chüro	ruh

• SIX has a palatal prefix, c- or ch-
• Mao has a neater clumping than Nruanghmei. Nruanghmei's runs do not include the "termini" ONE and TEN, and the highest run is discontinuous, broken by EIGHT.
• Mao has a distinctive root for ONE, shared e.g. by Kezhama (kele). See above 3.155.

[C] ZEME-KHOIRAO

	MZIEME	*ZEME*	*ZELIANG* [276]	*KHOIRAO*
ONE	ket	kat; hangkat	kat	khat
TWO	kena	kena	kena	kati
THREE	ketsum	kechum	kechum	kathum
FOUR	madai; mdai	medai	mdai	malhi
FIVE	mengei	mengeu	mengei	manga
SIX	heruk	seruk	heruk	saruk
SEVEN	hena	sena	sinna	sini
EIGHT	heset	desat	tesat	kachat
NINE	hekui	sekui	hekui	chaku
TEN	kerei	kereu	kerei	sara
	h-: 6-9	s: 6-7,9	h: 6,9	s: 6-7,10

• SIX has s- or h- prefix.
• ONE lacks a velar prefix.
• In the highest run, Mzieme has a perfect sequence 6-9, but Zeme lacks 8, Zeliang lacks 7-8, and Khoirao lacks 8-9.
• Khoirao has a distinctive form for TWO, kati.

---

276) Zeliang is a kind of composite dialect or lingua franca, an acronym for Ze-me + Liang-mai.

5.445   *Where the lowest run has a vocalic prefix*

In these languages the numerals 1-3 (or 2-3) typically have a vowel prefix; 4-5 have a labial stop or nasal; and the higher numerals 6-9 (or 6-10) have a dental or palatal prefix.

	LOTHA	YACHAM-TENGSA	SANGTAM	AO Mongsen	Chungli
ONE	ekha	khatu	khe; khürü	akha/ra	ka
TWO	eni/oni	anat	anyü	anet	ana
THREE	etham	asam	asang	asam	asem
FOUR	mezü	phale	müzyü	phüli	pezü
FIVE	mungo	phungu	münga	phanga	pungu
SIX	tirok	thelok	thüro	terok	trok
SEVEN	ti-ing	thanyet	thünye	teni	tenet
EIGHT	tiza	thesep;teset	ke	tsit	ti
NINE	toku	thaku	tüku	tüku	tuku
TEN	taro	thelu	thüre	tera	ter

- The runs are perfectly unbroken in Lotha and Yacham-Tengsa; in the other three languages the uppermost run is broken by EIGHT, which is prefixless and monosyllabic. (Cf. the concept of the "monosyllabic breather" introduced above, 5.21.) We may call such interrupted runs as these *discontinuous* runs.
- Sangtam shows vacillation in aspiration in the highest run.
- Note the distinctive words for TEN < *rok (above 3.234).
- Note the apocope in Ao Chungli TEN. This form bears no relationship to the similar looking AMD root *tel 'ONE' (above 3.151).

In this group also belong the following, where the highest run is also discontinuous, broken either by SEVEN (Tangsa, Kimsing) or subtractively by EIGHT and NINE (Meithei):

	TANGSA Moshang	Yogli	KIMSING	MEITHEI
ONE	ashi	ashi	ashi	ama
TWO	ani	anei	anai	ani
THREE	atum	adim	acam	ahum
FOUR	bali	bülai	balai	mari
FIVE	banga	banga	bangi	manga
SIX	taruk	türuk	tarok	taruk
SEVEN	mashi	mishi	mishi	taret
EIGHT	tachat	tüchat	techat	[nipal/nipan]
NINE	takru	tükau	tak(a)u	[mapan]
TEN	rok-shi	rauk-shi	ro-shi	tara

5.446　Where the lowest run has a fully syllabic CVC- prefix
NOCTE-MARAM

	NOCTE		MARAM
	[GEM]	[DUBEY]	
ONE	vanthe	wanthe	hang-li-ne
TWO	vanyi	wanni	hangna
THREE	vanram	wanrom	hangtum
FOUR	beli	bali	madai
FIVE	banga	bang	mingu
SIX	irok	iro:k	saruk
SEVEN	ingit	ingit	sina
EIGHT	isat	itse:t; iset	sachat
NINE	ikhu	ikhu	soki
TEN	ichi	ichi	kero

- Note the different roots for ONE in the two languages.
- The higher numerals show perfect runs of 6-9 (Maram) or 6-10 (Nocte). Nocte has the unusual i- prefix here.
- For similar syllabic prefixes in 1-3, see KOM REM (above 5.443).

5.447　Quadripartite runs: where the two lower runs are preserved, and the higher numerals show two successive innovative runs

Four is the maximum number of runs attested from 1-9:

1-3 *or* 2-3 / 4-5 / 6-7 / 8-9.

KEZHAMA-SEMA-RENGMA

	KEZHAMA	SEMA	RENGMA
ONE	kele	laki; khe	me
TWO	kenhi	kini	khohüng
THREE	katsü	küthu	keshan
FOUR	pedi	bidhi	pezi
FIVE	pangu	pongu	pfü
SIX	sarü	tsogho	tsaro
SEVEN	sinyi	tsini	tsanü
EIGHT	tiche	thache	tütse
NINE	tepfü	toku	tükhü
TEN	chiro	chüghi	tsarü

- Rengma has preemption in FIVE. The runs 2-3, 4-5 are less obvious in Rengma, because of aspiration differences in 2-3 and preemption in FIVE.
- In Sema the prefixes in the second run have a voicing difference, and those in the third run have an aspiration difference. In Kezhama, the prefixes fall into four perfect pairs, though the vocalism of the members of each pair is different.

5.5   Numeral Prefixes in Abor-Miri-Dafla

5.51   *Runs in the lower numerals*

With few exceptions, only two prefixes appear in these languages: (a) the velar **ka-** (or rather **kV-**) and (b) a naked vowel, usually **a-** (but also sometimes **e-** or **o-**). (An exception is MILANG, which has **pV-** for 4-5). Not only may we generalize with respect to the *repertoire* of prefixes here, but also with respect to the *domain* of the runs. Instead of the two separate runs 2-3 and 4-5, these languages exhibit enlarged or consolidated runs (usually also generalized backwards to include 1 and/or forwards to include 6: i.e. a single run from 1-5 or from 1-6). Sometimes this long lower run is discontinuous at some point. Runs in higher numerals are virtually non-existent. (Again, MILANG is an exception, with **ra-** for 7-8; also APATANI has **kV-** for 6-7, and IDU MISHMI has **i-** for 7-8). This is because of the bizarre replacive roots for higher numerals which are characteristic of AMD.

5.511   *Where the run has a velar prefix*

	IDU [Talukdar]	MISHMI [Dubey]	CHULIKATA[277] [LSI]	TARAON[278] [NEFA]	DENG DARANG [Sun et al. 1980]
ONE	khe(ng)ge	khege	eːkheː	khing	k'ɯn⁵⁵
TWO	kanyi	kani	kaːni	kaːing	kə²¹n⁵⁵
THREE	kasõ	kaso	kaːsh	kaːsaːng	kə²¹sɯng⁴⁵
FOUR	kapri	kapri	kaːppi	kaːpraːi	kə²¹pʁəi⁵⁵
FIVE	manga	manga	maːngaː	maːngaː	mə²¹ngə⁴⁵

In these languages the velar run is only from 2 to 4 (the velar in ONE seems to be the root initial — above 3.12). In the following language, whose dialects are known variously as MIJU, MIJU MISHMI, KAMAN, or DENG GEMAN, the velar run is extended in both directions, and extends all the way from 1 to 6:

	MIJU MISHMI [LSI]	MIJU [Das Gupta 1977a]	DENG GEMAN [Sun et al. 1980]
ONE	kɯoː/komoː	kumo	kɯ²¹mu⁵³
TWO	kaːning/kinnin	kinin	kɯ²¹jin⁵³
THREE	kaː-saːm	ksam	kɯ²¹săm⁵³
FOUR	kambrin	kambran	kɯ²¹bRɯn⁵³
FIVE	ka-liːn	klin	kɯ²¹len⁵⁵
SIX	kaːtaːm	katam	kɯ²¹təm⁵³

Note the characteristic nasal-finalled forms for FOUR (above 4.136), as well as the totally idiosyncratic forms for FIVE and SIX (above 4.147, 4.218).

---

277) = "Taying Mishmi".

278) Virtually identical to "Digaru Mishmi" (LSI III.1, 623), which has eːkhing, kaː-ying, kaː-sång, kaː-prei for 1-4.

5.512 Where the run has a vocalic prefix

(A) With a perfect 1-6 run in a-

	ABOR-MIRI	GALLONG	MINYONG
ONE	a-ko; a-ter/-tel	ako/aken	akon; atir/ayirr
TWO	a-nyi	anyi	anyi
THREE	a-um/a-ngum	aum	aum
FOUR	a-pi	appi	aki [279]
FIVE	a-ngo	ango	ango
SIX	a-keng/a-ke'	akke	akeng/akkeng

(B) With vocalic variation in the prefix (1-6 run)

	TAGIN [Das Gupta 1975]	LHOPA [Sun et al. 1980]
ONE	akin	ako; aken
TWO	anyi	anĩ
THREE	aum	afium
FOUR	epí	apiː
FIVE	ango	ongo
SIX	aké	akɯ

In Tagin FOUR has e-; in Lhopa FIVE has o-.   Tagin epí and aké are high tone (marked by acute accent).

(C) With a break in the 1-6 run

	PADAM [Dubey 1983]	NISHI [Dubey 1983]
ONE	akem/atel	lacking
		[Das Gupta 1977b]
TWO	ani	anni
THREE	aum	om
FOUR	appi	appi
FIVE	pilngo	aːngo
SIX	akke	akke

For FIVE Padam shows an interesting form with double prefix, < *b-l-ŋa. This is one case where STC does set up prefixal variation in a numeral at the PTB level, reconstructing both *b-ŋa and *l-ŋa. If anything the Padam evidence might suggest that the "inner" l- prefix is more primary than the "outer" labial one, so that only *l- should be set up for the PTB stage; but I do not feel that we can make such a rigid distinction between proto-variation and diachronic change in TB prefixes at the present state of our knowledge (and perhaps in principle).

For THREE, Nishi (like some other Dafla dialects: see below) has a monosyllabic form with o- vocalism, which clearly derives from a disyllabic

---

279) See above 4.136.

form with the **a-** prefix, as in Padam **a-um**. The fusing of the **a-** prefix with the root-vowel **-u-** to yield **o-** was made possible by the total loss of root-initial ***s-**, which is characteristic of many AMD languages.[280] This fusion of prefix and root into a monosyllabic unit breaks up the rhythm of the 1-6 run.

(*D*)   *Dafla dialects*

	DAFLA [LSI:ROBINSON]	*E. DAFLA* [LSI:HAMILTON]	*YANO DAFLA* [N.L.BOR 1938]	*DAFLA* [DAS GUPTA 1969]
ONE	aː-kin	akkin	akhin	aking/aku
TWO	aː-ni	anyi	anyi	anyi/ain[281]
THREE	aː-aːm	a-om	um	om
FOUR	aː-pli	a-pl	apli; appi	api
FIVE	aː-ngoː	aː-ng	ango	ango
SIX	akple	aː-kr	akke	aké

- Note the apocope in the E. Dafla forms FOUR, FIVE, SIX, paralleled also in its word for EIGHT **pliːn** (compare Yano Dafla **plönö**).

- Robinson's word for SIX has a strange consonant sequence **-kpl-**, where the **-p-** has perhaps crept into the form through contamination by FOUR. Similar interinfluence between successive numerals seems to be at work in Robinson's words for SEVEN and EIGHT: **kaːnag** '7', **plag-nag** '8' (see above 4.20).

5.52   Systems with no secondary prefixation in the lower numerals
AKA-APATANI-MILANG-SERDUKPEN

These languages form a miscellaneous group. Besides their prefixal paucity in the lower numerals, they are all characterized by highly idiosyncratic sets of *higher* numerals, though this trait is shared by many other AMD languages (above 4.2).

	AKA (= *Hruso*) [LSI]	*APATANI* [SIMON 1972]	*MILANG* [DAS GUPTA 1980]	*SERDUKPEN* [DUBEY 1983]
ONE	a	kũ	akan; atel	han
TWO	kshi	nĩ	ne	n(y)ik
THREE	zu	hĩ	ham	ung
FOUR	fi-ri	pe	pe	biːsi
FIVE	phum	ngo	pangu	khu

Only Aka preserves the velar prefix in TWO; there is no trace of any prefix in THREE; the labial prefix in FOUR is well-preserved (with preemption of the root initial in Apatani and Milang). In FIVE, the denasalized Serdukpen form perhaps reflects a lost **p-** prefix, preserved in Milang; in Aka **phum** the labial

---

280) See Matisoff 1978a:277-278 (n.258).

281) Several AMD languages show a tendency to metathesize the initial consonant and vowel of TWO. An intermediate stage is represented by the syllabic nasal in DENG DARANG (above 4.111).

stop prefix has become the root-initial via apocope of the root-vowel — i.e. the final **-m** seems to reflect the original root-initial (above 5.131).

Apatani has a special set of numerals used in counting humans, which seems more conservative with respect to prefix preservation, e.g. '3' **hingi**, '4' **pilye**, '5' **yango**.

6. SUMMARY AND AFTERWORD

The standard (STC) reconstructions for the PTB numerals stand up quite well, though we have nuanced them, especially with respect to the treatment of the prefixes. Reconstructions for several new numerical roots and allofams are offered, including ONE, SIX, NINE, and TEN. Several examples of previously attested variational patterns are provided by these new or revised reconstructions, e.g.: **-i- ⋇ -ya** (ONE *tik ⋇ tyak [3.14]; TEN *gip ⋇ gyap [3.21]; *tsiy ⋇ tsyay [3.22]; TEN/HUNDRED *liŋ ⋇ *lyaŋ [3.23]); **-ay ⋇ -an** (ONE *tay ⋇ tan [3.14]); **-u- ⋇ -a-** (THREE *-sum ⋇ *-sam [4.12]; TWENTY *m-kul ⋇ *kal [3.511]). We have emphasized the inter-influence of numerals in sequence, first in the context of general variational patterns in ST word families (1.11), then in more specific morphophonemic and semantic terms (4.01-4.02), finally focussing on prefixal behavior in numeral sets (5.2 et seq.). We have seen how some languages express their higher unit numerals (6-9) in additive, subtractive, or multiplicative formations (4.20).

Throughout we have not merely been concerned with reconstructing the etyma for the individual units (1-9) in ST numeral sets, but have paid particular attention to their *systemic structure* (1.12), as revealed by the various languages' methods of *TEEN-* and *ROUND-NUMBER* formation (3.3-3.5). We have pointed to striking cases of hesitation, flux, or transvaluation in the arithmetical bases of TB numeral systems, e.g. between ONE and TEN (3.4). Such phenomena are characteristic of Himalayish languages under strong contact influence (2.0-2.1), including hesitation between TEN and TWENTY as bases for the system (e.g. Sherpa, Lepcha, Dzongkha: 3.534), and even between TWENTY and TWELVE (Chepang: 3.535). We have noted traces of FOUR- (Boro, Kubhinde Dumi: 3.32c) and FIVE-based (Bantawa: 4.15) systems, which may well prove to represent ancient types of numerical organization in the family.

On the semantic side, we have found a few interesting cases of "transfield associations" between numeral concepts and roots from other semantic fields, e.g. between FIVE and *hand* (4.14-4.15), and FIFTY and *ridgepole* (3.522). For a diagrammatic representation of the various semantic interconnections uncovered among the numerals, see the semantic flowchart in Appendix I.

Aside from purely etymological problems like finding affiliations for the weird or isolated numeral forms that crop up here and there (especially in

AMD), there remain plenty of intriguing conceptual puzzles for further research. As a random example, one could cite the strange LAHU classifier lɛ̂, which in some dialects (including varieties of Black Lahu) occurs as the general classifier, but only after the numerals 3, 4, and 9. (The ordinary Lahu general classifier is mâ, functionally equivalent to Mandarin gè or Thai ʔan.) Roop [1970:62-63] reports a similar LISU general classifier lyö¹¹, occurring only after lyi⁵⁵ '4', which he characterizes as a "suppletive allomorph" of the ordinary general classifier ma³³. In the Qiangic language MUYA, a possibly cognate form lø⁵⁵ occurs as a citation classifier, but after *all* the numerals from 1-10 (e.g. so⁵⁵- lø⁵⁵ '3', ʐɯ³⁵-lø⁵⁵ '4', ŋa⁵⁵-lø⁵⁵ '5', etc. Another Qiangic language, SHIXING, has a different etymon for its ordinary general classifier, ko³³, but with a special allomorph ko³⁵, under a different tone, after the numerals 2, 4, and 9! What numerical sense does this make? What do 3/4/9/ or 2/4/9 have in common that would motivate these "special general" classifiers?

This paper should be viewed as part of the *Sino-Tibetan Etymological Dictionary and Thesaurus* project (STEDT), a longterm effort to reconstruct the lexicon of PTB/PST by semantic field, with the ultimate aim of recovering as much as possible of the semantic and phonological richness of the ancient lexicon.[282]

---

282) As this paper was going to press, I learned of a large-scale project on the "typology of numeral systems" being carried out at the University of Madrid, which includes data on the minority languages of East Asia. The present study should provide much grist for their mill. See Marcos-Marín 1993.

Appendix I.

METASTATIC FLOWCHART OF NUMERICAL SEMANTIC ASSOCIATIONS
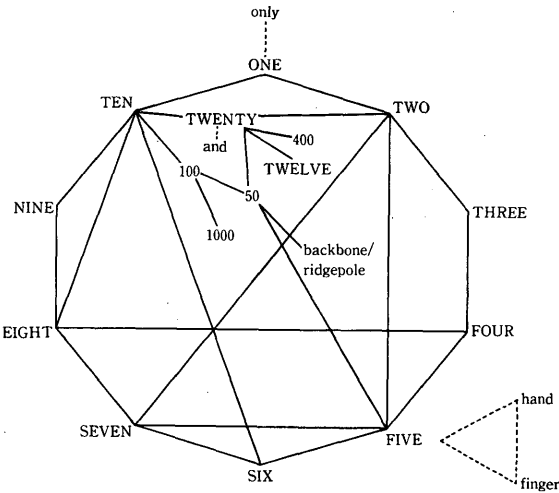
chart by Jonathan P. Evans

Appendix II.

INDEX OF RECONSTRUCTED ETYMA

ONE

TWO

THREE

FOUR

FIVE

SIX

SEVEN

EIGHT

NINE

TEN

TWENTY

HUNDRED

THOUSAND

Appendix III.

INDEX OF LANGUAGES AND SOURCES

Abor-Miri-Dafla

Abor-Miri: Lorrain, 1907. LSI III.1:622.
Apatani: Simon, 1972:9-11. Sun J.T., 1993.
Bengni: Sun J.T., 1993.
Bogaer Luoba: TBL.
Bokar: Sun J.T., 1993.
Bokar Adi: ZMYYC.
Chulikata (Taying Mishmi): LSI III.1:623.
Dafla (Nishi): Robinson, 1851. Hamilton, 1900. LSI III.1:622. Bor, 1938 (Yano Dafla). Das
    Gupta, 1969:2.
Damu: Sun J.T., 1993.
Deng (Kaman, Geman): Sun Hongkai, et al., 1980. ZMYYC. TBL.
Deng (Taraon, Darang): Sun Hongkai, et al., 1980:384-387. ZMYYC. TBL.
Gallong: Dubey, 1983. Das Gupta, 1963.
Hrusso (Aka): LSI III.1, 622-623.
Idu: Talukdar et al., 1962:15. ZMYYC.
Idu Luoba: TBL.
Lhopa: Sun Hongkai, et al., 1980:384-387.
Miji: Simon, 1979.
Miju Mishmi (Deng, Kaman, Geman): Das Gupta, 1977a:19-20.
Milang: Das Gupta, 1980. Sun J.T., 1993.
Minyong (E. Dafla): Das Gupta, 1977b:16-22.
Miri: Simon, 1976. Sun J.T., 1993.
Mishmi: Dubey, 1983.
Mising: Sun J.T., 1993.
Nishi: Dubey, 1983.
Padam: Das Gupta, 1977b:16-22. Dubey, 1983. Sun J.T., 1993.
Padam-Mishing: Sun J.T., 1993.
Serdukpen: Dubey, 1983.
Tagin: Das Gupta, 1975.
Taraon (Digaro): Chakravarty, et al., 1963.

Baic

Bai (Bijiang): ZMYYC.

Bai (Dali): ZMYYC.
Bai (Jianchuan): ZMYYC.
Bai (Minchia): Dell, 1981. TBL.

Bodo-Garo

Boro: Bible Society of India, 1972b, *Revelation* 21:19-20. Bhat, 1968:29-30.
Dimasa: Marrison, 1967.
Garo: Momin, n.d. Burling, 1961:57-58. Phillips, 1904.

Himalayish

Athpare: Gvozdanović, 1985.
Bahing: Gvozdanović, 1985.
Baima: Sun Hongkai, 1991 (p.c.).   ʳ
Bantawa: Gvozdanović, 1985.
Chepang: Hale (ed.), 1973.
Chourase: Gvozdanović, 1985.
Dumi: Gvozdanović, 1985.
Dzongkha: Mazaudon, 1985. Rinzin, 1984 (p.c.).
Gurung: Hale (ed.), 1973.
Hayu: Michailovsky, 1981:167. LSI III.1 (Vayu):384-385.
Jirel: Hale (ed.), 1973.
Kaike: Hale (ed.), 1973.
Kanawari: Joshi, 1909:2-3.
Khaling: Hale (ed.), 1973. Toba and Toba, 1975.
Kham (Nepal): Hale (ed.), 1973.
Kulung: Gvozdanović, 1985.
Lepcha: Mainwaring and Grünwedel, 1898.
Magari : Hale (ed.), 1973.
Mewahang: Gvozdanović, 1985.
Monpa: Das Gupta, 1968:101. Nishi, 1982. Dubey, 1983.
Monpa Cuona (Takpa): Sun Hongkai, et al., 1980. ZMYYC. TBL. Sun J.T., 1993.
Monpa Motuo (Tsangla): Sun Hongkai, et al., 1980:384-387. ZMYYC. TBL.
Newari: Hale (ed.), 1973.
Sharchop: Rinzin, 1984 (p.c.).
Sherpa: Hale (ed.), 1973.
Sikkim Bhutia (Dzongkha): Sandberg, 1895:59.
Sunwar: Hale (ed.), 1973. Gvozdanović, 1985:143.
Tamang: Hale (ed.), 1973.
Thakali: Hale (ed.), 1973.
Thulung Rai: Allen, 1975:102-103.
Tibetan (Lhasa): Goldstein and Nornang, 1970:395-396.
Tibetan (Written): Jäschke, 1881.
Tsangla (Northern and Southern): Nishi, 1982.
Yakkha: Gvozdanović, 1985.
Yakkhaba: Gvozdanović, 1985.

Jingpho-Nungish

Anong Nu: TBL. ZMYYC.

Dulung (Trung): Sun Hongkai, 1982b:54. ZMYYC. TBL.
Jingpho (Kachin): Hanson, 1906/1954. Maran (in prep.).
Jingpho: ZMYYC. TBL.
Nusu Nu: TBL.

Karenic

Kayah: Solnit, 1984.
Palaychi: Jones, 1961.
Pa-O (=Taungthu): Jones, 1961.
Pho (Bassein, Moulmein): Jones, 1961.
Sgaw (Bassein, Moulmein): Jones, 1961.
Kelun: TBL.

Kuki-Chin-Naga

Angami (Khonoma): Marrison, 1967.
Angami (Kohima): Marrison, 1967.
Ao (Chungli): Marrison, 1967.
Ao (Mongsen): Marrison, 1967.
Bawm (=Laizo): Osburne, 1975.
Chakhesang: Nagaland Bhasha Parishad, 1972a.
Chang: Marrison, 1967.
Chokri: Marrison, 1967.
Gangte: Bible Society of India, 1972a:512.
Hmar: Bible Society of India, 1970, *Revelation* 21:19-20.
Kheja: Nagaland Bhasha Parishad, 1974.
Khezhama: Marrison, 1967.
Khoirao: Marrison, 1967.
Kimsing: Das Gupta, 1978:12.
Kokborok (Tripuri): Karapurkar, 1976:45-48.
Kom Rem: Bible Society of India, 1976, *Revelation* 21:19-20.
Konyak: Marrison, 1967.
Kuki: Bible Society of India, 1973, *Revelation* 21:19-20.
Lakher (=Mara): Lorrain, 1951.
Liangmei: Marrison, 1967.
Lotha: Marrison, 1967.
Lushai: Marrison, 1967.
Manipuri (=Meithei): Marrison, 1967.
Mao: Marrison, 1967.
Maram: Marrison, 1967.
Maring: Marrison, 1967.
Meluri: Marrison, 1967.
Mikir: Marrison, 1967. Grüssner, 1979:63-64.
Mzieme: Marrison, 1967.
Nocte: Marrison, 1967.
Nruanghmei: Marrison, 1967.
Ntenyi: Marrison, 1967.
Paite: Bible Society of India, 1974, *Revelation* 21:19-20.
Phom: Marrison, 1967.
Pochury: Nagaland Bhasha Parishad, 1972b: 15-16.

Puiron: Marrison, 1967.
Rengma: Marrison, 1967.
Sangtam: Marrison, 1967.
Sema: Marrison, 1967.
Tangkhul: Marrison, 1967.
Tangsa (Moshang): Marrison, 1967.
Tangsa (Muklom): Dubey, 1983.
Tangsa (Yogli): Marrison, 1967.
Thado: Thirumalai, 1972.
Tiddim Chin: Henderson, 1965. Bible Society of India, 1979, *Revelation* 21:19-20.
Vaiphei: Bible Society of India, 1971, *Revelation* 21:19-20.
Wancho: Marrison, 1967. Das Gupta, 1979:27-28. Dubey, 1983.
Yacham-Tengsa: Marrison, 1967.
Yimchungrü: Marrison, 1967.
Zeliang: Nagaland Bhasha Parishad, 1973.
Zeme: Marrison, 1967.

Lolo-Burmese

Achang: ZMYYC. TBL.
Ahi: Yuan Jiahua, 1953.
Akha: Lewis, 1968.
Bisu: Bradley, 1979.
Bola: TBL.
Burmese (spoken): ZMYYC. TBL.
Burmese (Written): Judson, 1893/1953/1966.
Gazhuo: TBL.
Hani (Caiyuan = Biyue): ZMYYC.
Hani (Dazhai): ZMYYC.
Hani (Lüchun): TBL.
Hani (Mojiang): TBL.
Hani (Shuikui = Haoni): ZMYYC.
Hani: Hu Tan and Dai Qingxia, 1964. Gao Huanian, 1955.
Jinuo: ZMYYC. TBL.
Lahu (Lancang): TBL.
Lahu: Matisoff, 1973a. ZMYYC.
Langsu (Maru): ZMYYC. TBL.
Leqi: TBL.
Lisu: Fraser, 1922. ZMYYC. TBL.
Luquan: Ma Xueliang, 1949.
Mpi: Srinuan, 1976:538-541.
Nasu: Gao Huanian, 1958.
Naxi (Lijiang): ZMYYC.
Naxi (Yongning = Moso): ZMYYC.
Naxi: Rock, 1963. TBL.
Nusu: ZMYYC.
Phunoi: Bradley, 1979:338-341.
Sani: Ma Xueliang, 1951:81. TBL.
Ugong (Kanburi Lawa): Bradley, 1978.
Woni: Yuan Jiahua, 1947.
Xiandao (Achang): TBL.

Yi (Dafang): ZMYYC.
Yi (Mile = Axi): ZMYYC.
Yi (Mojiang): ZMYYC.
Yi (Nanhua): ZMYYC. TBL.
Yi (Nanjian): ZMYYC.
Yi (Weishan): TBL.
Yi (Wuding): TBL.
Yi (Xide): ZMYYC. TBL.
Zaiwa (Atsi): ZMYYC. TBL.

Qiangic

Daofu: TBL.
Ergong: ZMYYC.
Ersu (Tosu): Sun Hongkai, 1982a. ZMYYC.
Guiqiong: ZMYYC. TBL.
Muya: ZMYYC. TBL.
Namuyi (Namuzi): ZMYYC. TBL.
Pumi (Qinghua): ZMYYC. Lu, 1983:37, 128.
Pumi (Jiulong): TBL.
Pumi (Lanping): TBL.
Pumi (Taoba): ZMYYC. Lu, 1983:37, 128.
Qiang: Sun Hongkai, 1981:88. TBL.
Qiang (Mawo): ZMYYC.
Qiang (Taoping): ZMYYC.
Queyu (Zhábā): ZMYYC. TBL.
rGyarong (Jiarong): Nagano, 1979. ZMYYC. TBL.
Shixing: ZMYYC. TBL.
Zhábà: ZMYYC. TBL.

Other

Sulong (Sulung): ZMYYC.
Tujia: ZMYYC. TBL.
Nepali (Indo-European): Hale (ed.), 1973.

BIBLIOGRAPHY

ALLEN, Nicholas J.
  1975 *Sketch of Thulung Grammar, with Three Texts and a Glossary.* Cornell University East Asia Papers, #6. Ithaca: Cornell University.
BARNARD, J.T.O.
  1934 *A Handbook of the Rawang Dialect of the Nung Language.* Rangoon.
BAXTER, William
  1985 Tibeto-Burman Cognates of Old Chinese *ij and *ɨj. In G. Thurgood, J.A. Matisoff, and D. Bradley (eds.), pp. 242-263.
BENEDICT, Paul K.
  1972 *Sino-Tibetan: A Conspectus.* Contributing Editor, James A. Matisoff. Princeton-Cambridge Studies in Chinese Linguistics II. Cambridge: University Press. ("STC")
  1979 Four Forays into Karen Linguistic History. *LTBA* 5.1: 1-35.

1983   Qiang Monosyllabization: A Third Phase in the Cycle. *LTBA* 7.2:113–114.

BHAT, D.N.S.

1968   *Boro Vocabulary, with a Grammatical Sketch.* Poona: Deccan College Postgraduate and Research Institute.

1969   *Tankhur Naga Vocabulary.* Poona: Deccan College Postgraduate and Research Institute.

BIBLE SOCIETY OF INDIA

1970   *Pathien Lekhabu Inthieng.* [Holy Bible in HMAR: Old and New Testaments.] Bangalore. 1710 pp. (O.T.) + 533 pp. (N.T.), paginated separately.

1971   *Thuthung Thak le Sam.* [New Testament and Psalms in VAIPHEI.] Bangalore. 524 pp. (N.T.) + 137 pp. (Psalms), paginated separately.

1972a   *Thuthunthak leh Sam.* [New Testament and Psalms in GANGTE.] Bangalore. 513 pp. (N.T.) + 189 pp. (Psalms), paginated separately.

1972b   *Godan Rodai.* [New Testament in BORO.] Bangalore. 657 pp.

1973   *Pathen Lekhabu Theng.* [Holy Bible in KUKI.] Bangalore. 1296 pp. (O.T.) + 322 pp. (N.T.), paginated separately.

1974   *Laisiangthou.* [Holy Bible in PAITE.] Bangalore. 973 pp. (O.T.) + 307 pp. (N.T.), paginated separately.

1976   *Tongtep Ther leh Lha.* [New Testament and Psalms in KOM REM.] Bangalore. 568 pp.

1979   *Lai Siangtho.* [Holy Bible in TIDDIM CHIN.] Bangalore. 1072 pp. (O.T.) + 350 pp. (N.T.), paginated separately.

BOLTZ, William

1969   The Numeral ONE in Old Chinese. *Unicorn* IV: 56-59.

1977   The Etymology of the Old Chinese Numeral TWO: Grammatical and Semantic Consideration. Paper presented at Tenth Sino-Tibetan Conference, Georgetown University, Washington, D.C.

BOR, N.L.

1938   Yano Dafla Grammar and Vocabulary. *JRASB* 4:217-281.

BRADLEY, David

1978   Identity, Dialect, and Sound Change in mBisu and Ugong. *Working Papers in Linguistics, University of Melbourne* 4:37-46.

1979   *Proto-Loloish.* Scandinavian Institute of Asian Studies Monograph Series, #39. Copenhagen and London.

BRADLEY, David, E.J.A. HENDERSON, and Martine MAZAUDON (eds.)

1989   *Prosodic Analysis and Asian Linguistics to Honour R.K. Sprigg.* Pacific Linguistics C-104. Canberra: Australian National University.

BURLING, Robbins

1961   *A Garo Grammar.* Deccan College Postgraduate and Research Institute Monograph Series, #25. Poona.

1968   *Proto-Lolo-Burmese.* IJAL Special Publication 33.2, Part II. Issued simultaneously as Indiana Publications in Anthropology and Linguistics, #43. The Hague: Mouton and Co.

CAUGHLEY, Ross C.

1972   A Vocabulary of the Chepang Language. 40 pp. mimeo. Kirtipur, Nepal: Summer Institute of Linguistics.

1989   Chepang: A Sino-Tibetan Language with a Duodecimal Numeral Base? In Bradley et al. (eds.), pp. 197-199.

CHAKRAVARTY, L.N. et al.

1963   *A Dictionary of the Taraon Language.* Shillong: Philology Section, Research Department, Northeast Frontier Agency.

CHEN Weidong
　　1990　[Numerals in the Naxi Language.] In Chinese.　Unpublished research report, in a collection entitled *Research on Numerals in Tibeto-Burman Languages.*

CHINESE ACADEMY OF SOCIAL SCIENCES
　　1991　*Zang-Mianyu Yuyin he Cihui.* [Tibeto-Burman Phonology and Lexicon.]　Compiled and edited by Xu Jufang, Liu Guangkun, Sun Hongkai et al.　1420 pp.　Beijing. ("ZMYYC")

DAI Qingxia
　　1982　*[Outline Grammar of the Achang Language.]* In Chinese.　Beijing: People's Publishing Co.

DAI Qingxia et al.
　　1981　= Nhkum Seng Hkum et al. 1981.
　　1983　= Xu Xijian et al. 1983.

DAI Qingxia and HUANG Bufan
　　1992　*Zang-Mian Yuzu Yuyan Cihui.* [A Tibeto-Burman Lexicon.] In Chinese. Beijing: Central Institute of Nationalities.

DAS GUPTA, K.
　　1963　*An Introduction to the Gallong Language.* Shillong: Research Dept., NEFA.
　　1968　*An Introduction to Central Monpa.* Shillong: Philology Section, Research Department, Northeast Frontier Agency.
　　1969　*Dafla Language Guide.* Northeast Frontier Agency.　114 pp.
　　1975　A Few Features of the Tagin Language. *Resarun* 1.4: 6 pp.
　　1977a *A Phrase Book in Miju.* Published by Director of Information and Public Relations, Arunachal Pradesh.　Shillong.
　　1977b A Few Aspects of the Minyong Language. *Resarun* 3.4: 16-22.
　　1978　A Note on the Tangsa Language. *Resarun* 4.2: 6-16.
　　1979　A Note on the Wancho Language of Arunachal Pradesh. *Resarun* 5.1: 25-37.
　　1980　A Note on [the] Milang Language. *Resarun* 6.2: 14-18.

DELL, Francois
　　1981　*La langue Bai, phonologie et lexique.*　Paris: Ecole des Hautes Etudes en Sciences Sociales.

DIFFLOTH, Gérard
　　1976　Mon-Khmer Numerals in Aslian Languages. In Gérard Diffloth and Norman Zide (eds.), pp. 31-37.

DIFFLOTH, Gérard and Norman ZIDE (eds.)
　　1976　*Austroasiatic Number Systems.*　Special issue of *Linguistics: An International Review,* #174.　The Hague and Paris: Mouton.

DUBEY, Shail Kumari
　　1983　*Linguistic Study of Arunachal Dialects: Phonology.*　(Contains data on Gallong, Mishmi, Monpa, Muklom Tangsa, Nishi, Padam, Serdukpen.) Unpublished D. Litt. dissertation, University of Calcutta.　413 pp.

EMENEAU, Murray B.
　　1957　Numerals in Comparative Linguistics (with special reference to Dravidian). *BIHP* 29: 1-10.

EVANS, Jonathan P.
　　1993　[Fieldnotes on the Sanlong Dialect of Qiang.] Unpublished.

FRASER, J.O.
　　1922　*Handbook of the Lisu (Yawyin) Language.* Rangoon.

FRENCH, Walter T.
　　1983　*Northern Naga: A Tibeto-Burman Mesolanguage.* Ph.D. dissertation, City University of New York.　737 pp.

GAO Huanian
    1955   [A Preliminary Investigation of the Hani Language of Yang-wu.] In Chinese. *Journal of Zhongshan University*, pp. 175-231.
    1958   *[A Study of the Grammar of the Yi (Nasu) Language.]* In Chinese. Beijing: Scientific Publishing Co.

GOLDSTEIN, Melvyn C. and Nawang NORNANG
    1970   *Modern Spoken Tibetan: Lhasa Dialect.* Seattle and London: University of Washington Press.

GREENBERG, Joseph H.
    1987   *Language in the Americas.* Stanford: Stanford University Press.

GRIERSON, Sir George A. and Sten KONOW (eds.)
    1903-28   *Linguistic Survey of India.* Vol. III, Parts 1-3, *Tibeto-Burman Family.* Reprinted 1967 by Motilal Banarsidass (Delhi, Varanasi, Patna). [Vol. III was actually the sole responsibility of Sten Konow, though the whole *Survey* appeared under Grierson's name.]

GRÜSSNER, Karl-Heinz
    1979   *Arleng Alam: Die Sprache der Mikir.* Wiesbaden: Franz Steiner Verlag.

GVOZDANOVIĆ, Jadranka
    1985   *Language System and its Change: On Theory and Testability.* Trends in Linguistics Studies and Monographs, #30. Berlin, New York, Amsterdam: Mouton de Gruyter. [Sections on numerals in Kiranti and other TB languages, pp. 130-192.]

HALE, Austin (ed.)
    1973   *Clause, Sentence, and Discourse Patterns in Selected Languages of Nepal.* 4 vols. Publication of the Summer Institute of Linguistics. Kathmandu: University Press, Tribhuvan University.

HAMILTON, R.C.
    1900   *An Outline Grammar of the Dafla Language, as Spoken by the Tribes Immediately South of the Apa Tanang Country.* Shillong.

HANSON, Ola
    1906   *A Dictionary of the Kachin Language.* Rangoon. Reprinted (1954) by Baptist Board of Publications.

HE Jiren and JIANG Zhuyi
    1985   *[Outline Grammar of the Naxi Language.]* In Chinese. Beijing: People's Publishing Co.

HEIMBACH, Ernest E.
    1969   *White Meo-English Dictionary.* Data Paper #75, Southeast Asia Program, Cornell University. Ithaca: Cornell University.

HENDERSON, Eugénie J.A.
    1965   *Tiddim Chin: A Descriptive Analysis of Two Texts.* London: Oxford University Press.
    1986   Some Hitherto Unpublished Material on Northern (Megyaw) Chin. In John McCoy and Timothy Light (eds.), *Contributions to Sino-Tibetan Studies,* pp. 101-134.
    [in prep.]   *Bwe Karen Dictionary and Texts.* Edited by Anna J. Allott, David B. Solnit, and James A. Matisoff. London: School of Oriental and African Studies.

HODGSON, Brian H.
    1880   *Miscellaneous Essays Relating to Indian Subjects.* 2 vols. London: Trübner. [Data on Chepang numerals, Vol. I, pp. 166-167.]

HU Tan and DAI Qingxia
    1964   [Plain and Constricted Vowels in the Hani Language.] In Chinese. *Zhongguo Yuwen* 128: 76-87.

HUFFMAN, Franklin E.
    1970   *Modern Spoken Cambodian.* New Haven and London: Yale University Press.

JASANOFF, Jay H.
　　1994　Italo-Celtic: The Second Coming. Paper presented at Fifth Spring Workshop on Theory and Method in Historical Reconstruction, University of Pittsburgh (April 10).

JÄSCHKE, H.A.
　　1881　*A Tibetan-English Dictionary*. London. Reprinted (1958) by Routledge and Kegan Paul Ltd.

JONES, Robert B.
　　1961　*Karen Linguistic Studies: Description, Comparison, and Texts*. University of California Publications in Linguistics, #25. Berkeley and Los Angeles.

JOSHI, Tika Ram
　　1909　*A Grammar and Dictionary of Kanawari*. Ed. by H.A. Rose. Journal of the Asiatic Society of Bengal (n.s.), Vol. V, extra number.

JUDSON, Adoniram
　　1953　*Judson's Burmese-English Dictionary*. Centenary edition, revised and enlarged by R.C. Stevenson and F.H. Eveleth. Rangoon: Baptist Board of Publications. 2nd Printing 1966 (1123 pp.).

JUI I-fu
　　1948　[Notes on the Sounds of the Lisu Language, with Remarks on the Lisu Script.] In Chinese. *Bulletin of the Institute of History and Philology* 17: 303-327.

KARAPURKAR, Pushpa Pai [Ms.]
　　1976　*Kokborok (Tripuri) Grammar*. Central Institute of Indian Languages Grammar Series, #3. Mysore. 144 pp.

KARLGREN, Bernhard
　　1957　*Grammata Serica Recensa*. Stockholm: BMFEA 29.1: 1-332. ("GSR")

KITAMURA, Hajime, Tatsuo NISHIDA and Yasuhiko NAGANO (eds.)
　　1994　*Current Issues in Sino-Tibetan Linguistics*. Osaka: The Organizing Committee of the 26th International Conference on Sino-Tibetan Languages and Linguistics.

KNOWLTON, Edgar C. Jr.
　　1976　Serial Enumeration in Malayan Semai. In Gérard Diffloth and Norman Zide (eds.), p. 99.

KOEFOED, H.A.
　　1958　*Teach Yourself Danish*. London: English Universities Press Ltd.

LEWIS, Paul
　　1968　*Akha-English Dictionary*. Data Paper #70, Southeast Asia Program, Cornell University. Ithaca, NY.

Lo Ch'ang-p'ei
　　1945　A Preliminary Study of the Trung Language of Kung Shan. *HJAS* 8.3/4: 343-348.

LORRAIN, J. Herbert
　　1907　*A Dictionary of the Abor-Miri Language*. Shillong.

LORRAIN, Reginald A.
　　1951　*Grammar and Dictionary of the Lakher or Mara Language*. Shillong.

LU Shaozun
　　1983　*[An Outline Grammar of the Pumi Language.]* In Chinese. Beijing: People's Publishing Co.
　　1986　*[An Outline Grammar of the Cuona Menba Language.]* In Chinese. Beijing: People's Publishing Co.

LUCE, G.H.
　　1985　*Phases of Pre-Pagán Burma: Languages and History*. 2 vols. Oxford and New York: Oxford University Press.

MA Xueliang
　　1949　*[Annotated Translation of the Lolo Classic* Rites, Cures, and Sacrifices: *A Text in the*

Luquan Dialect of Lolo.] In Chinese.   Bulletin of the Institute of History and Philology, #20.   Beijing.

1951   *[Studies in the Sani Yi Language.]* In Chinese.   Chinese Academy of Sciences Linguistics Monograph Series, #2.   Shanghai.

MAINWARING, G.B. and A. GRÜNWEDEL

1898   *Dictionary of the Lepcha Language.* Berlin.

MARAN, LaRaw

[in prep.]   *The Jingpho Dictionary: A Dictionary of Modern Jingpho.*   [Revised and enlarged version of Hanson 1906, with tones indicated.] 1441 pp. MS.

MARCOS-MARÍN, Francisco

1993   Numerals and Typology in Minority Languages of China. *Proceedings of the XVth International Congress of Linguists,* Vol. IV: 169-172.

MARRISON, Geoffrey E.

1967   *The Classification of the Naga Languages of North East India.*   Vol. I, 292 pp.   Vol. II, 460 pp.   Doctoral dissertation, School of Oriental and African Studies, London.

MATISOFF, James A.

1972a   *The Loloish Tonal Split Revisited.*   Research Monograph, #7.   Berkeley: Center for South and Southeast Asia Studies, University of California.

1972b   Tangkhul Naga and Comparative Tibeto-Burman. *Tonan Azia Kenkyu* (Kyoto) 10.2: 1-13.

1973a   *The Grammar of Lahu.*   University of California Publications in Linguistics, #75. Berkeley and Los Angeles: University of California Press.   Reprinted 1982.

1973b   Tonogenesis in Southeast Asia.   In Larry M. Hyman (ed.), *Consonant Types and Tone,* Southern California Occasional Papers in Linguistics, #1, pp. 71-95. Los Angeles.

1975   Rhinoglottophilia: The Mysterious Connection between Nasality and Glottality. In C.A. Ferguson, L.M. Hyman, and J.J. Ohala (eds.), *Nasálfest: Papers from a Symposium on Nasals and Nasalization,* Stanford, CA: Language Universals Project, Department of Linguistics, Stanford University, pp. 265-287.

1978a   *Variational Semantics in Tibeto-Burman: The 'Organic' Approach to Linguistic Comparison.*   Philadelphia: Institute for the Study of Human Issues. ("ISHI")

1978b   Mpi and Lolo-Burmese Microlinguistics. *Monumenta Serindica* (Tokyo) 4:1-36.

1979   Problems and Progress in Lolo-Burmese: Quo Vadimus? *LTBA* 4.2: 11-43.

1980   Stars, Moon, and Spirits: Bright Beings of the Night in Sino-Tibetan.   *Gengo Kenkyu* (Tokyo) 77:1-45.

1985a   Out on a Limb: *Arm, Hand,* and *Wing* in Sino-Tibetan.   In G. Thurgood, J.A. Matisoff, and D. Bradley (eds.), pp. 421-450.

1985b   God and the Sino-Tibetan Copula, with Some Good News Concerning Selected Tibeto-Burman Rhymes.   *Journal of Asian and African Studies* (Tokyo Foreign Languages University) 29: 1-81.

1986   The Languages and Dialects of Tibeto-Burman: An Alphabetic/Genetic Listing, with Some Prefatory Remarks on Ethnonymic and Glossonymic Complications.   In John McCoy and Timothy Light (eds.), pp. 3-75.

1988a   *The Dictionary of Lahu.*   University of California Publications in Linguistics, #111. Berkeley, Los Angeles, London: University of California Press.

1988b   Universal Semantics and Allofamic Identification — Two Sino-Tibetan Case Studies: STRAIGHT/FLAT/FULL and PROPERTY/LIVESTOCK/TALENT.   In Akihiro Sato (ed.), *Languages and History in East Asia: Festschrift for Tatsuo Nishida on the Occasion of His 60th Birthday,* Kyoto: Shokado, pp. 3-14.

1989   Palatal Suffixes in Sino-Tibetan.   Paper originally presented at 22nd International Conference on Sino-Tibetan Languages and Linguistics, University of Hawaii.   Second

draft MS, 63 pp. To appear (1995).

1990a On Megalocomparison. *Language* **66**.1: 106-120.

1990b The Dinguist's Dilemma: l/d Interchange in Sino-Tibetan. Paper presented at the 23rd International Conference on Sino-Tibetan Languages and Linguistics, University of Texas (Arlington).

1991 Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology* 20:469-504.

1994a Regularity and Variation in Sino-Tibetan. In Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds.), pp. 36-58.

1994b Sangkong of Yunnan: Secondary 'Verb pronominalization' in Southern Loloish. In Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano (eds.), pp. 588-607.

1995a [to appear] Watch Out for Number *One*: Jingpho ŋāi 'I' and ləŋâi 'One'. *LTBA* 17.1.

1995b [to appear] Sino-Tibetan Palatal Suffixes Revisited. In Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano (eds.), *New Horizons in Tibeto-Burman Morphosyntax*, Senri Ethnological Studies No. 41. Osaka: National Museum of Ethnology.

[in prep.] *Languages of Mainland Southeast Asia.* Cambridge University Press.

MAZAUDON, Martine

1985 Dzongkha Number Systems. In Suriya Ratanakul et al. (eds.), *Southeast Asian Linguistic Studies Presented to A-G. Haudricourt* (Mahidol University, Bangkok), pp. 124-157.

McCoy, John and Timothy LIGHT (eds.)

1986 *Contributions to Sino-Tibetan Studies.* Leiden: E.J. Brill.

MEGU, Arak

1985 Numerals of the Adis: A Comparative Study. *Resarun* 11.2: 75-77.

MEI Tsu-lin and Jerry NORMAN

ca.1968 The Numeral SIX in Old Chinese. *Unicorn* 2: 22-32.

MENNINGER, Karl

1969 *Number Words and Number Symbols: A Cultural History of Numbers.* Cambridge, Mass.: MIT Press. Translated from the revised German edition (1958).

MICHAILOVSKY, Boyd

1981 *La langue Hayu: phonologie, morphologie, syntaxe.* Thèse de Doctorat de Troisième Cycle, University of Paris III; also accepted as Ph.D. dissertation, University of California, Berkeley.

MICHAILOVSKY, Boyd and Martine MAZAUDON

1992 Preliminary Notes on the Languages of the Bumthang Group (Bhutan). Paper presented at the 25th International Conference on Sino-Tibetan Languages and Linguistics, University of California, Berkeley. To appear in Proceedings of the 6th Conference of the International Association of Tibetan Studies (Fagernes, Norway).

MOMIN, K.W.

n.d. *English-Achikku [Garò] Dictionary.* Calcutta. 270 pp.

NAGALAND BHASHA PARISHAD [Linguistic Circle of Nagaland]

1972a *Hindi-Chakhesang-English Dictionary (Chokri Dialect).* Kohima. 138 pp.

1972b *Hindi-Pochury-English Dictionary.* Kohima. 100 pp.

1973 *Hindi-Zeliang-English Dictionary (Zemi Dialect).* Kohima. 95 pp.

1974 *Hindi-Kheja-English Dictionary.* Kohima. 199 pp.

NAGANO, Yasuhiko

1979 A Historical Study of rGyarong Rhymes. *LTBA* 5.1: 37-47.

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

NHKUM Seng Hkum et al.

1981 *A Chinese-Jinghpo Dictionary* [Hàn Jǐng Cídiǎn = *Miwa-Jinghpo Ga Ginsi Chyum*]. Kunming: The Yunnan Nationalities Publishing House.

NISHI, Yoshio
    1982   Swadesh 100-Word Lists for Some Languages of the Monpa Group. 6 pp. MS.
NISHI, Yoshio, James A. MATISOFF and Yasuhiko NAGANO (eds.)
    1995   [to appear] *New Horizons in Tibeto Burman Morphosyntax.* Senri Ethnological
            Studies No. 41. Osaka: National Museum of Ethnology.
NISHIDA, Tatsuo
    1964   [Burmese and the Lolo Languages: A Comparative Study of Their Tone-Systems.] In
            Japanese. *Tonan Azia Kenkyu* 4:13-40.
    1966/1967 [A Comparative Study of the Bisu, Akha, and Burmese Languages.] In Japanese.
            Part I, *Tonan Azia Kenkyu* 4.3: 42-68 (Dec. 1966); Part II, *TAK* 4.5: 52-68 (March
            1967).
ONO, Toru
    1965   [The Reconstruction of Common Kuki-Chin: (I) Word-Initial Consonants.] In
            Japanese. *Gengo Kenkyu* 47:8-20.
OSBURNE, Andrea G.
    1975   *Transformational Analysis of Tone in the Verb System of Zahao (Laizo) Chin.* Ph.D.
            dissertation, Cornell University.
PHILLIPS, E.G.
    1904   *Outline Grammar of the Garo Language.* Shillong.
RAI, Agam Singh Dewsa
    1944   *Thulung Rāī Bhāsā.* [The Thulung Rai Language.] In Nepali. Darjeeling.
ROBINSON, W.
    1851   Notes on the Dophlás and the Peculiarities of Their Language. *JASB* 20:126-137.
ROCK, J.F.C.
    1963   *A Na-Khi-English Encyclopedic Dictionary, Part I.* Serie Orientale Roma #28, Istituto
            Italiano per il Medio ed Estremo Oriente. Rome.
ROOP, DeLagnel Haigh
    1970   *A Grammar of the Lisu Language.* Ph.D. dissertation, Yale University. 312 pp.
SANDBERG, Graham
    1895   *Manual of the Sikkim Bhutia Language or Denjongke.* Second and enlarged edition.
            Westminster. 144 pp.
SCHMIDT, Ruth L.
    1993   *A Practical Dictionary of Modern Nepali.* Delhi: Ratna Sagar P. Ltd.
SHARMA, S.R.
    1990   Numeral and System. In Shree Krishan (ed.), *Linguistic Traits across Language Boun-*
            daries, Report of All-India Linguistic Traits Survey, Calcutta: Anthropological Survey
            of India, pp. 34-50.
SIMON, I.M.
    1972   *An Introduction to Apatani.* Gangtok, Sikkim: Philology Section, Department of
            Research, Arunachal Pradesh.
    1976   *Hill Miri Language Guide.* Published by Director of Information and Public Rela-
            tions, Arunachal Pradesh.
    1979   *Miji Language Guide.* Gangtok, Sikkim: Philology Section, Department of Research,
            Arunachal Pradesh.
SOLNIT, David B.
    1979   Proto-Tibeto-Burman *r in Tiddim Chin and Lushai. *LTBA* 4.2: 111-121.
    1984   Kayah Numerals. 9 pp. MS.
    [in prep.] *A Grammar of Eastern Kayah.* (Will include a collection of texts and a glossary.)
SRINUAN Duangkhom
    1976   *An Mpi Dictionary.* Edited by Woranoot Pantupong. Working Papers in Phonetics
            and Phonology, Vol. I, #1, Indigenous Languages of Thailand Research Project.

Bangkok.

SUN Hongkai

1981 *[An Outline Grammar of the Qiang Language.]* In Chinese. Beijing: People's Publishing Co.

1982a [A Brief Introduction to the Ersu (Tosu) Language.] In Chinese. *Yuyan Yanjiu* 2:241-264.

1982b *[An Outline Grammar of the Dulong Language.]* In Chinese. Beijing: People's Publishing Co.

1983 *[Outline of the Idu Luoba Language.]* In Chinese. *Minzu Yuwen* 6:63-79.

1985 [Peoples and Languages of the Six River Valleys and Their Genetic Classification.] In Chinese. *Minzu Xuebao* 3:99-273.

1986 Notes on Tibeto-Burman Consonant Clusters. *LTBA* 9.1: 1-2.

SUN Hongkai and LIU Lu

1986 *[An Outline Grammar of the Language of the Nu People (Nusu Dialect).]* In Chinese. Beijing: People's Publishing Co.

SUN Hongkai, LU Shaozun, and OUYANG Jueya

1980 *[The Monpa, Lhopa, and Deng Languages.]* In Chinese. Beijing: Chinese Social Sciences Publishing Co.

SUN, Jackson Tianshin

1992 Review of *Zang-Mianyu Yuyin he Cihui.* [Tibeto-Burman Phonology and Lexicon.] *LTBA* 15.2: 73-113.

1993 *A Historical-Comparative Study of the Tani (Mirish) Branch of Tibeto-Burman.* Ph.D. dissertation, University of California, Berkeley. 524 pp.

TALUKDAR, G.C. et al.

1962 *A Phrase Book in Idu.* Shillong: Philological Section, Research Department, Northeast Frontier Agency.

THIRUMALAI, M.S.

1972 *Thaadou Phonetic Reader.* Central Institute of Indian Languages Phonetic Reader Series, #6. Mysore.

THURGOOD, G., J.A. MATISOFF, and D. BRADLEY (eds.)

1985 *Linguistics of The Sino-Tibetan Area: The State of the Art.* Pacific Linguistics Series C-87. Canberra: Australian National University.

TOBA, Sueyoshi and Ingrid TOBA

1975 *A Khaling-English and English-Khaling Glossary.* Kathmandu: Tribhuvan University Press.

TURNER, Ralph L.

1931 *A Comparative and Etymological Dictionary of the Nepali Language.* London: Routledge and Kegan Paul Ltd.

1966 *A Comparative Dictionary of the Indo-Aryan Languages.* London: Oxford University Press.

WEIDERT, Alfons

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

WEN Yu

1950 [An Abridged Ch'iang Vocabulary (Chiu Tzu Ying Dialect).] In Chinese. *Studia Serica* 9:17-54.

WIERSMA, Grace

1990 *Investigation of the Bai (Minjia) Language along Historical Lines.* Ph.D. dissertation, University of California, Berkeley.

WU Zili

1993 [A Preliminary Study of the Mo'ang Speech in Funing County, Yunnan.] In Chinese. *Minzu Yuwen* 2: 53-63.

XU Xijian et al.
    1983   *A Jinghpo-Chinese Dictionary* [Jǐng Hàn Cídiǎn = *Jinghpo Miwa Ga Ginsi Chyum*].
            Kunming: The Yunnan Nationalities Publishing House.
XU Xijian and XU Guizhen
    1984   *[An Outline Grammar of the Zaiwa Language of the Jingpho Nationality.]* In Chinese.
            Beijing: People's Publishing Co.
YABU, Shiro
    1982   *[A Classified Dictionary of the Atsi or Zaiwa Language (Sadon Dialect), with Atsi,
            Japanese, and English Indexes.]* In Japanese. Tokyo: Institute for the Languages and
            Cultures of Asia and Africa, Tokyo Foreign Languages University.
YUAN Jiahua
    1947   *[Preliminary Investigation of the Woni Language of Er-shan.]* In Chinese. Nankai
            State University, Frontier Peoples' Culture Department, Publication, #4. Tianjin.
    1953   *[The Folksongs and Language of the Ahi People.]* In Chinese. Chinese Academy of
            Sciences, Linguistics Research Institute Monograph Series, #5. Beijing.
ZHAO Yansun
    1982   [The Problem of the Genetic Affiliation of the Bai Language.] In Chinese. In *Minzu
            Yuwen Yanjiu Wenji*, Xining: Qinghai People's Publishing Co., pp. 150-188.

シナ・チベット諸語の数詞と前接辞の役割

J. A. マティソフ

　チベット・ビルマ共通祖語における数詞の再構成形式はかなり安定しているが，そこでの前接辞の扱いには微妙な差異がある。本論文で私は幾つかの新しい数詞の語根要素と単語族の括り方を指摘した。既に検証されている変異様式に対しても，それらの新再構成形式などによって，実証性を持たせることに成功した。私は数詞相互の影響に特に留意し，三つの段階を踏んで検討した。すなわち，まずシナ・チベット諸語の単語族に一般的に認められる変異様式のコンテキストでの検討（1.11），次に，より個別的な形態音韻論的・意味論的観点からの検討（4.01-4.02），最後に数詞のセットにおける前接辞の役割に焦点をあてた検討，である。我々は，幾つかの言語がどのように「足す」，「引く」，「掛ける」を用いてより高次の基本数（6から9まで）を表現するかを解明した。

　シナ・チベット祖語段階での個々の基本数（1から9まで）の再構成形式はここでは取り扱わなかったが，様々の言語が用いている「10＋…」や10進法／20進法に見られるような体系的構造に特に注意した（3.3-3.5）。我々は例えば1と10の間に見られるようなチベット・ビルマ系の数詞体系における著しい揺れ，融合，及び再評価を見てきた（3.4）。このような現象は言語接触の強い影響下にあるヒマラヤ諸語に顕著に観察される（2.0-2.1）。これには，シェルパ語，レプチャ語，ゾンカ語に見られる10と20の間の揺れ（3.533）やチェパン語に見られる20と12の間の揺れが含まれる（3.535）。我々はまたボロ語，クビンデ語，ドゥミ語などの4を基盤とする体系（3.32C）や，バンタワ語の5を基盤とする体系（4.15）の痕跡に注目した。これらはチベット・ビルマ語族の数組織の古い類型を代表している蓋然性があるからである。

　小稿はカリフォルニア大学における筆者の長期研究プロジェクト，Sino-Tibetan Etymological Dictionary and Thesaurus, の成果の一部である。このプロジェクトは意味領域にしたがってチベット・ビルマ祖語ないしシナ・チベット祖語を再構成し，意味と音韻共に豊富な太古の語彙体系を再生させようとする試みである。