Sino-Tibetan numeral systems: prefixes, protoforms and problems

PACIFIC LINGUISTICS

Pacific Linguistics is a publisher specialising in linguistic descriptions, dictionaries, atlases and other material on languages of the Pacific, the Philippines, Indonesia and southeast Asia. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world.

Pacific Linguistics is associated with the Research School of Pacific and Asian Studies at the Australian National University. Pacific Linguistics was established in 1963 through an initial grant from the Hunter Douglas Fund. It is a non-profit-making body financed largely from the sales of its books to libraries and individuals throughout the world, with some assistance from the School.

The Editorial Board of Pacific Linguistics is made up of the academic staff of the School's Department of Linguistics. The Board also appoints a body of editorial advisors drawn from the international community of linguists. Publications in Series A, B and C and textbooks in Series D are refereed by scholars with relevant expertise who are normally not members of the editorial board.

To date Pacific Linguistics has published over 400 volumes in four series:

- **Series A:** **Occasional Papers**; collections of shorter papers, usually on a single topic or area.

- **Series B:** **Monographs** of intermediate length.

- **Series C:** **Books**; publications of greater length, especially reference books such as dictionaries and grammars, and conference proceedings.

- **Series D:** **Special Publications**; including archival materials, pedagogical works, maps, audiovisual productions, and materials that do not fit into the other series.

SINO-TIBETAN NUMERAL SYSTEMS: PREFIXES, PROTOFORMS AND PROBLEMS

James A. Matisoff

TABLE OF CONTENTS

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 (Beijing; 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 1903-28)
LTBA	Linguistics of the Tibeto-Burman area (Berkeley)
MC	Middle Chinese (= Karlgren's 'Ancient Chinese')
NBP	Nagaland Bhasha Parishad (Linguistic Circle of Nagaland, Kohima)
OC	Old Chinese (= Karlgren's 'Archaic Chinese')
PGmc	Proto Germanic
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/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)

PREFACE

This study was originally prepared for the Seventeenth 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 about a decade before being published in 1995 as 'Sino-Tibetan numerals and the play of prefixes' in the Bulletin of the National Museum of Ethnology (Osaka) 20/1:105-252. It is here reprinted with some revisions with the persmission of the Bulletin. 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 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 at Appendix 3.

A fine-tuned subgrouping of the TB languages is an essential long-term 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 protolanguage. 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, there are vast differences,

---

[1] I would like to thank Nicholas C. Bodman and David Bradley for their cogent comments on the first version of this monograph. My thanks also to Chang Kun for the numerals of the Zida (Tzuta) dialect of Rgyalrong; Gérard Diffloth for information on numerals in Mon-Khmer languages; Robert Goldman for helping me read a Hindi source on Pochury; David Solnit for data on Kayah Karen; Purna C. Thoudam, for sending me over a dozen translations of the Bible into TB languages of India; Graham Thurgood for forms from Idu and Rawang; Chhewang Rinzin for making clear tape recordings of the numerals in Sharchop and Dzongkha; and Gabrielle Yablonsky for putting me in touch with Mr. Rinzin. For more recent helpful comments and practical assistance I am much obliged to Lon Diehl, 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.

[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 has appeared as Volume 2 of the STEDT Monograph Series (Matisoff 1996).

[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 recognise the Qiangic languages as a separate subgroup at all, not surprisingly since most of them have only recently been put into relief by Chinese scholars. I have suggested the term 'Kamarupan' (from Kāmarūpa, 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,

both qualitative and quantitative, in the data available for particular TB languages or language-groups.
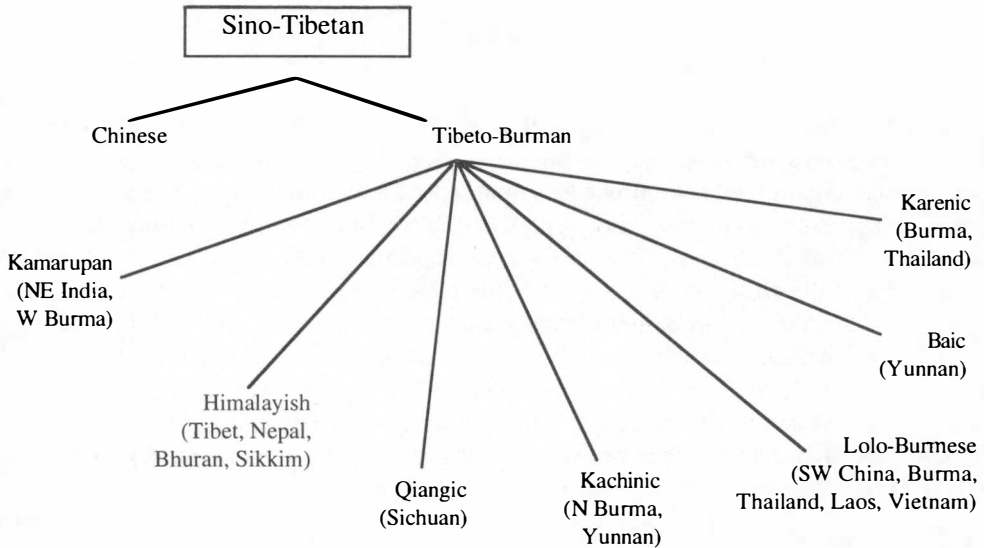
FIGURE 1: THE BRANCHES OF TIBETO-BURMAN

For the Naga languages we are still heavily dependent on G.E. Marrison (1967) ('GEM'), 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 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.

---

and is the 'centre 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.

[4] For full references to all works mentioned see the reference list.

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 monograph was written, it will be discussed in appropriate contexts below (§2.1, §3.5.3 and §4.0.2). 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 monograph was composed (see §3.5.3.4).

Reliable data on the Qiangic languages of Sichuan is now becoming available in quantity, thanks to the efforts of scholars like Sun Hongkai (for example, 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 preparation) and two excellent dictionaries produced by Dai Qingxia, Xu Xijian, et al. (Chinese/Jingpho 1981; Jingpho/Chinese 1983). For Nungish, older sources like Barnard (1934) and Lo Ch'ang-p'ei (1942) are now vastly enriched by Sun Hongkai 1982 (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 (for example, 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 (1966/67) (Bisu); Srinuan (1976) (Mpi); Rock (1963), Chen Weidong (1990) (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).

---

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.

CHAPTER 1

INTRODUCTION

1.1 ISSUES IN THE RECONSTRUCTION AND SYSTEMATIC BEHAVIOUR OF THE TIBETO-BURMAN NUMERALS

In a sense this monograph 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[1] as in Figure 2, raises a variety of interesting morphophonemic and lexico-semantic issues.

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-lə y[2] (#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-kə w ~ *d-ga w (#13)
TEN	*gip (#16); *ts(y)i(y) ~ *tsyay (#408)
TWENTY	*m-kul (#397)
HUNDRED	*r-gya (#164)
THOUSAND	*s-toŋ (#32)

FIGURE 2: PROTO TIBETO-BURMAN NUMERALS

1.1.1 PROTO-VARIATION

Proto-variation must be recognised 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

---

[1] See especially STC, §16, pp.93-95. References preceded by '#' refer to the numbered cognate sets in STC; other references are to page numbers.

[2] 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, 20-21), hence the 'equals' signs in the chart. Quite distinct from the above are cases where Benedict (explicitly or implicitly) recognises phonological variation at the proto-stage. These are marked with a tilde in the chart. The case of TEN poses a special problem (see below).

1

around a basic phonological shape and a core of meaning. Variability is observable in all parts of the TB syllable: rhymes, initials, prefixes, tones.[3] To some extent STC is prepared to recognise cases of proto-variation, and its labyrinthine pages contain many more 'allofamic reconstructions' than is at first apparent. An examination of Figure 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.

1.1.1.1 VARIATION OF PROTO-RHYME

The STC recognises a PTB alternation *-uw ~ *-aw in NINE, on slender evidence, rejecting as secondary the better-attested variant in -wa (Lushai pakua, Angami thepfə).[4] For TEN, the first version of STC recognised 'vowel gradation' between *-ai and *-i to account for Written Burmese (WB) ʔəchai versus, for example, Jg. shi. This view was later changed (note 272) in favour of introducing a complication into the reconstruction of the initial consonant.[5]

TWO furnishes an example of variation of final consonant. Alongside the principal allofam in -s, a variant in *-k (underlying, for example, WB hnac < *s-nik) is also attested independently in several branches of TB, but is denied PTB status in STC. (See §4.1.1.)

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- and *-ya-, or between the rhymes *-ay and *-an.[6]

1.1.1.2 VARIATION OF PROTO-PREFIX

All the numerals from 2 to 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 recognises 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.

---

3   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 'x' is there introduced to stand for the allofamic relationship: X x Y 'X and Y are co-allofams; X and Y both belong to the same word-family'.

4   See the discussion in Matisoff (1980:15-17), and §4.2.4.

5   Matisoff (1985b:5, 32), and §3.2.2.

6   See §3.1.1, §3.1.2, §3.1.4, §3.2.1, §3.2.2, §3.2.3.3, etc.

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 recognise (Garo *sku*, Kanauri *zgui*, Jingpho *džə khû*, Pumi *sgiuh*). As for SIX, the initial velars in Himalayish, Jingpho and Lolo-Burmese (for example Magari *kruk*, Jg. *krú?*, WB *khrok*) are treated as secondary developments from the dental prefix before root-initial *r-, that is *d-r > kr. Yet tonal developments in Loloish, where the word appears in the low-stopped tone despite its voiceless initial (for example, Lahu *khɔ̀?*), clearly point to a doubly-prefixed prototype *d-k-rok.[7] This syntagmatic view of the prefixal dynamics in SIX makes its behaviour 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' (§5.2ff.), whereby consecutive numerals acquire the same prefix. It is undeniable that many modern TB 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 to 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*[8]

Yet STC has to recognise 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-ŋa.[9] 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.

1.1.2 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 (such as additive and subtractive, multiplicative) 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

---

7   See Matisoff (1972a 35, 71), *The Loloish tonal split revisited*.
8   Lushai *pasarih* '7' is a doubly prefixed form, with the younger *pa- attached before the older *sa-* (< PTB *s-nis). Contra Matisoff (1980:16-17), the Lushai form *pakua* '9' furnishes no support for a PTB *b- prefix with this numeral.
9   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 §5.4.4.

even suffixes) of consecutive numerals (§4.0.1). 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 mathematical relationship between '8' and '100'.[10]

Although wholesale borrowing of numerals is by no means unheard of,[11] and has reached critical proportions in many of the TB languages of Nepal (§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.[12]

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.[13] 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.[14] 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.[15]

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 x 2'. (See §4.2.0.)[16] 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 (§4.1.1, §4.2.2), but it seems likely that the TB numeral system once related them conceptually.

---

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

11   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.

12   See the 'profiles' of the primary numerals at §4.1–§4.2.4. 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' (§4.2.2). The AMD branch of TB (§1.2.5) has the most aberrant-looking numeral sets in the whole family.

13   Besides the three roots for ONE and the two roots for TEN that are reconstructed in STC (see Figure 2), several additional etyma have been discovered for both. See §3.1.5, §3.2.3.

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

15   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 1994c I identify the former with the Jg. first person pronoun *ŋāi* 'I'.

16   Subtractive formations are also occasionally encountered in higher numerals, e.g. EIGHT expressed as '9–1'.

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 monomorphemic form for *twenty*, **m-kul*, is reconstructible for PTB (see Figure 2), and a number of modern languages have thoroughgoing VIGESIMAL systems of 'round-number formation' (§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* (§3.5.1, §3.5.3.4). 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* (§3.5.3.5).

Actually the phenomenon of numeral transvaluation is surprisingly widespread in TB, a testimony to the multiple simultaneous conceptual interconnections among the numbers themselves (§4.0.2).

In sum, this monograph 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 (§1.2.4, §5.4, §5.5). On the conceptual side, the Himalayish languages are of particular interest, especially because of the hesitation between decimality and vigesimality in their higher numerals (§1.2.3, §3.5.3). Yet all the subgroups of the family have their characteristic numerological flavour, and it is worthwhile to do a quick rundown of the various branches, giving a representative specimen of the kinds of numeral systems to be found in each.

1.2.1 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).

	Written Burmese	Lahu (C. Loloish)
ONE	*tac*	*tê*
TWO	*hnac*	*nî*
THREE	*sûm*	*šɛ̂ʔ* ⚹ *šɛ̄*[17]
FOUR	*lê*	*ɔ̂*
FIVE	*ŋâ*	*ŋâ*
SIX	*khrok*	*khɔ̂ʔ*
SEVEN	*khu'-hnac*	*šɨ̄*
EIGHT	*hrac*	*hí*
NINE	*kûi*	*qɔ̂*
TEN	*tə-chay*[18]	*tê-chi*
TWENTY	*hnə-chay*	*nî-chi*

FIGURE 3: SOME LOLO-BURMESE NUMERALS

Another route by which a prefix could survive was by 'pre-empting' or driving out a weak (non-obstruental) root-initial[19], 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.[20]

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.[21] Compare 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 (pers.comm. 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').[22] 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 (for example, *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').

---

17   *šɛ̂ʔ* 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* (§3.3).

18   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, that is '10' is expressed as 'one tenworth', '20' as 'two tensworth', and so on. See §3.3.2(B), and the Kayah form for ONE (Figure 4).

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

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

21   Karen dialects mentioned as having such composite numerals in STC (p.130) include "White Karen, Bwe, Brek, Red Karen, Yintale, and Mano".

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

	Sgaw (Jones 1961)	E. Kayah (Solnit 1984)
ONE	tə̄	tə-
TWO	khī̄[23]	nʌ̄
THREE	θə̄	sō
FOUR	lwì	lwī
FIVE	jὲ	ŋē ~ ñē
SIX	xỹ	sō swá?
SEVEN	nwī	sō swá? tə-
EIGHT	xɔ̀?	lwī swá?
NINE	khwī	lwī swá? tə-
TEN	ʃĭ	chʌ́ ~ chʌ̄[24]

FIGURE 4: SOME KAREN NUMERALS

Other Karen dialects, especially Pa-O (Taungthu) have developed secondary dental suffixes with certain numerals: *Pa-O līt* '4', *ngāt* '5', *kūt* '9'. In the case of *nūt* '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.[25] We consider these final dentals to constitute a 'suffix run', one of the many manifestations of the interinfluence of consecutive numerals (§4.0.1). As we shall soon see (§1.2.3), numeral suffixes are also characteristic of Himalayish, but there they tend to be fully syllabic.

1.2.2 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, *ŋ-*. The old causative prefix **s-* has been preserved and generalised as *šə- ~ džə-*. Younger strata of prefixation are much in evidence, with a tendency to create fully syllabic prefixes out of sub-syllabic ones, for example, **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 Figure 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-*[26], 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'.

---

23   Note the preemption of the root-initial by the velar prefix, **g-nis > khī̄.*
24   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, for example, *chʌ́ sō swá* '16', i.e. 10 + (3 x 2), where TEN is in an *additive* relationship to the following numeral, with *chʌ̄ sō swá* '60', i.e. 10 x (3 x 2), where TEN stands in a *multiplicative* relationship with it.
25   See STC, p.131, and Benedict (1979:18-20). For more discussion see §4.2.2.3.
26   The term *sesquisyllabic*, referring to words 'a syllable and a half long', was introduced in Matisoff (1973b).

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).

	Jingpho	Nusu	Ergong
ONE	ləŋâi	thi^{53}	ẓau
TWO	ləkhôŋ	m̩^{55}	wnɛ
THREE	məsūm	sɔ^{35}	wsu
FOUR	məlī	vɹi^{35}	wʑɛ
FIVE	məŋā	ŋa^{55}	wŋuɛ
SIX	krúʔ	khɹu^{53}	wtɕhau
SEVEN	sənìt	ŋə̄$^{.55}$	sŋie
EIGHT	mətsát	ṣa$^{.53}$	ɣiɛ
NINE	jəkhû	gɯ^{35}	ŋgɛ
TEN	šī	tshe^{35}	zʁa/sqha

FIGURE 5: KACHIN-NUNG AND QIANGIC NUMERALS

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 to 5. As mentioned above (footnote 15), the Jingpho forms for ONE and TWO are innovations which require a special explanation.

1.2.3 HIMALAYISH AND RGYALRONG

Himalayish shows fairly good preservation of the proto-prefixes, but by 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 (§3.5.3.4). 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 (§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, for example, Tamang (Nepal) or Kanauri (Simla Hill States, Punjab), which only show prefixes with a few of the 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

FIGURE 6A: SOME HIMALAYISH NUMERALS

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.[27] Even more exuberant in this respect is Rgyalrong (= 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 Rgyalrong is not 'core Himalayish' at all, but rather a transitional language, with suggested affinities to Abor-Miri-Dafla (see Nagano 1984),[28] and/or Qiangic (as maintained by Sun Hongkai 1985).[29] See Figure 6b.

	Lepcha	Rgyalrong (Zida dialect)
ONE	*kat*	*tšek*
TWO	*ɲət*	*kenes*
THREE	*sam*	*kesom* � *kesam*
FOUR	*fəli*	*kewdži*
FIVE	*fəŋo*	*kemɲa*
SIX	*tərək*	*keʈa*
SEVEN	*kəkyək*	*keʂñit* ⋈ *keʂñis* ⋈ *keʂñes*
EIGHT	*kəku*	*warže(t)*
NINE	*kəkyot*	*keŋgu*
TEN	*kəti*	*ʂtʂi*

FIGURE 6B: HIMALAYISH INNOVATORS: LEPCHA AND RGYALRONG

It is characteristic of many languages of Nepal to have *suffixes* attached to their numerals, for example, Dumi *-po* (*tūk-po* '1', *sak-po* '2', *suk-po* '3'), Bantawa (*ūk-pok* -pok '1', *hūa-pok* '2', *sum-ka-pok* '3'), Yakkhaba *-ci/-ji* (*nic-ci* '2', *sum-ji* '3', *ri-ji* '4'), etc. (Gvozdanović 1985:135–136). These suffixes are fully syllabic (unlike those of Pa-O Karen, §1.2.1), so one may surmise they are (or once were) classifiers, or even gender markers, rather than meaningless formatives. See §2.1.

1.2.4 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 behaviour will be discussed in the most detail (§5.4). Like Himalayish, Kuki-Chin-Naga shows a complex interplay of decimal and vigesimal characteristics (§3.5.2).

Bodo-Garo (= Barish) displays occasional cases of reprefixation (for example, Garo *ge-gni* '2'), but in general is not so extreme in this respect as KCN, Qiang, or Rgyalrong. Boro can definitely be shown to have a *quaternary* or 4-based numeral system, very unusual for TB.[30]

---

[27] As we shall see (§4.0.2, §4.2.3, §4.2.4), these Lepcha forms for EIGHT and NINE seem to have undergone an 'etymological flipflop'.

[28] The possibility of a special AMD-Rgyalrong relationship is vigorously criticised in J.T. Sun (1993:379-389).

[29] This Rgyalrong run is reminiscent of the 2-6 run of the *w-* prefix in the Qiangic language Ergong (§1.2.2), though Ergong shows only one prefix per numeral.

[30] See §3.3.2(C) 'Teen formation in Barish' and §4.2.0.1 'Multiplicative phenomena'. Elsewhere in TB, the closest thing I have found to the Boro quaternary system is the *duodecimal* system of Chepang (§3.5.3.5).

10

1.2.5 ABOR-MIRI-DAFLA

This relatively obscure branch of TB harbours 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 also the systems reveal peculiarities of internal structure (for example, 'multiplicative' forms for EIGHT: see §4.2.0, §4.2.3.7). Several new roots for ONE and TEN have been unearthed in this subgroup (§3.1.5, §3.2.3).

The numeral prefixes that appear with the highest frequency in AMD consist of a vowel alone: *a-*, *o-*, *e-*. All other prefixes (for example, *kV-*, *pV-*, *ra-*) are quite rare with AMD numerals. See §5.5.

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

	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*[31]
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*

FIGURE 7: SOME ABERRANT ABOR-MIRI-DAFLA NUMERAL SYSTEMS[32]

1.2.6 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.[33] See Figure 8.

---

[31] See §4.2.1.7.
[32] Aka/Hruso from LSI III.1, Miju from Das Gupta (1977a); Milang from Das Gupta (1980); Serdukpen from Dubey (1983).
[33] See also the section 'Chinese numerals' in STC, pp.161-162. Special studies have been devoted to the Chinese numerals ONE (Boltz 1969), TWO (Boltz 1977), and SIX (Mei and Norman 1968).

	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	*ts'əm ~ *səm (GSR #647, 648)	*g-sum
FOUR	*si̯əd[34] (GSR #518)	*b-ləy
FIVE	*ngo[35] (GSR #58)	*l-/b-ŋa
SIX	*li̯ôk[36] (GSR #1032)	*d-ruk
SEVEN	*ts'i̯ĕt (GSR #400)	*s-nis
EIGHT	*pwăt[37] (GSR #281)	*b-g-ryat/ *b-r-gyat
NINE	*ki̯ug > MC *ki̯ə w (GSR #992)	*d-kə w
TEN	*ḑi̯əp (GSR #686)	*g(y)ip
HUNDRED	*păk[38] (GSR #781)	*r-gya (but WT brgya)

FIGURE 8: OLD CHINESE AND PTB NUMERALS

---

[34] STC derives this from pre-Archaic *p-səy.

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

[36] 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").

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

[38] 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)".

LANGUAGE CONTACT AND THE WEIGHT OF NUMBERS

2.0 INTRODUCTION

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, for example, 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."[1] 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' *hɔ́k*, '7' *kjīt*, '8' *pɛt*, '9' *kāw*, '10' *sīp*.[2]

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 to 10 with Chinese ones.[3] 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.

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 (§4.0.2), foreign incursions

---

[1]   Henderson (1986:112).
[2]   See Nishida (1966/7).
[3]   In Matisoff (forthcoming), 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 preserved only in a few obscure 'outlier Kadai' languages like Li (Hainan). For the introduction of the terms 'Sinosphere' and 'Indosphere', see Matisoff (1990a and 1991).

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.

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

THOUSAND *hajaar*

FIGURE 9: NEPALI NUMERALS

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.[4] Many languages (for example, 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.

	Kham	Sunwar	Magari	Chepang
ONE	*tobo*	*ka:*	*kaʈ*[5]	*ya:t-jo?*
TWO	*nehplo*	*niksyi*	*nis*	*nis-jo?*
THREE	*sohmlo*	*sā:*	*som*	*sum-jo?*
FOUR			*puli*[6]	*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*

FIGURE 10: PRESERVED TB NUMERALS IN SOME LANGUAGES OF NEPAL[7]

---

4   See §1.1.2, §2.0. For the exceptional situation in Jingpho in this regard, see Chapter 1, note 15 and Figure 5.

5   I believe Magari *kaʈ* 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, §3.1.

6   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.

7   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

Similarly, Michailovsky (1988:123) 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 (1975:102) 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 (1985:140) attempts an explanation in terms of grammatical function, claiming that "the process of numeral decay is at each stage characterised by a language-specific cut-off point, defined by the highest numeral which is actively used in numeral constructions". This is a merely circular explanation, 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 (for example, 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.[8] 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.[9]/[10] 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, for example, 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 (1975:102-103) notes 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.

ONE	*ko*		SIX	*ru*
TWO	*nə*		SEVEN	*yet*
THREE	*sium*		EIGHT	*let*[11]
FOUR	*blə*		NINE	*gu*
FIVE	*ŋo*		TEN	*kodium*

FIGURE 11: THE MORIBUND TB-DERIVED NUMERALS OF THULUNG RAI

---

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 §4.1.5.

[8] For young children, big numbers are mysterious undifferentiated jumbles, so many 'forty-levens'es'.

[9] 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!*'

[10] 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.

[11] Note the convergence of the rhymes in SEVEN and EIGHT.

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 (1988) with the forms to be found in Hodgson's (1880) 19th century material. See Figures 12a, 12b.

(a) Michailovsky (1988:167)

ONE	*kolu*
TWO	*nakpu* (human) / *na?ung* (non-human)
THREE	*tshukpu* (human) / *tshu?ung* (non-human)
FOUR	*b(l)i?ung*

(b) Hodgson (ca. 1860, cited in LSI III/1:384).[12]

	Masculine	Feminine	'Irrational'
ONE	*kom-pu/kwong-pu*	*kwo-mi/kwong-mi*	*ko-lu*
TWO	*na:k-pu*	*na:ng-mi*	*na:-yung*
THREE	*chhuk-pu*	*chhung-mi*	*chhu-yung*
FOUR	*blik-pu*	*blig-mi*	*bli-ning*

FIGURE 12: SURVIVING TB NUMERALS IN HAYU

What Hodgson found was considerably more elaborate than the vestigial system reported by Michailovsky after 120 more years of intense pressure from Nepali. In fact, the 3-way gender distinction in Hodgson's (1880) data furnishes a possible clue as to the original function of the *suffixes* which are such a characteristic feature of Himalayish numeral systems (§1.2.3).

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 Yakkha *kolok* '1' (with *-k* suffix).

(b) *-pu/-po*. Besides Hayu *nakpu* '2', *tshukpu* '3', *blikpu* '4', cf. Kham *tobo* '1'; Sunwar *sa:hpu* '2' and *suhpu* '3'; and especially the Saptesar dialect of Dumi, which has generalised the *-po* with all the numerals from 1 to 9 (*tɨkpo* '1', *sakpo* '2' ... *ɔmpo* '8', *rekpo* '9').[13]

(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', *reṭkapok* '4').[14]

---

[12] LSI also cites forms for FIVE and SIX, which appear, however, to be multiplicative in origin (§4.2.0.1).

[13] 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 (§4.0.2) the Kubhinde words for '2' and '3' have been 'transvalued' from their original meanings of '3' and '4', respectively.

[14] Other Bantawa dialects have a suffix with retroflex *ṭ*, e.g. *ɨkṭa(k)* '1', *hɨaṭa* '2', *sumkaṭ* '3', *reṭkaṭaṭ* '4'. See Gvozdanović (1985:155).

(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 generalised with the whole set of numerals from 2 to 9 or 2 to 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 generalised with all the numerals from 2 to 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, for example, Chepang *-jo?* (Figure 10) and Hayu *-?ung*, remain a mystery in comparative terms. The old Hayu feminine suffix *-mi*, however, is relatable to a general TB root **mi(y)* 'woman, female' (for example, Lahu *yâ-mî* 'daughter', *ɔ̀-mî-ma* 'wife').

The generalisation of a particular suffix to a succession of adjacent numerals may be referred to as a 'suffix run'.[15]

---

[15]   See §1.2.1, with respect to the non-syllabic dental-suffix run in Pa-O Karen.

ONE AND TEN AND TEENS AND TWENTIES

3.1 PROFILE OF NUMBER ONE

As STC (p.94) observes, 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. (Compare 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, for example, Thai *nỳŋ* '1', *sìp* '10', but *sìp-èt* '11', *rɔ́ɔj-èt* '101'. (This *-èt*, like the rest of the Thai numerals from 2 to 9, is of Sino-Tibetan origin (§3.1.1).)

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 x 1'; §3.1.4); and *chi-kung* '10' (presumably '10 x 1'; §3.1.3(b)).

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

3.1.1 *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 *ʔjĕt (p.162).

To these I would now like to add Chepang *ya:t(-jo?)*, 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 (for example, EYE, PHEASANT[1]), and, strikingly enough, with several other numerical etyma as well, as we shall see.[2]

---

1 See Matisoff (1978a:40-41). ('VSTB')
2 Compare *tik ⪢ *tyak '1' (§3.1.4), *gip ⪢ *gyap '10' (§3.2.1), and perhaps *ring ⪢ *ryang '10' (§3.2.3.3).

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

Mehawang	*ek*-*ku*[3]
Yakkhaba	*ik*-*ko*
Khulung	*i-bum* ~ *i-bim*
Bantawa	**ʉk-** (as in **ʉk**-*ʈai*, **ʉk**-*ʈak*, **ʉk**-*pok*, **ʉk**-*ʈa*, all meaning 'ONE' in various dialects: 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).

3.1.2 **k-(y)at ✕ *k-(y)it ✕✕ *k-yan ✕ *k-(y)in*

STC (p.94) laconically sets up a PTB root **kat* on the basis of 'Lepcha *kat* and Kuki-Naga **khat*'. 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*).

3.1.2.1 **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 1983) *khege*; Chulikata *e:khe:*; 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), *a:-kin* (Robinson 1851); Taraon (*e:-*) *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 perhaps 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 §4.1.1.4.

---

3   For the second element in these apparently pleonastic Mewahang and Yakkhaba formations, see §3.1.3.

3.1.3 *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 §3.1.1 and §3.1.2. Several Kamarupan languages have two velar-initial words for ONE, one with a back vowel and the other with a front one (for example, 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-*, §3.1.5.2); Yacham-Tengsa *kha-tu* (for second syllable see §3.1.4.3); Tangkhul *akha*, *khatkha*[4]; 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 §3.4).

Himalayish

Thulung Rai *ko* '1', *ko-* '-teen' (see §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* (§3.1.1); 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 (§3.1.4)). 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' (§3.1.6).

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

Abor-Miri-Dafla

Milang *akan*; Minyong *akon* (also *atîr*, *ayirr*); Darang Deng *k'un*[55]; Apatani *kun* (non-humans), *kon* (humans); Idu *khun*[55] (Sun 1983:69).[5]

Other TB forms which seem to belong here are Bahing (Himalayish) *kong* '1', and Garo (Bodo-Garo) *chi-kung* '10' (lit. '10 x 1').

3.1.4   *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 *thit-ta*, and 'Rai' *tik-pu*).[6] To these

---

4   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* versus *kat*? Or is the -*t* a suffix (*ka-t*)?

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

6   See STC, pp.84, 94, 169 and 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

may be added Dumi *tɨk-po*, *tɔk-pu*; Lohorong *ʈhik-ko* (for the second element see §3.1.3); Athpare *ʈhik*; Limbu *lot-thik*; Dzongkha *ci*; Kaike *ʈi*; and the second element of Sikkim Bhutia *khe-chik* '20' (lit. '20 x 1').

Allofamic variations involve all parts of this etymon:

(a) A velar prefix is reconstructed on the basis of WT and other Himalayish languages (for example, Rgyalrong *kətɛk* (ZMYYC #911)), but other prefixes are attested elsewhere (for example, 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 (for example, 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 *ki:h* (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 (§3.1.1). The variant with *-ya-*vocalism, *tya-k* (STC, note 271, p.94), is reflected by Chinese 隻 *tśjäk* '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).[7] 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*.[8]

(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 (for example, in Lolo-Burmese) we are forced to recognise 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 (§3.2.2), a resemblance that may be more than accidental. (See §3.4, 'Interchange 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.[9] Any given language is likely to have

---

forms like Karok *ti:k* '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).

[7] 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.

[8] 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 x 1'), suggesting a connection with Jg. *maʔ⁵⁵ khra³¹* 'all; whole' (Dai et al., 1983:455). (An obvious alternative source for the Yogli syllable would be *ka* (§3.1.3).

[9] 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 *ti* occurs otherwise. See also the unstressed form for ONE in Kayah (§1.2.1), as well as Dulung *tiʔ⁵⁵* (Sun 1982:244-245), and Karenic forms like Pa-O *tə́ʔ-ba*, Palaychi *tə-*, Sgaw *tə-*.

developed several co-existent variants (much like English *one*, *an*, *only*, and so on; see §3.1), for example, Lahu *tê* 'one', *dê-dê* 'all', *tí* 'only', *tè?-chí* 'nothing', *a-cí* 'little bit' < PLB **day*, **nday*, **?dik*, **dek*, and **?gyik*, respectively.[10]/[11]

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, it is 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'.[12] 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).[13]

A group of forms with *tu* are perhaps distinct from the above:[14]

(Himalayish)	Khaling	*tu*
(Naga)	Yacham-Tengsa	*kha-tu*
	Wancho	*tu-ta*

3.1.5 NEW ROOTS FOR ONE IN ABOR-MIRI-DAFLA AND ELSEWHERE IN TIBETO-BURMAN

3.1.5.1 **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; also Nishi 1982)
Written Tibetan	*thor-bu* 'single; separate' (Jäschke p.289)

---

10 Supporting forms for **?gyik* '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 *tci⁵⁵*.

11 Several PLB variants are reconstructed in Matisoff (1972a) ('TSR') #31/#48 and #70: **-tik ⋈ **ti ⋈ **?dik ⋈ **?-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 recognised 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.

12 Other candidates for membership in the **day ⋈ **tay* group include Monpa (Dubey) *thee*, Nocte *wan-the* (Dubey, Das Gupta), van-*the* (GEM) (for the first syllable see **haŋ*, §3.1.5.3), Ersu *te⁵⁵*, Pumi *ti¹³* (Qinghua); *ti⁴⁵* (Taoba). For an alternative etymology of the Nocte form see §3.1.5.2.

13 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 hypothesised for this etymon is shown to be paralleled in several other cases. For detailed further discussion of these complex word families, please see Matisoff (1989/1995, #27).

14 Compare also Kham *to-bo*. Michailovsky observes that Khaling *-u* can be a reflex of **-ik* (pers.comm. 1995).

We reconstruct this etymon as *tir ⩨ *tur (-u- ⩨ -i- is a well-established variational pattern in TB; see VSTB pp.41–42).[15]

3.1.5.2 *(t)se

This group of Kamarupan forms meaning 'ONE' seems to have undergone 'contamination' with a root meaning 'TEN'. (See §3.2.2; §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[16]
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 §4.2.0.

Distinct from the above is another group of Kamarupan forms with -a vocalism, apparently from *sa or *tsa:

Garo	sa (Burling, Phillips); gesa (Momin)
Kokborok	-cha ~ -sa ~ -ca
Lakher	sá 'one' [17]

This Lakher morpheme is also used as a prefix before all the numerals 2–10, for example, sa-pali '4' (lit. '1 x 4'), sa-pangaw '5' ('1 x 5'). Also perhaps reflecting this etymon are Tiddim a-ma-sa 'first', Lotha ma-tsa-nga 'one'.

3.1.5.3 *han or *haŋ

Serdukpen	han (Dubey)
Zeme	hangkat
Maram	hangline

---

15 J.T. Sun (1993:234) reconstructs Proto Eastern-Tani *tel on the basis of the Mising-Padam and Milang forms.

16 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 (footnote 12).

17 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).

This morpheme seems to function as a fully syllabic prefix[18] 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' |

3.1.5.4 *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)[19]	*a*[21] (Sun 1981:217)
Qiang (Mawo)	*a* (Sun 1981:217)

3.1.5.5 *(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/rin ≈ l/ryaŋ* (§3.2.3.3; §3.4):

Pwo Karen	*lə́n* (Jones 1961:618)
Yimchungru	*khü-**lang*** (first syllable is a prefix)
Sangtam	*khürü* (also *khe*)
Kezhama	*ke**le*** (*ke-* is a prefix, part of a 1–3 run)
Mao	*ka**li*** (*ka-* is a prefix, part of a 1–3 run)
Sema	*la ki* (also *khe*)
Maram	*hang-li-ne* (cf. *hang, §3.1.5.3)

It it not clear whether these forms are relatable to a group of phonologically similar Himalayish words for ONE, for example Chourase *kolo/kwalo*; Yakkha *kolok*; Hayu (Hodgson) *ko-lu* (§2.1; §3.1.3b).

3.1.5.6 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:

---

18 See §5.4.4.6 'Where the lowest run has a fully syllabic CVC- prefix'.
19 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 (§3.1.4).

Deng Geman	$kuu^{31}mu^{53}$ (Sun et al. 1980:252)
Kaman (Miju Mishmi)	ku-mo (Das Gupta 1977a)
	kmo: ~ kŏmo: (LSI 3.1:623)[20]
Rengma	me 'one'
Tiddim	a-ma-sa 'first'
Lotha	ma-tsa-nga 'one'
Meithei	ama 'one'

Compare 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', §4.2.0.3.[21]

(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 #1224(i))), but since Jirel is a Bodish dialect, it would be well to find a cognate in WT before going out on a limb.

3.1.6 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' (§3.1.3(b)). This same etymon appears as a suffix in the vestigial numeral systems of Kiranti languages like Mewahang, Yakkhaba, and Lohorong (§3.1.3(b)), and has been generalised 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 (§3.1.5.2).

In fact nothing is more natural than for a language to develop a generalised counter or an indefinite article by semantic bleaching of the numeral for ONE. This is of course what has happened in English,[22] 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 (p.94) observes, "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.[23] Often a language will maintain several

---

[20]   kuu-/ku-/k- is now a meaningless prefix, part of a secondary 1-6 'prefix run' (§5.5), but is perhaps itself a reflex of *ko 'one' discussed above in §3.1.3. Compare also Mishmi (Dubey) mu-ou '10', especially in the context of the interchange between ONE and TEN (§3.4).

[21]   As a long shot we might compare these forms with the Lahu 'general classifier for objects', mà (Matisoff 1973a:91-92, 1988a:975-976).

[22]   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.

[23]   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').

etymologically distinct morphemes for '10', one used as the independent numeral, and the other(s) for the TEENS and/or ROUND NUMBERS.

3.2.1 *gip ✕ *gyap

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

In fact, however, this seems to be still another root where we must posit -i ✕ -ya- variation[24], as witness these forms from an AMD language: Kaman (Miju Mishmi) kyap-mo (LSI has kap), Deng Geman kiap^{55} mu^{54}.

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

3.2.2 *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,[25] Garo tśi, Dimasa dźi, Miju si (in composition), Karen (Taungthu) tśi, (Pwo and Sgaw) shi.

The vocalism of WB əchai poses a problem,[26] which the original version of STC (p.94) glosses over with the remark that it "appears to be related to this root through vowel gradation". In a new footnote (272), Benedict suggests that the solution is to change the reconstruction to *tsyay, but I feel it is better to recognise 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-thamba:q, Tsangla/Sharchop s(h)e, Kanawari sáí.[27] Here belong several other Himalayish forms with -u vocalism:[28] WT bću (Lhasa cu), Kaike chyu, Gurung cyuq, Thakali cyu, Jirel cyu-ta:mba:q, Sikkim Bhutia chu-tamba, Dzongkha (Mazaudon) cu-thãm.[29]

(Kamarupan) Monpa (Cuona) tɕi^{53}, (Dubey) chi; Monpa (Motuo) se; Garo chi-kung (Phillips), ci-king (Burling); Kokborok ći

(Baic) Bai (= Minchia) tsw^8 (Dell)

---

24 See *tik ✕ *tyak (§3.1.4), *it ✕ *yat (§3.1.1).
25 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 Moshang root meaning TEN, not mentioned in STC (§3.2.3.3(c)). Ultimately, however, I believe that the meaning ONE for forms like shi, she might actually be a transference from an original meaning TEN (see §3.1.5.2).
26 Note that Burmese has reflexes of both *gip (§3.2.1) and *tsyay, with some repartition of function. According to Judson (1953:215), "kyip is substituted for chai in the numbering of rational beings."
27 Perhaps ✕ Kanawari sa'e- '10 in additive higher round numbers'; see §3.5.3.3[D].
28 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.
29 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 §3.2.3.5.

(Qiangic) Ersu *tshɛ⁵⁵*, Proto Rgyalrong*syʈsye* (Nagano 1984) < **s-tsyiy* (JAM)

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

(Nungish) Dulung *tsăl⁵⁵*, Rawang *hti sel* , Nujiang *ts'i⁵⁵ tshăn⁵⁵*[30] (first syllables mean 'one'). Have these curious Nungish forms with final *-l* developed from **-y*, or do they point to an allofam **tsyal*?

3.2.3 NEW ROOTS FOR TEN IN KUKI-CHIN-NAGA, ABOR-MIRI-DAFLA AND ELSEWHERE

3.2.3.1 PROTO-KUKI-CHIN **som* (< **tsom*)

This root is widespread in Kuki-Chin,[31] 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-hni*? 'twenty'; Gangte, Hmar, Paite, Tiddim, Vaiphei *sawm*; Anal, Lai, Laizo, Ngawn, Thado *sɔm*; Maring *chip* 'ten' (< **gip* (§3.2.1), but *som-thum* '30', *som-li* '40';[32] 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, for example, *sot-bri* 'forty', *sot-bonga* 'fifty' (§3.5.1). 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 (for example, Bengni *čam-ni* 'twenty', *čam-pi* 'forty'; Hill Miri *čom-oum* 'thirty', *čaŋ-ŋo* 'fifty', *čem-piŋ* 'eighty').[33]

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

3.2.3.2 **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 **bo:n*, though he suggests that this might be a "loan from Austroasiatic into Northern Naga", citing Khasi *ši pón* 'ten' (*ši* means 'one'[34]).

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 §4.2.0). These forms seem to indicate that the original final consonant in this root was **-l*.[35]

Also undoubtedly to be assigned to this etymon are Phom *püan-* ('plus ten'), used in the odd round numbers of its vigesimal system, for example, *pinyi-püan* '50' ('(2 x 20) + 10'); and perhaps also Ntenyi *apyam-*, used in the decimal formation of its round numbers from 60 to 90 (see §3.5.2.2).

---

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

31   See Ono (1965).

32   GEM (gives Maring *som-nga* for both '20' (p.279) and '50' (p.79), but the former seems to be an error. See §3.5.1.3, §4.14.

33   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 §3.2.3.3.

34   Note the fortuitous similarity of this Khasi morpheme to some of the TB forms for ONE cited in §3.1.5.2.

35   The independent word for TEN in Meithei is *tara*, whose affiliations are elsewhere (§3.2.3.3(a)).

3.2.3.3 *s-r/liŋ and *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), ra:ng (Robinson), il-lyi (E. Dafla), Apatani alyā (for humans) ≉ lya (for non-humans; < *lya-ŋ); Gallong i'ri°~ i'yi°; Padam (Dubey) i:yi, i:i; Aka (Hruso) rhi, rɯ; Taraon ha:long, Darang Deng xa⁵⁵lɯng⁵⁵; Idu hū (Talukdar et al. 1962), hong⁵⁵ɦong⁵³ (Luoba: Sun 1983); Chulikata hush (< *hu-shV, with vowel of second syllable apocopated[36])

J.T. Sun (1993:144) sets up Proto Tani *rjɯŋ on the basis of Bengni ɯ-rjɯŋ, Lhopa/Bokar ɯjɯng, and the above Abor-Miri (=Padam-Mising) forms, also citing Dhammai lin, Bangru rəŋ⁵³, and Idu/Luoba (ZMYYC) hɹoŋ⁵⁵ (used in multiples of ten, for example, ni⁵⁵hɹoŋ⁵⁵ '20', a³¹soŋ³⁵hɹoŋ⁵⁵ '30').[37]

Several of the above AMD forms with h- or voiceless sonorant initials point to a possible *s- prefix on this root (Aka rhi, Idu hū ≉ 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- (for example, lha-pangu '50'[38]), Tangkhul həŋ- (for example, həŋ-phənga '50'), Southern Rengma hē (for example, hem-pfũ '50'), Angami (Kohima) hie- (for example, hie-pengou '50'), and Chokrü (=Chokri) he- (for example, hie-püngu '50' (GEM); we should add Angami (Khonoma) lhi- (for example, 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 (§3.2.3.4):

(a) With velar prefix:

Angami kerü, Chokri küri, Chakhesang keri, Liangmai kariu, Maram kero, Mzieme/Zeliang kerei, Zeme kereu, Mikir kre-[39]

---

36   This is an apocopating language. Compare Chulikata kāsh 'three' < *g-sum (§4.1.2, §5.1.3.1).
37   These AMD forms are phonologically quite similar to another, probably distinct root for HUNDRED (§3.5.4.6).
38   All these illustrative forms meaning '50' are from GEM, not Weidert.
39   This is the Mikir combining form for teens, as in kre-isi '11', kre-hini '12'; the independent Mikir numeral '10' is kep (§3.2.1).

28

(b) With dental prefix:

Ao (Mongsen) *tera*, Ao (Chungli) *ter* (with *apocope*)[40]; Khoirao *sara*, Lotha *taro*, Meithei *tara*, Meluri *tera*, Ntenyi *dagha*, ta'a[41]; Pochury *türa*, Rengma *tsarü*, Sangtam *thüre*, Tangkhul *thara*, Yacham-Tengsa *thelu*, Yimchungru *thürü*

(c) With palatal affricate word-initially:

Kezhama *chiro*, Mao *chüro*, Sema *chüghi*

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* (§3.2.2). 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 *chüghi* (where the '*gh*' presumably stands for [ɣ]) agrees well with other Naga forms (for example, 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 (§3.5.3.3).

3.2.3.4 **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*[42] '10' ('10 x 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'[43]

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* and *mia-hraw* '10'.[44] The Lakher voiceless liquid clearly points to an **s-* prefix at an earlier stage.

---

40   This form bears a merely accidental resemblance to some AMD forms descending from **tir* 'ONE', e.g. Mising *a-ter*, Padam *a-tel* (§3.1.5.1).
41   The apostrophe probably means glottal stop.
42   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. See footnote 25, above.
43   See Weidert (1987:413) and §3.5.2.1.2.
44   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').

We may thus combine the etyma discussed in §3.2.3.3–§3.2.3.4 into a single word family comprising both nasal- and stop-finalled allofams, and displaying both -*i*- ✶ -*ya*- and -*y*- ✶ -*w*-variation:

*s-riŋ ✶ *s-ryaŋ ✶✶ *s-ryak ✶ *s-rwak

There is some evidence of phono-semantic interchange between TEN and HUNDRED/THOUSAND in this root, which once might have meant something more vague, like 'BIG NUMBER'. See §3.5.4.6, §3.5.4.7.

3.2.3.5 *d(y)am ✶ *t(y)am 'ten; a full decade'

We have already mentioned (§3.2.2, footnote 29) a morpheme meaning FULL that occurs in several Himalayish compounds for TEN, evidently signifying something like the *completion of a full decade*, for example, Sherpa *ci-tham-baːq*, Jirel *cyu-taːm-baːq* and Sikkim Bhutia *chu-tam-ba*.[45]

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 1985:135); Khaling *tadam* (the first syllable looks like a reduction of *tu* 'one'); Thulung Rai *ko-dium* (glossed 'one-zero' in Allen 1975); Lepcha *ka-ti* (*kat* 'one'; see §3.1.2).[46] 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) *ka⁵⁵ti⁵⁵*, Pumi (Qinghua) *qa⁵⁵stiẽ⁵⁵*, Qiang (Taoping) *xa²¹dy³³* and Qiang (Mawo) *hədiu*.

3.2.3.6 *p/boŋ

Several Kiranti languages (East 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', and so on; Yakkhaba *ip-poŋ* '10'; Limbu *thi-booŋ* '10', *ni-booŋ* '20', *sum-booŋ* '30', and so on (Gvozdanović 1985:136, 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* (§3.1.4) and *ik-ko* (§3.1.1), respectively. See §4.0.2.

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

---

45   This root has several reflexes in Tibetan, e.g. WT *ltams-pa* 'be full', *tham-pa* ✶ *them-pa* 'complete, full', *ldem-pa* 'straight, upright'. As demonstrated in Matisoff (1988b), STC #226 'full' and #227 'straight/flat' really represent one and the same etymon. The presence of the 'infinitive' or nominalising suffix -*ba*/-*pa* in these Bodish forms indicates that the preceding morpheme is inherently verbal.

46   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 §3.3.2[A]. But see footnote 58, below.

3.2.3.7 APPARENT ISOLATES

There remain a number of isolated forms meaning TEN in individual languages that so far resist attempts at etymologisation, including:

(a) Boro *khao-she* '10' (*-she* 'one')[47]

(b) Mishmi *muou* '10'

   Is the *mu-* segmentable off with the meaning 'one'? Compare Geman Deng *kiap-mu* 'ten' (for the first syllable, see §3.2.1).

(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 analysing teen-systems include:

• Does the TEN morpheme come before or after the UNIT morpheme? For example, 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 English *-teen*), or is it a totally separate etymon (for example, Mikir *kep* '10', *kre-* '-teen')?

• Is the complex numeral agglutinative, easily segmentable into the TEN part and the UNIT part (French *dix-huit*, German *achtzehn*), or is it fusional (French *onze, douze, treize, quatorze, quinze*)? Does the UNIT morpheme undergo morphophonemic change when combined with the TEN morpheme (for example, English 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 (for example, Latin *duodēvīginti* '18', *undēvīginti* '19')?

---

47   There is an interesting lookalike in Hmongic, for example, White Hmong: *kaum [kao˧˩]* '10', *(nees) nkaum [nkao˧˩]* '20', *(peb) caug [kẏạo˧˩]* '30'. See Heimbach (1969:9, 77, 152).

- 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 refer to the numbers from 21 to 39 as a group – I suggest the term TWENTEENS. The twenteens 21–39 of vigesimal systems correspond to the teens 11–19 of decimal ones. (See, for example, the discussion of Sherpa, §3.5.3.4[B].)

- In the excessively rare duodecimal type of system represented by Chepang (§3.5.3.5), the numbers 13–23 correspond to the teens of decimal systems. We might as well call them the TWELVEENS.

3.3.1 TEEN FORMATION IN KUKI-CHIN-NAGA[48]

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

3.3.1.1 WHERE THE COMBINING FORM ('-TEEN') IS IDENTICAL OR MORPHO-PHONEMICALLY 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[49]	*ta'a*	*ta'a-*	*keching*	*ta'akecham*	*munga*	*ta'amanga*
Sangtam	*thure*	*thure-*	*asang*	*thureasang*	*munga*	*thuremunga*
Sema	*chughi*	*chughi-*	*kuthu*	*chughikuthu*	*pongu*	*chughipongu*

---

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

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

3.3.1.2 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[50]	*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*
Wancho	*ban*	*-ba-*	*ajam*	*banbajam*	*aga*	*banbaga*
Y-Tengsa	*thelu*	*-le-*	*asam*	*talulesam*[51]	*phungu*	*talulephungu*
Yimchungru	*thuru*	*-kheak-*	*asam*	*thurukheakasam*	*phungu*	*thurukheakphungu*
Zeme	*kereu*	*-ze-*	*kechum*	*kereuzekechum*	*mengeu*	*kereuzemengeu*

3.3.1.3 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[52]	*ara*	*-thoi*	*ahum*	*tarahumthoi*	*manga*	*taramanga*
Mzieme[53]	*kerei*	*-ngkei*	*ketsum*	*kerieketsumngkei*		
Puiron[54]	*som*	*-to*	*thum*	*somthumto*		
Rengma	*tsaru*	*-chu*	*keshan*	*tsarukeshanchu*	*pfu*	*tsarupfuchu*

3.3.1.4 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*

---

50   '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-*).

51   Note the combining form *talu-*.

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

53   11–14 are the only Mzieme teens that appear in Marrison. Note the combining form *kerie-* vs. independent *kerei* '10'.

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

3.3.1.5 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.[55] The morpheme *-ko* apparently means '*-teen*'. Note the intrusive *-ng-* in TWELVE.

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

	Angami (Khonoma)	Ao (Mongsen)	Meluri
SIXTEEN	*(kerosuru)*	*mukyimupenterok*	*mukweshuntaro*
SEVENTEEN	*mekupomothena*	*mukyimupenteni*	*mukweshunteru*
EIGHTEEN	*mekupomothetha*	*mukyimupentsit*	*mukweshuntuze*
NINETEEN	*mekupomotheku*	*mukyimupentuku*	*mukweshuntokhu*
TWENTY	*meku*	*mukyi*	*mukwe*

	Ntenyi	Pochury	Rengma
SIXTEEN	*kwushetuo*	*mkeshuntoro*	*nkipamotsaro*
SEVENTEEN	*kwushetughu*	*mkeshunturu*	*nkipamotsanu*
EIGHTEEN	*kwushetuza*	*mkeshuntuze*	*nkipamotutse*
NINETEEN	*kwushetukhu*	*mkeshuntoku*	*nkipamotukhu*
TWENTY	*mekweru/mukwung*	*mke*	*nki*

These formations are subtractive in a different sense from, for example, Latin *duodēvīginti* '18' and *undēvīginti* '19', literally '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,* and so on.[56]

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.'[57]

---

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

56   This is explicitly stated (in Hindi) in the Pochury source (Nagaland Bhasha Parishad 1972b:16).

57   See §4.2.0, 'Additive, subtractive, and multiplicative formations'.

3.3.2 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, for example, Tibetan and its dialects, Newari, Thakali, Kanauri and Dzongkha). Lepcha seems to be an exception to both of these generalisations:

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, and the *t'əp* takes over its semantic load.[58] 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', and so on; 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-*, for example, *song-sam* '13' and *song-zon* '17'.

[B] LOLO-BURMESE

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

tê chi tê	'11' '(1 x 10) + 1'
tê chi khɔ̀ʔ	'16' '(1 x 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-tìʔ* '11', *tshé-kòʔ* '16'.

---

58 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' (§3.2.3.5).

[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 §3.1.1)
> *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 (§3.5.2.2).

Kokborok[59]

> *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'.[60]

Boro (Bhat 1968)

Boro has the most thoroughgoing QUATERNARY (four-based) system that I am aware of in TB.[61/62] As D.N.S. Bhat (1968:29) says, "The system is basically a quadruplous one":

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 x 3) + 1'
FOURTEEN	zokkay-tam kanəy	'(4 x 3) + 2'
FIFTEEN	zokkay-tam katam	'(4 x 3) + 3'
SIXTEEN	zokkay-brə	'4 x 4'

---

59 Data from Pushpa Pai Karapurkar (1976).
60 Compare the mysterious first syllable of the etymologically distinct though semantically similar WB form *khu'-hnac* '7', alongside *hnac* '2'.
61 According to Ian Maddieson (pers.comm. 1984), quaternary numeral systems are widespread in African cultures that hold markets every four days.
62 Other possible reflections of an original quaternary system are Abor-Miri-Dafla multiplicative formations for EIGHT of the form '4 x 2' (see §4.2.3.7).

SEVENTEEN	*zokkay-brə kase*	'(4 x 4) + 1'
EIGHTEEN	*zokkay-brə kanəy*	'(4 x 4) + 2'
NINETEEN	*zokkay-brə katam*	'(4 x 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' and *thai-**ne**-nia* '9th' *(Revelation 21:19-20)* .

However, the quaternary system seems to be on the way out in Boro, judging 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' and *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', *khə mətay* 'increase, multiply', *akətay* 'remnant'; Wancho *a-tai* 'far', *tai-hu* 'many', and s o on.

[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 *-gu:* (§3.1.3). 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 the next section.)

3.4 INTERCHANGE AND CONFUSION BETWEEN ONE AND TEN[63]

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

---

[63] See §3.1.4 **t(y)ik*; §3.1.5.2 **s(h)e*; §3.1.5.5 **(k-)lV(N)*; §3.2.3.3 **s-riŋ* ⤨ **s-ryaŋ*; also §3.2.3.5, §3.5.4.6, §3.5.4.7. The Proto Mayan forms for ONE (**xu:n*) and TEN (*la:xu:n*; lit. 'end of one') are obviously related morphophonemically and conceptually (pers. comm. Terrence Kaufman 1994). According to Ives Goddard (pers. comm. 1994), there is also ONE/TEN interchange in Algonkian.

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...).[64]

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

3.4.1 *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' (§3.1.3(c)). But *-ko* means something quite different in Miju Mishmi, viz. -TEEN: *kumo* '1', *kumu-ko* '11'; *kinin* '2' and *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'.

3.4.2 *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 x 1'; for the first syllable see §3.2.1). 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* (§3.1.5.6), probably to be segmented *mu-ou*, with one syllable meaning ONE and the other meaning TEN – but which is which?[65]

3.4.3 *tšek / ş(ə)-* IN RGYALRONG

In the Zida dialect of Rgyalrong(data from Chang Kun), '10' is *ştşi* ; what seems to be the same initial element occurs in *şətšek* '11', *şənes* '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 *ştşi* meant '1 x 10' (i.e. < *tšek-tşi*); in the teens, however, where it co-occurs with the UNIT morphemes, the *şə-* 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.)

---

64   See §3.5.1.2–§3.5.1.3 'Decimal multiplicative TWENTY'.
65   This is the same problem STC faced (p.94) in analysing Moshang *rok-shi* '10' as '1 x 10', when actually its structure is '10 x 1'. See §3.2.2, §3.2.3.4.

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

Given the high degree of phonetic similarity between *s(h)e '1' (§3.1.5.2) and *ts(y)iy '10' (§3.2.2), 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

3.5.1 VIGESIMALITY IN TIBETO-BURMAN AND ELSEWHERE

Is vigesimality a primitive characteristic? Consider that well-known primitive language, French: *quatre-vingts* '80' ('4 x 20'), *quatre-vingt-un* '81' ('(4 x 20) + 1'), *quatre-vingt-dix* '90' ('(4 x 20) + 10'), *quatre-vingt-quatorze* '94' ('(4 x 20) + 14'), and so on. Many Francophones outside France (Belgium, Switzerland, Québec) sensibly prefer decimal alternatives to the higher round numbers, viz. *septante* '70', *octante* '80', *nonante* '90'.[66/67]

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

Danish[68]

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 ten') and EIGHTY (originally *firsindstyve*, lit. '4 times ten') were once segmentable as *tre-sinds-tyve*, *fir-sinds-tyve* (*sinds* 'times'), but have now been shortened (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

---

66 Note a bit of orthographic pedantry here: for the round number '80', the plural grapheme *-s* is used after *vingt* (*quatre-vingts*), emphasising 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.

67 Eric Hamp (pers.comm. 1994) believes that French vigesimality reflects a Celtic substratal influence.

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

numerical expressions at least make sense in terms of their constituent morphemes: FIFTY *halvtredsindstyve*, literally 'half-from-three times twenty', that is 'two and a half times twenty'; SEVENTY *halvfjerdsindstyve*, literally 'half-from-four times twenty', that is 'three and a half times twenty'; NINETY *halvfemsindstyve*, literally 'half-from-five times twenty', that is '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 – that is 'four (tens)', 'three (times (twenty))' and 'four (times (twenty))'!

In Tibeto-Burman, hesitation between decimality and vigesimality is apparent in several subgroups – Himalayish, Barish (e.g. Garo) and Kuki-Naga. Some languages have both kinds of systems in more or less free variation, with the vigesimal one apparently older.[69]

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, for example, English *eighty* (< EIGHT - TEN) versus *fourscore*. In Garo, for example, '20' is *kol-grik* ('20 x 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' and *sot-dok* '60'.[70] Conservative speakers use a vigesimal system throughout (for example, *kol-chang-gni* or *wakma-gni* '40'). (See Phillips 1904; Momin n.d.; Burling 1961.)

In a 'pure decimal' system (for example, Nocte, §3.5.1.3) even the word for TWENTY is analysable 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 (§3.5.2.4).

3.5.1.1 TWENTY AS A UNITARY, UNANALYSABLE MORPHEME

STC reconstructs a monomorphemic, unanalysable word for '20', like English *score*, of the shape **m-kul*[71], on the basis of the following forms: Jingpho *khun*, Garo *khol ~ khal*, Dimasa *khon*, Mikir *ingkol ~ ingkoi*, Siyin *kul*, Haka *kul ~ kwe*.[72]

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*

---

69 French itself belongs to this 'mixed' category of languages, with both decimal and vigesimal features. All its round numbers are decimal except for '80' ('4 x 20') and '90' ('(4 x 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 this 'vigesi-decimal vacillation', see §3.5.3.3.
70 This *sot-* is possibly related to Chin *som* '10' (§3.2.3.1).
71 STC #397, pp.15, 18, 83, 119, 120.
72 This etymon is discussed in Matisoff (1980:17-18), 'Stars, moon, and spirits...'

Lotha	*mekwi*	Tangkhul	*makɯ* (Bhat 1969)[73]
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 (#10, p.18), Benedict had reconstructed a distinct root **kun* 'all', based on WT *kun* 'all' and WB *kun* 'come to an end; be used up' amd *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',[74] 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. §3.5.3.3; §3.5.3.4) seem to reflect a prototype with **-a-*, i.e. **-kal*:

Sherpa *khal-jik*; Jirel *khalq*; Tamang *kha:l*; Khaling *(k)ha:el*; Lepcha *k'a*; and perhaps Sikkim Bhutia *khe-chik*, Sharchop *khye*[75].

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

Kaike (Him.) *tha:l*; Miju (AMD) *katal-mo* (*-mo* 'one', *ka-* is a secondary 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 (1985:136) points out in her excellent study of the Dzongkha numeral system, 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' (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, for example, 'a complete load; everything that can be placed on a beast of burden at one time'. The variational pattern *-a-* ⋇ *-u-* is grudgingly recognised even in STC (e.g. #405 **b-suŋ* ⋇ **b-saŋ* 'fragrant').[76/77]

---

73   Marrison (1967) has Tangkhul *maga*.

74   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'.

75   Mazaudon (1985:154) cites several additional Himalayish forms, including Gongar (Bhutan) *khay /khel*, Dungkarpa *khɛ*, Thakali *khal*, and Tamang *⁺pokal*, as well as Tipra (=Kokborok) *khol*.

76   Several other such etyma are discussed in VSTB (Matisoff 1978a:43-44). See also our posited allofamic alternation **sam* ⋇ **sum* 'three' (§4.1.2). Note that the Garo alternants cited in STC (*khol ~ khal*)

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

[A]   Meithei *-phu*; Wancho *pu-*[78]; 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 (§3.5.2.2; §3.5.2.4).

These etyma are used multiplicatively to form the higher twenties (including HUNDRED) (§3.5.4.2).

3.5.1.2 DECIMAL MULTIPLICATIVE TWENTY, WITH THE UNIT FIRST

‘20 = 2 x 10’

This is the universal pattern in Lolo-Burmese, for example, Lahu *nî chi*.

3.5.1.3 DECIMAL MULTIPLICATIVE TWENTY, WITH THE UNIT SECOND

‘20 = 10 x 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*[79]	*somthum*	*somli*	*somnga*

Note that there are many Kuki-Chin languages where *som* is the independent numeral ‘10’ (for example, 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 x 1’, TWENTY is of course also expressed as ‘10 x 2’. See §3.4; §3.5.2.1.3.

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:

---

confirm the reality of the variational pattern with this root, as perhaps do the Tangkhul variants *maga* and *makɯ* (footnote 73).

[77]   Aficionados of worldwide look-alikes 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 *preglottalised initial) and Proto Otomanguean **kala*! (Pers.comm. Terrence S. Kaufman, April 1994.)

[78]   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 §3.5.2.4), but perhaps these may all be referred back to an earlier ***k(y)al*.

[79]   Marrison (1967:279) has ‘*somŋa*’ 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.

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 x 2*'
>
> *sa-ki* 'twenty'
>> *sa-* '1'; *ki* = '20' : '*1 x 20*'
>
> *hlei-hraw* 'twenty'
>> *-hraw* '10'; *hlei* seems to mean 'pass, exceed, be extra'[80]

3.5.2 ROUND NUMBER FORMATION IN KUKI-CHIN-NAGA

3.5.2.1 DECIMAL SYSTEMS OF ROUND NUMBER FORMATION IN KCN

3.5.2.1.1 WHERE THE MORPHEME FOR TEN IN THE COMPOUND NUMERALS IS THE SAME AS (OR A MORPHOPHONEMIC VARIANT OF) THE INDEPENDENT NUMERAL '10'

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

Yimchungru has a special word for '50' (§3.5.2.2).

3.5.2.1.2 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 to 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 x 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 x 4, 10 x 5...10 x 9*, with the UNIT morpheme *following* this TEN morpheme. The morpheme for TEN used in composition is usually etymologically distinct from the independent numeral.

---

[80] 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	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[81]	*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[82]	*chüghi*	*küthu*	*sheghi*	*lho-*
Tangkhul	*thara*	*kathum*	*thumra*	*hang-*
Zeme	*kereu*	*kechum*	*himreu*	*he-/re-/ riak-*[83]

	40	50	60
Angami (Khonoma)	*lhida*	*lhipengu*	*lhisuru*
Angami (Kohima)	*hiede*	*hiepengou*	*hiesorou*
Chokri	*hieda*	*hiepungu*	*hieshwuru*
Kezhama	—	*lhapangu*	—
Khoirao	*ramri*	*renga*	*reksaruk*
Liangmai	*atai*	*ringiu*	*riacharuk*
Mao	*ridei*	*ripongo*	*richoro*
Maram	*ragdai*	*rengo*	*reksaruk*
Mzieme	*riakdai*	*riangngei*[84]	*riakheruk*
Nruanghmei	*rekdai*	*rekngu*	*rekcüruk*
Rengma	*henzi*	*hempfü*	*hentsaro*
Sema	*lhobdhi*	*lhopongu*	*lhotsogho*
Tangkhul	*hangmati*	*hangphanga*	*hangtharuk*
Zeme	*hedai*	*rengeu*	*riakseruk*

---

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

82   Note *küthu* 'three' ≈ *she-* 'thir-'.

83   Zeme '70, 80, 90' have *riak*: *riaksena, riakdesat, riaksekui*.

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

(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'; that is instead of a monomorphemic word for '20', it too is composed of the special morpheme for TEN[85] (along with the unit morpheme TWO): *ruaknyi* '20'.

3.5.2.1.3 WHERE TEN IS A CLASSIFIER, NOT A NUMERAL

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

(a) This structure is common in Lolo-Burmese (see §3.3.2[B]):

	1	10	2	20	3	30
Lahu	*tê*	*tê chi*	*nî*	*nî chi*	*šêˀ*	*šêˀ chi*

tê chi	'10' (Num + Clf)	'1 x 10'
nî chi	'20'	'2 x 10'
šêˀ chi	'30'	'3 x 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 analysable into two morphemes, ONE and TEN, but merely consists of a unitary root for TEN preceded by a meaningless prefix.

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 unanalysable monomorphemic word for '20'.

3.5.2.2 LANGUAGES WITH A SPECIAL WORD FOR FIFTY

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

---

[85]   For the etymology of this special combining form for TEN, see §3.2.3.3(c).

	5	10	20	50
Ao (Chungli)	*pungu*	*ter*	*metsü*	*tenem*
Ao (Mongsen)	*phanga*	*tera*	*mükyi*	*tünam*
Lotha	*mungo*	*taro*	*mekwi*	*ti-ingya*[86]
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-* and *-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 x 5* or *5 x 10*).

(a) First syllable:

There is some basis for hypothesising 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.[87] However, these eight languages have generalised a dental prefix with the numerals all the way from 6 to 10,[88] 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 far-fetched, we shall soon see (§3.5.2.3) 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 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[89]) 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).[90]

---

86   Compare Lotha *ti-ing* '7', *ekhati-ing* '70'.

87   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 §4.1.4.1.1.

88   A slight exception is Sangtam, which lacks a dental prefix in '8' (*thuro, thunye, ke, tuku, thure*). See §5.4.4.

89   Among the puzzling sidelights here is the similarity between FIFTY (*ti-ingya*) and SEVEN (*ti-ing*) in Lotha. Is this merely accidental? See §4.2.2.8.

90   This is similar to the case of those languages discussed above (§3.5.2.1.2(a)) where it was the word for THIRTY whose constituent order was out of step with all the higher round numbers.

	30 (3 x 10)	40 (4 x 10)	50 (10 x 5 ?)	60 (6 x 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 x 10)	80 (8 x 10)	90 (9 x 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).[91] 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 'middle of the roof in houses' (i.e. 'ridgepole')."[92] It is possible that forms like Sangtam *thunyang* or Lotha *tiingya* are also to be analysed 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*) x 2 (*nyü*)': *zyure-re-anyü*. Ao Mongsen has an identical formation for '80': 40 (*lira*) x 2 (*anet*): *lira-anekhi*.

---

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

[92]   Pers.comm., 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 sǝru* (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'.

- 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 x 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 (§3.2.3.2). 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 (§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',[93] 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.[94]

- The Ao Chungli round numbers (except '20' and '50') are all formed multiplicatively of UNIT x 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 x 2', perhaps to avoid the near-pernicious homophony found in Mongsen. We should note that the Chungli independent numeral '4' is *pezii*; a more ancient-looking form (< PTB **b-ləy*) now survives only in *lir* '40'.

3.5.2.3 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

---

93   It is perhaps to be analysed as *tiing-ya* (above).
94   Compare French *chambre* 'room' < Latin *camera* (Vulg. Lat. *camra*).

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) x 10', that is '9 x 10.'[95]

3.5.2.4 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 x 3) + 10' (i.e. 60 + 10) or '-10 + (20 x 4)' (i.e. 80 – 10).[96]

In what we might call 'super-vigesimal' languages (for example, Wancho and Chang in the following chart), the word for HUNDRED is also expressed in terms of TWENTY ('20 x 5') – that is HUNDRED is morphologically complex.

	Wancho	Phom	Konyak	Chang
TEN	ban	an	pen	an
TWENTY (indep.)	tsa/ca[97]	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

---

95   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 (§3.5.2.2). For more on the Mikir system, see §5.4.2.1.

96   Compare French quatre vingts '80', quatre-vingt-dix '90' ('(4 x 20) + 10'), and the Danish system discussed above (§3.5.1).

97   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".

FIFTY	*punyi-ban*	*pinyi-püan*	*teija-pen*	*anchinsam*
SEVENTY	*puram-ban*	*betjam-püian*	*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' x *-i-* '2'/*-lem-* '3'/*-peli-* '4' x *-ja* '1'.

That is, '40' = 20 x 2 x 1; '60' = 20 x 3 x 1; '80' = 20 x 3 x 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 x 1'. Chang is supervigesimal, in that HUNDRED is also expressed in terms of TWENTY ('20 x 5'). W.T. French derives the form *sau-* from Proto Northern Naga **ja* (§3.5.1.1), 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 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').[98] 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', that is '60 minus 10'; *an-chin-lei* '70' is '10 (from) 20 times 4', that is '80 minus 10', and so on.

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*

---

98   See §4.2.0.

- 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' and '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, §3.5.2.2!) 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 generalisation 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*.

3.5.3 ROUND NUMBER FORMATION IN HIMALAYISH[99]

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.

3.5.3.1 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 §2.1.1.)

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: for example, Nepali *cha*, *sa:ṭh*, *a:ṭh*, *nau* '6–9', but *sa:ṭhi*, *sattari*, *asi*, *nabbe* '60–90'. Although these were

---

[99]   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/ Grünwedel (1898) (Lepcha), Mazaudon (1985) (Dzongkha), Chhewang Rinzin (pers.comm. 1984) (Dzongkha, Sharchop), and Sandberg (1895) (Sikkim Bhutia).

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.

3.5.3.2 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:

	10	3	30	4	40	5	50
Wr. Tibetan	*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:*[100]	/ < 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, that is '2 x 10': WT *nyi-shu*, Thakali and Gurung *ngih-syu* and Newari *ni:-* / < ni + i/. Note that the unit morpheme precedes the TEN morpheme in all these systems.

3.5.3.3 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 neighbouring multiple of 20:

[A] CUONA MENBA (MAMA (SOUTHERN) DIALECT: LU SHAOZUN 1986:184-185)

ten	$t\varphi i^{53} \sim t\varphi i\eta^{55} \sim t\varphi i p^{53} \sim t\varphi i k^{53}$[101]	
twenty	$kh_A{}^{55}li^{55}$	
forty	$ch\varepsilon^{\text{？}53}\text{-}n_Ai^{53}$	('20 x 2')
sixty	$ch\varepsilon^{\text{？}53}\text{-}sum^{53}$	('20 x 3')
eighty	$ch\varepsilon^{\text{？}53}\text{-}pli^{53}$	('20 x 4')

---

[100]   *-gu:* is a suffix which occurs with all Newari numerals (§3.1.3), and is irrelevant to the present discussion.

[101]   The final of this morpheme assimilates to the initial of the following unit numerals in teen-formation: $t\varphi i^{53}the^{\text{？}53}$ '11', $t\varphi i\eta^{55}nAi^{35}$ '12', $t\varphi i k^{53}sum^{53}$ '13', $t\varphi i p^{53}pli^{53}$ '14'.

hundred	$ch\varepsilon\textipa{P}^{53}$-$le^{31}\eta e^{53}$	('20 x 5')
thirty	$kh\textsc{a}^{55}li^{55}$-$t\textctc i^{53}$	('20 + 10')
fifty	$ch\varepsilon\textipa{P}^{53}$-$n\textsc{a}i^{53}$-$t\textctc i^{53}$	('(20 x 2) + 10')
seventy	$ch\varepsilon\textipa{P}^{53}$-$sum^{53}$-$t\textctc i^{53}$	('(20 x 3) + 10')
ninety	$ch\varepsilon\textipa{P}^{53}$-$pli^{53}$-$t\textctc i^{53}$	('(20 x 4) + 10')

In this dialect the word for '20', $kh\textsc{a}^{55}li^{55}$, 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, $ch\varepsilon\textipa{P}^{53}$-, 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 $kha^{55}li^{55}$, but the variant that occurs with 40–100 is $kh\textsc{a}i^{55}$.) 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 x 1'). Tamang apparently goes so far as to maintain the vigesimal system all the way up to 1,000, which is expressed as *kha:l-paca:s* ('20 x 50'), with the second element from Nepali. (Gurung and Thakali, so closely related to Tamang, are completely decimal, further demonstrating that decimality versus vigesimality is a useless criterion for linguistic subgrouping.)

[C] KAIKE AND JIREL

	Jirel	Kaike
10	cyuta:m:baq	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 x 10'. Jirel uses the inherited monomorphemic TB *khalq* for everything above 20 (for example, *khal-jik-tangq-nyiq* '22'), but Kaike uses *nhi-chyu* throughout the twenteens (for example, *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 x 1, 2, 3...'), Kaike puts the UNIT before the TWENTY ('1, 2, 3... x 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, §3.2.3.3c.) 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 x 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' and *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 x 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 x 20); 90 = 10 + (4 x 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' (§3.2.2).

[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, for example, *cha-bis-e-das* '130' (see §3.5.4.5).

In Khaling, TWENTY and its multiples are expressed by the inherited TB morpheme -*(k)ha:el* (*tu-ha:el* '20', *sa:h-kha:el* '40', *suk-kha:el* '60', *bha:el-kha:el* '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: *paca:s* '50', *sattari* '70' and *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.

3.5.3.4 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 (x 1)'	*kə-nyət*	'10 x 2'
FORTY	*k'a-nyət*	'20 x 2'	*kə-fəli*	'10 x 4'
SIXTY	*k'a-sam*	'20 x 3'	*kə-tərək*	'10 x 6'
EIGHTY	*k'a-fəli*	'20 x 4'	*kə-kək'u*[102]	'10 x 8'
HUNDRED	*k'a-fəngo*	'20 x 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 x 1) + 10'	*kə-sam*	'10 x 3'
FIFTY	*k'a-nyət sə kəti*	'(20 x 2) + 10'	*kə-fəngo*[103]	'10 x 5'
SEVENTY	*k'a-sam sə kəti*	'(20 x 3) + 10'	*kə-kəkyək*	'10 x 7'
NINETY	*k'a-fəli sə kəti*	'(20 x 4) + 10'	*kə-kəkyot*	'10 x 9'

---

[102] This form is missing from Mainwaring, and is my guess.
[103] 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'.

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-* in *kəti* is only part of a 'prefix run' that extends from 6 to 10 in Lepcha (§5.3.4).

[B] SHERPA

Sherpa has a classically vigesimal form for TWENTY, *khal-jik* ('20 x 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(thamba:q)* ('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 recognised: people from 21 to 39 could be called *twenteenagers*. See §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 'decimalised' system continues through and beyond HUNDRED: *khal-citambaq* '100' ('10 x 10') (etymologically '20 x 10'), *khal-cupsum* '130' ('10 x 13') (etymologically '20 x 13'). Compare Jirel *khalq-Thuk-cyutambaq* '130' ('(20 x 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 x 10') in Rinzin, but *khe-chik* in Sandberg (first syllable < **m-kul* '20', second syllable < **tyik* '1' (§3.1.4)). The higher round numbers are all formed decimally with *chu-* or *chu-tamba* '10', for example, *zhib-chu* '40', *ngab-chu* '50', *Tuk-chu* '60' (note the *-b* at morpheme boundary in '40' and '50').[104] 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 (1985:150) demonstrates that "Dzongkha exhibits a coherent vigesimal system equal in complexity and extension to any vigesimal system described in any part of the world". Not only does the

---

[104] These forms are obviously resyllabifications of compounds where the second element began with prefixal *b-* (cf. WT *bcu* '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'.

language retain the general monomorphemic TB root for '20' ($khe < *m\text{-}kul$), but it also has lexemes for the next three powers of twenty:

20^1	*khe*	20
20^2	*ɲicu*[105]	400
20^3	*kheche*[106]	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-cu* ('2 x 10')	*khe-ci* ('20 x 1')
21	*tsa-ci*[107] ('20 + 1')	*khe-ci (da) ci* ('(20 x 1) + 1')
22	*tsa-ɲi*	*khe-ci (da) ɲi*
30	*sum-cu* ('3 x 10')	*khe pɟhe da ɲi* ('20 x (-1/2 + 2)')
31	*so-ci* ('thir- + 1')[108]	*khe-ci da cu-ci* ('(20 x 1) + 11')
35	*so-ŋa* ('thir- + 5')	*khe ko da ɲi* ('20 x (-1/4 + 2)')
40	*zi-p-cu* ('4 x 10')	*khe-ɲi* ('20 x 2')
50	*ŋa-p-cu*	*khe pɟhe da sum* ('20 x (-1/2 + 3)')
55	*ŋa-ŋa*	*khe ko da sum* ('20 x (-1/4 + 3)')
60	*ɖhuk-cu*	*khe-sum* ('20 x 3')
70	*dyn-cu*	*khe pɟhe da zi* ('20 x (-1/2 + 4)')
80	*ge-p-cu*	*khe-zi* ('20 x 4')
90	*gu-p-cu*	*khe pɟhe da ŋa* ('20 x (-1/2 + 5)')
100	*cik-ɟa* ('1 x 100')[109]	*khe-ŋa* ('20 x 5')
400	*zip-ɟa* ('4 x 100')	*ɲicu-ci* ('400 x 1')
500	*ŋap-ɟa* ('5 x 100')	*ɲicu-ci da khe-ŋa* (('400 x 1) + (20 x 5)')
600	*ɖhuk-ɟa* ('6 x 100')	*ɲicu pɟhe da ɲi* ('400 x 1 1/2')

In the vigesimal system, the even round numbers are expressed straightforwardly as multiples of *khe* '20', but the odd tens and fives are formed by what Mazaudon (following

---

[105] This form, etymologically '2 x 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 §4.0.2 below.

[106] The second syllable is identified by Mazaudon (1985:137) with WT *che-ba* 'large', so that the compound means 'a large twenty', much as French *une grosse* (> English *gross*) '12 dozen; 144' derives from *une grosse douzaine* 'a big dozen'). The etymology of the first syllable of *jãche* remains obscure. Michailovsky (pers.comm. 1995) suggests a connection with Tibetan *yaŋ* 'again', that is 'even bigger'.

[107] The morpheme *tsa* is a fascinating example of a radical but entirely natural semantic slippage. As Mazaudon shows (1985:129), it derives from the WT conjunctive particle *rtsa* used to connect the tens to the units, e.g. *nyi-shu-rtsa-gcig* '21' ('2 x 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, §4.0.2. The morpheme *da* (< WT *daŋ*) then took over connective function in the Dzongkha vigesimal system (e.g. *khe-ci da ci*).

[108] See footnote 104 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).

[109] One can also say *ɟa-thampa* ('100 full'), with the same 'full' morpheme as found, e.g. in Sherpa and Jirel (§3.2.2; §3.2.3.5; §3.5.3.3[C]; §3.5.3.4[B]).

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'), that is 'which with an additional one-half, would be two'), the odd round numbers are formed subtractively in Dzongkha: for example, '30' *khe pɟhe da ɲi*, that is 'twenty times one-half-less-than-two', or 'twenty times one-and-a-half').[110] 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 (1985) 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).

3.5.3.5 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: *ya:t-joʔ, nis-joʔ, sum-joʔ, play-joʔ, ponga-joʔ*.[111] SIX through TEN are expressed by Nepali numerals: *cha-gota:, sa:t-gota:, ʔa:t-gota:, naw-gota:, das-gota:*. The crucial form for ELEVEN is missing in CSDPN, but presumably it too is simply the unanalysable Nepali word.

The Chepang word for TWELVE is truly unique: *ya:t-ha:le*. The first syllable means ONE (§3.1.1), and the second element is clearly the inherited TB word for TWENTY (< **kal ≋ *m-kul*; cf. Khaling (*k)ha:el*). 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.[112]

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

ya:t-ha:le ʔa:t-gota:	20	('(1 x 12) + 8')
ya:t-ha:le das-gota:	22	('(1 x 12) + 10')
nis-ha:le	24	('2 x 12')
nis-ha:le ponga-joʔ	29	('(2 x 12) + 5')
nis-ha:le sa:t-gota:	31	('(2 x 12) + 7')
sum-ha:le play-joʔ	40	('(3 x 12) + 4')
play-ha:le nis-joʔ	50	('(4 x 12) + 2')
ponga-ha:le	60	('5 x 12')

---

110 This is not so different from the German and Russian way of expressing the half-hour when telling time, for example, German *halb vier*, Russian *polovino cetvortogo* '3:30', i.e. 'half of four'.

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

112 Compare the case of Sherpa, §3.5.3.3, where this same etymon has been transvalued from TWENTY to TEN.

· 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 (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.)[113]

Since the first version of this monograph was written (in 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 x 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).

3.5.4 HUNDRED AND THOUSAND

3.5.4.1 DECIMAL MULTIPLICATIVE EXPRESSIONS FOR HUNDRED ('10 X 10')

Gallong (AMD) *cam-ri* '100' (*cam-* '-TY', as in *cam-nyi* '20' ('-TY x 2'), *cam-um* '30', *cam-ke* '60'; *i-ri* '10'). (See 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' (§3.5.3.4[B]; §4.0.2).

Lakher (KCN) *sy-hraw* '100' (*sy-* '-TY', as in *sy-pali* '40'; *pa-hraw* '10'). Lakher has two other ways of expressing '100', either as an independent unanalysable numeral (*za*), or as '100 x 1', where the morpheme HUNDRED functions as a classifier (*za-kha* '100', *kha* '1').

3.5.4.2 VIGESIMAL MULTIPLICATIVE EXPRESSIONS FOR HUNDRED ('20 x 5' OR '5 x 20')

	HUNDRED	TWENTY	FIVE
Kuki-Naga			
Chang	*sau-ngau*	*sau*	*ngau*
Wancho[114]	*pu-ga*	*pu-*	*ga*
Yacham-Tengsa	*mesung-phung*	*mesung-*	*phung*

---

[113] 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).

[114] 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).

Abor-Miri-Dafla
    Monpa Motuo   *k'ai-nga*   *k'ai*   *nga*

Himalayish
    Tamang   *kha:l-nga:h*   *kha:l*   *nga:h*
    Sharchop   *khye-nga*   *khye*   *nga*
    Lepcha   *k'a-fəngo*   *k'a*   *fəngo*
    Kaike   *nga:-thal*   *thal*   *nga:*

3.5.4.3 WHERE HUNDRED HAS A SPECIAL RELATIONSHIP WITH FIFTY

In Miju Mishmi (Das Gupta 1977a), '100' is *waie-mo*[115] (*-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.

3.5.4.4 WHERE HUNDRED FUNCTIONS AS A CLASSIFIER

In languages with this formation, HUNDRED is expressed as '100 x 1' or '1 x 100'.

(a) '1 x 100'

Lolo-Burmese
    Lahu *tê ha* '100' (*tê* '1'); *ha* < PLB **hra^1* or **?ra^1*

Qiangic
    Ersu *tə^{55}zə^{55}*

(b) '100 x 1'

Himalayish
    Dzongkha *ja-ci* (*ci* '1')

Barish
    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)[116]; Lakher *za-kha* (alongside two other expressions: see §3.5.4.1); Pochury *mza-ke*; Garo *ritcha-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

---

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

116 The apparent homophony of these two syllables reminds one of the Lahu number *hí hí* '8,000', where even the tones are identical. This is pure accident, however: the first syllable is from PLB **?rit* 'eight', but the second is a loan from Shan *hiŋ* 'thousand'. See Matisoff (1988a:1070).

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.

3.5.4.5 WHERE HUNDRED IS A MONOMORPHEMIC REFLEX OF PTB *b-r-gya* ⍨ *b-g-rya*[117]

It seems clear that this etymon has been remodelled by analogy with EIGHT (§4.2.3). For EIGHT, STC recognises 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*; Rgyalrong *perzhe*; Aka (Hruso) *phogwa*, *purrwa*

b-rya   *with labial prefix*
Mikir *paro*; Nruanghmei *phai* (with preemption of the initial)
Gurung *pra:hq*; 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*[118]; Puiron *raja*; Kom Rem *raza*; Phom *gho* (presumably '*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* (§3.5.3.3[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*; Yimchungru *chhi*
Lushai, Lakher, Hmar, Gangte, Paite, Tiddim *za*; Vaiphei *ja*
Boro *jou* ; Dulung *çya*
Pumi *shə*[55]; Qiang *tshi*[55]
Karen (Taungthu) *rja*, (Pwo) *ja*, (Palaychi) *sa*

---

117   STC #164 *r-gya* (pp.45, 54, 57, 89, 94-95, 109, 131, 137, 151, 161-162).
118   Alongside Jg. *mətsát* 'eight'.

3.5.4.6 A SPECIAL ABOR-MIRI-DAFLA ROOT FOR HUNDRED AND ITS RELATIONSHIP TO TEN

$*m\text{-}li(\eta)$ (? $< *m\text{-}l\text{-}\eta ya$)

Abor-Miri	*li-ko ~ ling-ko*
Minyong	*ling-ko*
Dafla	*leng-go*
Nishi	*lunkh*
Apatani	*lange, laŋ-e*
Monpa Cuona	$c'\varepsilon^{\text{?}53} le^{21} nge^{53}$
Lhopa	*lɯng*
Darang Deng	$m\partial lum^{55}$
Chulikata	*malu:*
Mishmi	$malo^{119}$
Gallong	*hamyi* ($< *s\text{-}mli\eta$?)

It will be observed that these forms bear more than a slight resemblance to a set $*ri\eta \asymp *yi\eta$ 'TEN' discussed above (§3.2.3.3(a), (b)): for example, Abor-Miri *eying-ko* '10', and so on. However, my present view is that these two roots are distinct (cf. pairs of reflexes like Lhopa *ɯjɯng* '10' and *lɯng* '100'),[120] though they may have 'contaminated' each other.[121] It is possible that these forms for HUNDRED are ultimately to be derived from $*m\text{-}l\eta ya$ (ult. $< *b\text{-}rgya$) via apocope of the root vowel.

3.5.4.7 THOUSAND

(a) $*s\text{-}to\eta$

STC #32 (pp.21, 94) sets up the root $*s\text{-}to\eta$ 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	$st\bar{\imath}^{55}$
Qiang	xto^{55}
Dulung	tu^{55}
Karen	(Pwo) *thon*, (Sgaw) *kə?tho*

---

119   Compare perhaps Mishmi (Dubey) *muou* '10' (§3.2.3.7, §3.4.2).

120   J.T. Sun (1993:121) sets up a Proto Tani root $*lɯ\eta$ 'hundred', distinct from PTani $*rjɯ\eta$ 'ten' (p.144), citing forms like Bengni and Bokar *lɯŋ*, Bangru $l\partial\eta^{53}$, Dhammai *bɯ-loŋ*, and Hruso *phu-yu*.

121   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'.

(b) *s-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 (§3.2.3.3; §3.5.4.6):[122]

(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 (Taungthu) *tə̀ʔ-rèng* (*tə̀ʔ* '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]* '1,000' looks like the root for TWENTY found in a few languages (Meithei *phu*, Wancho *pu-*, Phom *pü*). See §3.5.1.1.

- 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, p.279).

- Tiddim *tul* and Mikir *suri* look as if they are related to each other, but not obviously to anything else.

3.5.4.7.1 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 x 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	lā-lyā	(lā '100', lya '10')
Abor-Miri	li-ying-ko	(li ~ ling '100', eying '10')
Adi[123]	ling-ko-iying	(ling-ko '100', iying '10')

---

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

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

(b) '10 x 10 x 10'

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

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

(c) '20 x 50'

    Tamang   *kha:l paca:s*   (*kha:l* '20', *paca:s* '50' (< Nepali))

(d) '20 x 10 x 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 '1,000'![124]

3.5.4.7.2 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 *sahasra* '1,000' (*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 *hāqjār*

    Kanawari *hanzár* (apparently with rhinoglottophilia in the first syllable; see Matisoff 1975.)

    Konyak, Phom *haja*

    Sangtam *hajar*

---

[124]   Could *puban* be a typo in Marrison (1967:267) for *puga-ban* ('100 x 10')?

THE PRIMARY NUMERALS: TWO TO NINE

4.0 INTRODUCTION

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, for example, Apatani *lya* '10' (of non-humans) – *alyā* '10' (of humans) (§3.2.3.3); Dzongkha *tsa* '20' (for counting objects from 21 to 29, for example, *tsa-ɲi* '22') versus *ɲer* '20' (for dates, for example, *ɲer-ɲi* 'the 22nd of the month'; Mazaudon 1985:129).

(1) *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 (§5.2ff.), for example, Jingpho *məsūm* '3' (< *g-sum) <-- *məlī* '4' (< *b-ləy).[1] 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' (§1.1.2).

(b) Influence on root-initial consonant

Consecutive numerals may influence each others' initial consonants, with several famous examples in Indo-European.[2] Compare, for example, Phom *shüt* '8' --> *shü* '9'; Serdukpen *khu* '5' <-- *khit* '6'; Khaling *sa:hpu* '2' <-- *suhpu* '3'; Khoirao *kati* '2' (< *kani) <-- *kathum* '3'.

---

[1] In this discussion we use the symbols --> and <-- to indicate the direction of influence.

[2] For example Germanic FIVE influenced FOUR (we would expect English *whour); Russian *desjat'* '10' influenced *devjat'* '9' (instead of the expected *nevjat'*). See §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).

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, for example, 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'.[3]

(d) Generalisation of final consonant to suffixal status

We have mentioned how some Karen dialects, especially Pa-O (Taungthu) have generalised 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 §1.2.1, 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.[4]

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 generalising 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, for example, Mikir *throk* '6' --> *throk-si* '7' ('6 + 1') (§4.2.2.9); Meithei *nipan* '8' ('2 from 10'), *mapan* '9' ('1 from 10') (§3.2.3.2; §4.2.3; §4.2.4). See §4.2.0 below.

(2) *Transvaluation of numerals (semantic)*

Throughout this monograph we are concerned with pointing out indisputable or possible cases of 'numerical transvaluation', that is a shift in the referent of a numeral from its etymological meaning. In this section the most interesting examples are listed together for convenience's sake.[5]

- Interchanges and confusions between 'ONE' and 'TEN' (§3.4), and between 'TEN' and 'HUNDRED' (§3.5.4.7b).

---

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

4   See my note 413 in STC (p.152), and §4.1.2.2.

5   In Appendix I we shall offer a semantic diagram or 'flowchart' that schematises all these shifts in meaning.

66

- Yacham-Tengsa *tamong* '20' looks like the same etymon that means 'FIFTY' in most other languages where it occurs (§3.5.2.2).

- The Mikir reinterpretation of *throk* 'SIX' to mean 'TEN' in the higher round numbers (above §3.5.2.3).

- The change in the value of **m-kul* 'TWENTY' to 'TEN' in Sherpa (§3.5.3.4[B]), and to 'TWELVE' in Chepang (i.e. from '2 x 10' to '2 + 10'; §3.5.3.5).

- The Dzongkha numeral *ɲiɕu*, which sometimes means the etymologically correct 'TWENTY', but is often used to mean '20 x 20' or 'FOUR HUNDRED' (§3.5.3.4[C]).[6]

- In Lepcha an etymological flipflop between 'EIGHT' and 'NINE' seems to have occurred: *kăkŭ* '8' (but cf. PTB **d-kəw* 'nine', §4.2.4), *kăkyót* '9' (but cf. PTB **-gyat* '8', §4.2.3).

- 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 (§4.1.1; §4.1.4; §4.2.0; §4.2.2).

- 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') (§3.2.3.6; Gvozdanović 1985:162).

  •• The Moli dialect of Bahing has etymologically correct forms for '6' (*rukhu*) and '7' (*cūni*), but in the Biguṭar 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').

  •• 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' (§4.2.3.6).

---

6   As we have seen (§3.5.3.4), 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'.

4.1 THE LOWER NUMERALS: 2–5

4.1.1 PROFILE OF NUMBER TWO

TWO *g-ni-s/k

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, Rgyalrong(=Gyarung=Jyarong) 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, for example, Lahu nî < PLB Tone *2. (Compare similar open-syllable derived forms in Loloish words for SEVEN, §4.2.2.4.) The STC, without much discussion (see note 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'.[7] 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[8]
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).[9] There is, however, no evidence to link SIX with ONE, EIGHT with 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.[10])

---

7   See STC #4, and pp.16, 75, 94, 130, 131, 147, 162, 168, 169, 185, 186.

8   Lepcha here exhibits the variational pattern -i- ~ -ya-, that is so well documented for TB as a whole (see STC note 251, p.84; VSTB pp.40-43).

9   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 §4.1.5.

10   For an explanation of this form as an allofam of the Jg. first-person pronoun ŋāi, see Matisoff (1994) 'Watch out for number one', LTBA 17.1.

4.1.1.1 FORMS WITH VELAR PREFIX

Angami (Khonoma) *kena*, (Kohima) *kenie*; Chakhesang and Chokri *küna*; Kezhama *kenhi*; Khoirao *kati* (with denasalisation 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 for example, FIVE, §4.1.4.1.1); 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 (Deng Geman, Digaru Mishmi, Miju and Taraon) show a final nasal 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 palatalising 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 fricativisation of the palatal semivowel: **n- > ny > y > sh*.[11]

The distinctive initials in Karen (Palaychi) *chi* and (Sgaw) *khi* (Jones 1961) might well be preemptive survivals of the velar prefix. Other Karen dialects (Taungthu, Pwo *ni*) simply reflect the prefixless root.

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

4.1.1.2 FORMS WHICH SHOW NO TRACE OF A CONSONANTAL PREFIX

Dulung *a²¹ni⁵⁵*; Rawang *əni*

Apatani *nī*; Abor-Miri, Gallong, Lhopa, Minyong, Tagin *a-nyi*; Padam *a-ni*; Nishi *anni*; Dafla (Das Gupta) *anyi ~ ain*; Milang *ne*; Monpa (Cuona) *nAi²³*[12]

Boro *nôi*; Kokborok *nuy*

Chepang *nis-jo?*; 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 (= Minchia) *ne²*; Pumi *ni²³*; Qiang *nyi⁵⁵*

---

11 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).

12 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.

Ao (Chungli) *ana*; Kimsing *anai*; Chang *nyi*; Konyak *i* (with palatalisation 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 grammaticalised into a plural or collective suffix, for example *gwì nī* 'the dogs'. The independent Jingpho numeral for TWO is the mysterious *lekhoŋ* (§4.1.1.5).

4.1.1.3 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: Hmar *pahni*, Lushai *pahnih*, Vaiphei *pani* (§5.4.3).

4.1.1.4 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 *ʔrit* or *ʔryat*, 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:[13]

Hayu *nak-pu* (for humans), *naʔung* (for non-humans) (Michailovsky 1988:167); 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 amd Bahing forms certainly seem related to the *-tsing* in Sharchop and Monpa, though their wider affiliations are still unknown.

4.1.1.5 UNUSUAL FORMS

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

Jingpho	*ləkhông*	(no known etymology)
Khaling	*sa:h-pu*	(apparently contaminated by *suh-pu* '3')
Bumthang	*zon*	

---

[13] 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.

This Bumthang form is virtually identical in appearance to Sharchop *zon* and Central Monpa *zum* 'SEVEN', which we analyse as deriving from **s-nis* via apocope of the final (§4.2.2.4). 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'.

4.1.2 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 (§4.1.2.2). Most TB languages, however, now show 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.[14] Many daughter languages (including Chinese) have *-a-* vocalism, however, and perhaps this cannot be dismissed as secondary in all cases – that is we may ultimately be forced to recognise **-u-* ≍ **-a-* variation in this etymon, a pattern already noticed in a number of another etymologies (for example, 'fragrance' PTB **b-suŋ* ≍ **b-saŋ* (STC #405)).[15]

4.1.2.1 FORMS WITH VELAR PREFIX

Written Tibetan *gsum*

Chulikata *ka:sh* (with preemption via apocope of the rhyme); Deng Darang *kə²¹sung⁴⁵*; Deng Geman *kɯ²¹səm⁵³*; Idu (Sun 1983) *ka³¹song³⁵*; Miju Mishmi *ksam*; Taraon *ka:sa:ng*

Kezhama *katsii*; Khoirao *kathum*; Mao *kosii*; 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*

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

---

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

15 See STC note 486 (p.186–187). The Middle Chinese vocalism in this root is generally considered to be 'irregular' (cf. STC note 436, p.162). See also, for example, **s-riŋ* ≍ **s-raŋ* 'thousand', §3.5.4.7(b).

4.1.2.2 FORMS WITH NO OVERT TRACE OF A CONSONANTAL PREFIX

Miri (=Mising), Gallong, Minyong, and Tagin *a-um*; Abor (=Padam) *a-ŋum*; Aka *zu*; Apatani *hĭ*; Dafla (=Nishi) *(a-)om, um*; Bokar *a-hum*; Lhopa *afium*[16]; 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, Puiron, and Tiddim *thum*

Rawang (Barnard 1934) *ətsum*; Dulung *a²¹ suum⁵³*

Boro and Kokborok *tham*

Ersu *si⁵⁵*; Pumi *sãu²³*; Qiang *tshi⁵⁵*; Bai (= Minchia) *sa¹*

Chepang *sum-jo²*; Gurung *soq*; Hayu *tshuk-pu* (for humans), *tshu²-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 *šê² ~ šɛ̃*, etc.

The glottal stop in Lahu is secondary, a junctural feature that arose to separate THREE from FOUR in counting, since Lahu *ɔ* '4' has a vocalic onset.[17] For similarly secondary final laryngeals in this etymon, compare the Hayu and Khaling forms just cited.

4.1.2.3 FORMS WITH MISCELLANEOUS SECONDARY PREFIXES

Several Chin languages have developed a general numeral prefix, *pa-*, used with all the primary numerals: Hmar, Lushai and Vaiphei *pathum* (§5.4.3).

A couple of Naga languages have developed a fully syllabic CVC- prefix with the lower numerals (§5.5.1.2): 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, mə lī, 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 §4.1.3; §4.1.4.

---

16   The voicedness of the *fi* 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 §5.5.1.2) might have caused the voicing. J.T. Sun (1993:110) sets up Proto Tani **fium* for this root.

17   See §4.0.1(e).

4.1.3 PROFILE OF NUMBER FOUR[18]

FOUR *b-liy or *b-ləy

The following forms are cited in STC #410:[19]

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); Mikir phli.

The weak root-initial lateral in this etymon offered little resistance to the preemptive propensities of the prefix (§4.1.3.1).

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, puɯ-lje (also pe, with preemption); Monpa (Cuona) pli[53], (Dubey) blee, (Central) b(i)ci ~ p(i)ci; Serdukpen bi:si[20]

Ao (Chungli) pezɯ, (Mongsen) phüli; Kezhama pedi; Kimsing balai; Konyak peli[21]; 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.jo?; Hayu b(l)i?ung; Gurung plihq; Tamang and Thakali plih; Thulung Rai blə; Sharchop pshi; Dumi balɨkpi.

In several Chin languages (for example, Hmar, Lushai, Vaiphei pali) the pa- prefix is secondary, used as a general prefix with all the numerals. See §5.4.3.

4.1.3.1 FORMS 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[22]

Khaling bha:el (alongside bho:m '5', also with preemption); Kanawari pii; 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.)[23]

---

18   This numeral plays interesting conceptual roles in some TB languages. We have seen how the Boro system is basically quaternary (§3.3.2[C]); many AMD languages express EIGHT as a multiplicative formation '4 x 2' (§4.2.3.7).

19   See STC pp.33, 61, 88, 91, 94, 104, 111-112, 131, 152, 158, 171-172, 180, 196.

20   J.T. Sun (1993:124) reconstructs PTani *pri.

21   This is the only Konyak numeral that carries a prefix. W.T. French (1983:492) reconstructs Proto Northern Naga *bələy.

22   Many of these AMD forms have acquired a secondary vocalic prefix, after the preemption.

23   This point has been much discussed in the literature, sometimes with acrimony. See STC, p.60.

The same development has taken place in the moribund Ugong language, where a more conservative form *pli* now varies with *pi*.[24]

4.1.3.2 FORMS WITH NASAL (NOT STOP) PREFIX: **m-ləy* < **b-ləy*

As with FIVE (§4.1.4; §4.1.4.2), many languages have a labial nasal instead of a labial stop prefix with FOUR. STC regards this nasalisation 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*

4.1.3.3 FORMS WITH LABIAL FRICATIVE PREFIX:

> Aka (= Hruso) *fi-ri*; Lepcha *fəli*

4.1.3.4 FORMS THAT DO NOT OVERTLY REFLECT ANY CONSONANTAL PREFIX

Karen:	Taungthu *lit* (with suffix; §4.0.1(d)); Pwo *li*, *li?*; 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*

4.1.3.4.1 FORMS THAT INDIRECTLY REFLECT A CONSONANTAL PREFIX

> PLB **hləy²* > WB *lê*; Lisu *li⁵⁵*; YI (Dafang) *ɬi³³*; Naxi (Lijiang) *lu³³*; MPI *li⁶*; Lahu *ɔ̂(n)²⁵*; Akha *ø̀*; Bisu *ha*; Phunoi *hàn* (with rhinoglottophilia), and so on.

4.1.3.4.2 WITH REPLACEMENT OF THE ROOT-INITIAL LATERAL BY A STOP[26]

> Angami (Khonoma) *da*, (Kohima) *die*; Chakhesang *daa*; Chokri *da*

4.1.3.5 WITH FUSION OF PREFIX AND INITIAL TO A SPIRANT/AFFRICATE

> Naxi (Moso) *ẓv³³*

> Jirel *syi*; Sherpa *ji*; Sikkim Bhutia/Dzongkha *zyi ~ syi*

---

24   This Burmish language, formerly known by the misnomer 'Kanburi Lawa', has been rediscovered by D. Bradley. See, for example, Bradley (1978).

25   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 nasalisation of the vowel is due to 'rhinoglottophilia' after the zero-initial (see Matisoff 1975).

26   For a treatment of the relatively widespread phenomenon of *d* � *l* interchange in ST/TB, see Matisoff (1990b), 'The dinguist's dilemma'.

Ersu $z o^{33}$; Pumi (Taoba) $z\mathit{v}^{55}$, (Jinghua) $\mathit{3\varepsilon}^{55}$; Qiang (Taoping) $d\mathit{31}^{55}$, (Mawo) $g z \partial$; Ergong $w z \varepsilon$; Muya $z \underline{\mathit{v}}^{35}$; Queyu (Yajiang) $z i^{35} t\mathit{c\tilde{a}}^{53}$; Guiqiong $ts1^{55}$; Namuyi $z1^{33}$; Shixing $\mathit{3u v}^{33}$

Bai (Dell 1981) su^{2}, (Dali, Jianchuan) $\mathit{c}i^{44}$, (Bijiang) si^{44}

4.1.3.6 WITH VELAR PREFIX:

*g-ləy >	Minyong	aki (with preemption)
*g-b-ləy>	Rgyalrong(Zida)	kewdzyi
		(with lenition of the *-b- to -w-)
	Idu Mishmi	kapri
	Taraon	ka:pra:i
	Deng Daȑang	kə²¹pɹəi⁵⁵
	Chulikata	ka:ppi
	Digaro	kəprei

With this last allofam also belong Miju Mishmi *kambran*, Deng Geman $ku^{2l}b\mathit{u}un^{53}$, evidently with a fully syllabicised prefix and secondary nasalisation in both syllables: *g(N)-b-ləy-(N)*.[27]

4.1.4 PROFILE OF NUMBER FIVE

FIVE *l-ŋa ⋇ *b-ŋa

In STC #78[28] 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*[29].

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 (for example, 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.[30]

Since FOUR is also reconstructed with *b-, 4–5 constitutes a *proto-prefix run* (§1.1; §5.5.1).

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

---

27   Compare for example the syllabic prefix *gùm-* that Jingpho developed in its word for 'horse' *gùmrà(ŋ)* < *m-raŋ*.

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

29   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 (§5.4.3).

30   The case is somewhat less strong for positing *b- ⋇ *m- variation at the proto-level for FOUR (§4.1.3.2).

be the widely distributed root *lak (STC #86) 'hand', given the well-known fact that hands have five fingers.[31]

Other TB forms that point to a liquid prefix for FIVE are:

'Old Kuki' *r-ŋa (for example, Rangkhol ringa) (STC, ibid.)

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 (§4.0.1(b)). Compare also the inner lateral prefix in Padam pilngo (§4.1.4.4), and the Miju and Kaman forms in k-l- (§4.1.4.8).

4.1.4.1 FORMS WITH LABIAL STOP PREFIX

Dimasa and Garo bonga

Chepang ponga:-jo⁷

Milang pangu

Dulung puɯ²¹nga⁵³; Rawang (Barnard 1934) hpung-nga

The following Kuki-Naga forms all participate in FOUR/FIVE prefix runs (see §4.1.3; §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

4.1.4.1.1 WITH PREEMPTION OF THE ROOT-INITIAL BY THE LABIAL PREFIX:

*b-(ŋ)a >	Boro	ba³²
	Kokborok	ba³³
	Rengma	pfü

Several forms have been uncovered where the preemption served to compensate for the apocope of the root vowel (see §5.1.3.1):

Khaling (Him.)	bho:m
Puiron (KCN)	pang
Aka (AMD)	phum

---

[31] Compare Proto Austronesian *ka-lima 'five; hand', as well as evidence for a 'hand-based' quinary numeral system in some Himalayish languages (§4.1.5).

[32] Compare the other Boro preemptive form do '6' < *d-ruk.

[33] Kokborok, like Boro, also shows preemption in its form for '6', dɔk.

4.1.4.1.2 WITH SECONDARY LABIAL PREFIX

A few Chin languages have developed *pa-* as a universal numeral prefix, used with all the numerals even if they already bear a prefix inherited from PTB (§5.4.3). Naturally these languages cannot be used as evidence for the ancient status of **b-* with this numeral:

Hmar, Lushai, Vaiphei *panga*.

4.1.4.2 FORMS WITH LABIAL NASAL PREFIX

Idu (Sun 1983) *ma³¹nga³⁵*; Mishmi (Dubey) *manga*; Taraon *ma:nga:*; Deng Darang *ma²¹nga⁴⁵*; Chulikata *ma:nga:*; Rgyalrong(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; that is all these words for FIVE participate in 4–5 prefix runs (§4.1.3.2):

Khoirao *manga*; Liangmai *mangiu*; Lotha *mungo*; Maram *mingu*; Meithei, Meluri *manga*; Mzieme *mengei*; Ntenyi *münga*; Pochury *mnga*; Sangtam *münga*; Zeliang *mengei*; Zeme *mengeu*

4.1.4.2.1 WITH PREEMPTION OF THE ROOT-INITIAL BY *m-*

Bai (Dali) *mu³*

4.1.4.3 FORMS WITH LABIAL SPIRANT PREFIX:

Lepcha *fəngo* (alongside *fəli* '4')

4.1.4.4 DOUBLY PREFIXED FORMS:

*b-l-ŋga > Padam/Abor *pilngo* (see LSI III.1, p.622); Shimong Adi *pi-ri-ŋo*

4.1.4.5 FORMS THAT SHOW NO OVERT TRACE OF A CONSONANTAL PREFIX:

Dzongkha/Sikkim Bhutia *nga*; Gurung *nga:hq*; Jirel *nga:q*; Kaike *nga:*; Kanawari *nga*; Sharchop *nga*; Newari *nya:-gu:*; Sherpa, Tamang and Thakali *nga:q*; Thulung Rai *ngo*; Dumi *ŋo*; Kulung *ŋa-chi*; Limbu *ŋa-si*

Chang *ngau*; Konyak and Phom *nga*; Wancho *aga* (with denasalised initial)

Gangte, Kuki, Paite, Tiddim *nga*

Monpa (Motuo) *nga*; Apatani *ngo*; in this category we may include other AMD forms with vocalic prefix, for example, Abor-Miri, Dafla, Gallong, Minyong, Tagin *ango*; Nishi *a:ngo*; Lhopa *ongo*; Bangni *u-ŋu*

Ersu *nguar³³*; Muya *ŋa³⁵*; Queyu *ŋua-tɕã⁵³*; Guiqiong *ŋɛ̃³⁵*; Namuyi *ŋa³³*

Proto Lolo-Burmese **ŋa²* > WB *ŋâ*; Lahu *ŋâ*; Lisu *ŋwa*; MPI *ŋo²*; BISU *ngà ~ hà*; Phunoi *ʔàn* (with rhinoglottophilia), and so on.

Bai (Jianchuan, Bijiang) ηv^{33}

Karen (Pa-O) *ngat* (with suffix); other Karen dialects have lenited the nasal to a palatal semivowel: Pwo *jɛ, jaiʔ*, Palaychi and Sgaw *jɛ*

4.1.4.6 WITH SPIRANTISED INITIALS:

In several Qiangic languages the velar nasal has become a velar or 'uvular' fricative:

Pumi *ɣuã*; Qiang *ʁua³³*; Shixing *ɦã*

4.1.4.7 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*

4.1.4.8 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' (§4.0.1(c); §4.2.1.8).

4.1.5 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:[34]

Bantawa (Raniṭar dialect): *chuk* 'HAND', *ūkchuk* 'FIVE' ('1 x 5'), *hũachuk* 'TEN' ('2 x 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 in §4.2.0) which make it look quite a bit like Khmer.

We have noted (§4.1.1) 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,[35] they at least demonstrate that the idea of counting by fives still occupies a niche in TB conceptual space.

---

34   See the charts in Gvozdanović (1985:135-136). As noted above (footnote 31), such a constellation of ideas is also found in Austronesian (PAN **ka-lima* 'hand; five').

35   Such is the opinion of Gvozdanović (1985:137).

4.2 THE HIGHER NUMERALS: SIX TO NINE

4.2.0 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 x 20) + 10') through *quatre-vingt-dix-neuf* '99' ('(4 x 20) + 19').

TB languages also provide many examples of these phenomena, some of which have already been mentioned:

4.2.0.1 MULTIPLICATIVE FORMATIONS

[A] In the quaternary system of Boro (§3.3.2(c)), EIGHT is expressed as '4 x 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 (§4.2.0.3).

[B] Many AMD languages have similar multiplicative expressions for EIGHT, for example, Apatani *a-pi* '4', *nyi* '2', *pih-nyi ~ pryih-nyi* '8' (see §4.2.3.7).

[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, *reṭkapok* means '4'; but in Chhinamakhu subdialect of Bhojpur it means '8' (Gvozdanović 1985:136).

[D] As we have seen (§1.2.1), 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 x 2) + 1'; *lwīswá tə-* 'nine' '(4 x 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 x 2'). At this same period there was also a numeral *uː-ning* '5', that looks like it was influenced by '6'.

4.2.0.3 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 and 90). Examples may readily be found in Indo-European languages. Latin has subtractive forms for '18' and '19' (*duodēvīginti* '2 from 20' and *undēvīginti* '1

from 20').[36] In Sanskrit, pairs of alternative forms exist for '19' and the other higher numbers ending in nine. These are 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	*aṣṭadaśa*
19	*navadaśa* ('9 + 10') or *ūnaviṃśati* < *ekonaviṃśati* ('one diminished 20' (pers.comm. 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 (§3.3.1.6), as well as subtractive ways of expressing the odd round numbers in TB vigesimal systems, for example, in Chang (§3.5.2.4) and in Dzongkha (§3.5.3.4[C]). As far as the basic unit numerals themselves are concerned,[37] the best examples of subtractivity in TB are to be found in Mikir (§3.5.2.3; §5.4.2.1) and Meithei (§3.2.3.2; §5.4.4.5):

	Mikir	Meithei
ONE	*isi*	*ama*
TWO	*hini*	*ani*
EIGHT	*nirkep*	*nipal ~ nipan*
NINE	*sirkep*	*mapan*
TEN	*kep*	*(tara)*

4.2.0.4 ADDITIVE FORMATIONS OF THE BASIC UNIT NUMERALS[38]

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:

Khmer[39]

ONE	*muəy*	SIX	*prammuəy*
TWO	*pii*	SEVEN	*prampii*
THREE	*bəy*	EIGHT	*prambəy*
FOUR	*buən*	NINE	*prambuən*
FIVE	*pram*	TEN	*dap*

---

36   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'.

37   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' (pers.comm., Adam Jacobs 1992); Indonesian *sembilan* 'nine' ('taking one [from ten]' < *ambil* 'take away', *se-* 'one'); *delapan* 'eight' (prob. '[taking] two [from ten]' < *dua* 'two').

38   We have already mentioned additive formations involving the round numbers of vigesimal systems (§3.5.2.4); 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 (§3.3).

39   Huffman (1970:25).

As we have seen (e.g., §4.1.1 above), 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 §3.3.2[C]), '6' and '7' are expressed as '(4 x 1) + 2' and '(4 x 1) + 3', while '8' is '4 x 2' and '9' is '(4 x 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' (§4.2.2.9).[40] However, it is in the 'endangered' numeral systems of certain Kiranti languages that the most elaborate additive sets of basic numerals have been reported, for example, in the Raniṭar dialect of Bantawa:

Bantawa (Raniṭar dialect)

ONE	ūk-ṭa	SIX	bhan-ka-chuk
TWO	hūa-ṭa	SEVEN	bhan-hū-chuk
THREE	sum-kaṭ	EIGHT	bhan-sum-chuk
FOUR	reṭ-kaṭaṭ	NINE	bhan-reṭ-chuk
FIVE	ūk-chuk	TEN	hūa-chuk

As we have seen (§4.1.5 above), FIVE is here expressed as '1 x 5/HAND' and TEN is '2 x 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 *ūk* '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:

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 x 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' (that is the number of fingers on both hands), *hongbi* 'minus'). Logically enough, TWENTY is simply the compound 'HANDS and FEET' (*lang* 'foot')!

---

40   This is quite comparable to Russian *semj* '7', *vosemj* '8'.

4.2.1 PROFILE OF NUMBER SIX

<div align="center">

SIX *d-ruk* (STC)/*d-k-rok* (JAM)

</div>

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, that is 'prefixal *d-r-' > Written Burmese *khr-* (versus 'cluster *dr-' > WB *khy-*), but this is not very convincing.[41] In my view, both a velar and a dental element must be recognised at the PTB level.

4.2.1.1 FORMS THAT REFLECT A DENTAL (BUT NO VELAR) BEFORE THE -r

STC #411 cites the following 6 forms:

WT drug; Kanauri *ʈug*; 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 *ʈhuk*; Sherpa *ʈuk*; Sikkim Bhutia *ʈuk*; Gurung *ʈuhq*; Tamang *ʈuːh*; Thakali *ʈuh*; Kulung *tuk-chi*; Limbu *tuk-si*; Chamling *tukara*

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 §5.4.4):

Lakher *charu*; Liangmai *charuk*; Mao *choro*; Nruanghmei *ciinei*; Rengma *tsaro*; Sema *tsogho*

4.2.1.2 FORMS THAT REFLECT A VELAR (BUT NO DENTAL) BEFORE THE -r:

*kruk	>	Written Burmese	khrok
		Achang	$xʐo^{ʔ55}$
		Langsu	$khjauk^{55}$
		Zaiwa	$khju^{ʔ55}$
		Nusu	$kh^{ʲ}u^{53}$
		Naxi (Yongning)	$khə^{ʲ13}$
		Jingpho	$kru^{ʔ55}$
		Trung	$k^{ʲ}lu^{44}$

---

41 STC #411, and pp.41, 45, 75, 76, 82, 83, 88, 94-95, 114, 115, 116, 141, 146, 154, 161, 162, 171 and 182.

Newari	*khu-gu:*
Monpa (Cuona)	*kro?⁵³*, (Dubey) *gro*
Kom Rem	*karuk*
Puiron	*keruk*[42]

J.T. Sun (1993:132) reconstructs Proto Tani **krə*, on the basis of Apatani *xrɟɯ*, Bengni *a-kɟɯ*, Bokar (=Adi=Lhoba) *a-kɯ*, Padam (=Abor) *a-ke*, Gallong *ak-kə*, Nishi (=Dafla) *ax* and 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*.

4.2.1.3 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 TB family, so is there with SIX:

(a) **d-k-ruk* > Deng Darang   *tə⁴¹x¹o⁵⁴*
    Taraon   *ta:hro*
    Idu   *tarho*
    Mishmi   *tiaro*

In this category also belong Loloish forms with velar initials that reflect Proto Loloish **LOW*-stopped tone (for example, Lahu *khɔ?*). 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 *khɔ?*; Akha *ko.*; Ahi *tṣhu?⁴⁴*; Sani *khu?²²*; Hani (Gao Huanian 1955) *khu²¹*; Hani (Hu and Dai 1964) *ku²¹*; Lisu (Fraser 1922) *hchaw⁶*; Luquan (Ma Xueliang 1949) *tš'u⁵⁵*; Nasu (Gao Huanian 1958) *tšu?⁴⁴*; Moso *tṣ'wa⁵⁵*

(b) **k-d-ruk*

The Rgyalrong dialects reflect a double prefixation in the reverse order, **k-d-ruk*, with the velar being of demonstrably more recent origin. (All the Rgyalrong numerals from 2 to 9 have the velar prefix *ke-*: §5.2.) Rgyalrong forms for SIX include: *keʈa* (Zida dialect); *katruk*, *truk*, *keʈo*, *ki-trog*, *kə-tshuo*, *koco*, *ktru*, *ku-tok* (cited in Nagano 1984); and *kətsok* (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 *tṣhu⁵⁵*, Pumi (Taoba) *tṣhu³⁵*, Muya *tɕhyi³⁵*, Queyu *tṣhō*, Shixing *tɕho⁵⁵*) but Qiang

---

[42] Kom Rem and Puiron have a velar prefix with other high numerals as well: Kom Rem and Puiron *karet* '8'; Puiron *kakwa* '9' (see §5.4.4.2, §5.4.4.3).

(Mawo) *xtʂə*, (Taoping) *xtʂu³³* (ZMYYC #916) seem clearly to point to a complex proto-consonant group where the first element was a velar, **k-d-ruk.*[43]

Garo (Momin) *gedok* is another reprefixed form, paralleled by *gesa* '1', *gegni* '2', *gesni* '7'.

4.2.1.4 FORMS WITH INITIAL RESONANT, WITH NO OVERT SIGN OF A PREFIX

In this category belongs Chinese itself, with the Old Chinese form reconstructed as **ljôk* in GSR #1032 (see §1.2.6).

TB languages that also reflect the bare root **ruk* or **rok* are scattered around the family:

(Himalayish)	Kaike *ru*; Khaling *raː*; Thulung Rai *ru*
(Northern Naga)	Chang *lak*; Konyak *wok*; Phom *vok*
(Abor-Miri-Dafla)	Aka (LSI) *rieh*; Chulikata *ahe*; Sulong (ZMYYC #916) *yək³³*
(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*[44]

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,[45] so that these forms may also be referred back to the simple unprefixed allofam **ruk*.

4.2.1.5 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*.

4.2.1.6 MISCELLANEOUS FORMS WITH LABIAL INITIALS

Ergong (Qiangic group) *wtɕhau* 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 generalised with all the numerals (§5.4.3).

---

43 A couple of other Qiangic languages have simple velar stop initials (Namuyi *qhu³³* and Guiqiong *khɔ³³*).
44 Apparently with assimilation of the final stop to the roundedness of the vowel.
45 As convincingly demonstrated in Solnit (1979). See also §4.2.2.1.2.

A few other languages have forms with initial *f-*, for example, Yi (Xide) *fu*[55] and Bai (ZMYYC) *fv*[44], but, as is usually the case in TB, this labiodental consonant is a secondary development from an earlier sequence of consonant + resonant.[46]

4.2.1.7 ISOLATES

Milang *sap* is quite enigmatic. To it we may perhaps compare Miju *katam*/Deng Geman *kw²¹tam*[53]. The stop/nasal interchange is paralleled in EIGHT (Miju *grin*; see §4.2.3.6). 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 x 2'), like a well-established group of forms for EIGHT that derive from '4 x 2' (§4.2.3.7). 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'.

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 – that is 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 §4.0.1.)

4.2.2 PROFILE OF NUMBER SEVEN

SEVEN PTB **s-nis*[47] / PLB **s(n)i-t* (TSR #128)

As we have already pointed out several times,[48] 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.[49]

4.2.2.1 FORMS OVERTLY REFLECTING THE **s-* PREFIX, WITH RETENTION OF THE ROOT-INITIAL

STC lists Kanawari *stis*[50] (with denasalisation of initial), Rgyalrong *kĕnĕs* ~ *kĕsnĕs*, Garo *sni*, and Jingpho *sənìt*. To these we may add:

---

46 Lahu *f-*, for example, comes from PLB **hw-* or **?w-*. See for example, Matisoff (1979).
47 See STC #5, and pp.16, 79, 93-94, 130, 131, 147, 162, 168, 169, 185 and 186.
48 For example, §4.0.2; §4.1.1; §4.1.4 and §4.2.0.
49 See Matisoff (1985a:432), 'Out on a limb: *arm*, *hand*, and *wing* in TB'.
50 Given as *stish* in Joshi (1909:2-3).

Dulung (= Trung) *suɯ²¹ nyit⁵⁵*

Boro *shni*; Dimasa *sini*; Kokborok *chini*

Newari *nhae-gu:*

Written Burmese *khu'-hnac* (alongside *hnac* '2'); Atsi *nˀyit*; Maru *nˀat* (the glottalised 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 to 9)

Qiang (Mawo) *stə* (with denasalisation of the root-initial); Pumi (Taoba) *n̥ĩ³⁵*; Ergong *snie/sŋie*; Queyu *n̥ā⁵⁵*

4.2.2.1.1 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*

4.2.2.2 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-* (that is 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 *səgiˀ*; 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*,[51] but in my opinion they cannot be ostracised from this etymon. I believe these forms merely show 'rhotacism' – that is a 'liquefaction' of the nasal. We have already noted the regular development of **r* to *g* in many Kuki-Chin languages (§4.2.1.4). 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 favoured 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.

---

[51] 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*'.

4.2.2.3 FORMS WITH NO OVERT TRACE OF A PREFIX: *nis

Monpa (Cuona) *nis^{55}*; Limbu *nuu-si^{52}*

Gurung *ngiq*; Tamang *nyis*; Thakali *ngis^{53}*; Kaike *ne*

Karen[54] (Pa-O) *nə́t*, (Pwo) *nwè*, (Palaychi) *nwìq*, (Sgaw) *nwí*

Konyak *nyit*; Phom and Chang *nyet*

Guiqiong *nɹī^{55}*; Muya *nɹyi^{35}*

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 to 10);

Wancho *anat* (part of a run of *a-* from 5 to 9)

There is a strong tendency to palatalise 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 palatalisation.[55]

4.2.2.4 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/shi̤*; Hani (Gao Huanian 1955) *sɿ^{21}*.

Most Loloish languages, however, have forms pointing to an open syllable under PLB Tone *2, *si^2:

Lahu *šī*; Lisu (Fraser) *shi̤^{5}*, (Jui 1958) *sɿ^{11}*; Luquan *ši^{55}*,[56] and so on.

These latter forms are strikingly parallel to the Loloish words for TWO that also descend from open syllables under Tone *2 (§4.1.1), 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.

---

52 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).
53 Cited as "*nigs*" in CSDPN, an obvious typo.
54 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 FOUR, FIVE, and NINE it is a *suffix* (1979:19; see §1.2.1). This apparent contradiction is resolved by assuming that in 'SEVEN' the suffix had already become "welded" to the root at the PST level (1979:20). For Loloish forms which support the suffixal nature of the *-s* in SEVEN, see §4.2.2.5.
55 This form also shows convergence in rhyme with the next higher numeral, Thulung *let* 'eight'.
56 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 at §4.1.1.

(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' (§4.0.1(c)).

(c) Qiangic

Qiangic has treated the initial consonant sequence **s-n-* of this etymon in a variety of ways (summarised below §5.0), with several languages showing preemption of the nasal root-initial by the prefix:

Namuyi *ʂʅ³³*; Shixing *ʂɛ̃⁵⁵*; Pumi (Jinghua) *xiɛ̃¹³*.

In Shixing and Pumi(Jinghua), the nasal initial has been preserved by becoming 'prosodised' or 'suprasegmentalised' in the shape of vowel nasalisation. The Ersu form *ʃʅ⁵⁵ ŋ̩⁵⁵* shows an alternative evolutionary strategy: disyllabisation via vocalisation 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![57]

(d) Sharchop and Monpa

These two Himalayish languages have peculiar-looking forms for SEVEN that certainly appear related to each other:

Sharchop (=Tsangla) *zon*;[58] Monpa (Motuo) *zum*

There are at least two etymological possibilities here. Either these forms are related to WT *bdun* (§4.2.2.8); 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, that is **s-n-* > **s-* > **m-s-*.

4.2.2.5 AMD FORMS FOR SEVEN WITH A VELAR PREFIX

A number of AMD languages have developed a velar prefix with this etymon:[59]

---

[57] Monosyllabisation of disyllabic compounds is a strong tendency in Qiangic. See Benedict (1983) and Matisoff (1991:493). Compare similarly preempted and apocopated forms for FIVE (§4.1.4.1.1), the Nishi forms for '7' and '8' (§4.2.2.5), and the general discussion (§5.1.3.1).

[58] See also Bumthang *zon* '2', and 'Transvaluation of numerals', §4.0.2.

[59] J.T. Sun (1993:213) reconstructs Proto Tani **kV-nɯt*.

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*[60]

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 (for example, Abor-Miri, Dafla, Gallong, Minyong, Tagin *a-nyi*), or by innovating a new root for SEVEN (for example, Chulikata *joh*, Taraon *wē*, German Deng *nɯn*[53], Idu *īũ*, Miji *myah*, Milang *rangal*; see below).

4.2.2.6 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:

Rgyalrong(Zida dialect) *keṣnyit ~ keṣnyis ~ 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 §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'.

4.2.2.7 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) *nɯn*[54] (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 *grɯn*[53] '8'), even though the latter is to be reconstructed with PTB **-t*.[61]

A more far-fetched hypothesis would be to try to relate these forms to WT *bdun* (next section).

4.2.2.8 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': that is Tibetan dialects and a few other closely related Himalayish languages:

---

60 This monosyllabic Nishi form is another illustration of the process of 'prefix preemption 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*, §4.2.3.7); also such forms for FIVE as *phung* (§4.1.4), and some Qiangic words for SEVEN, §4.2.2.4(c). See §5.1.3.1.

61 As we have seen (§4.2.2.4), Thulung has also analogically levelled the rhymes of these two numerals, but in favour of a final stop rather than a nasal: *yet* '7', *let* '8'.

Lhasa *tüün*; Sherpa *din*; Jirel *duin*; Sikkim Bhutia (=Dzongkha) *düin* (the vowel in Chhewang Rinzin's speech is *ɨ*); Baima *de[13]* (pers.comm. Sun Hongkai 1991); also probably Khaling *ta:er*.

While most dialects of Rgyalrong have doubly-prefixed versions of the normal root **s-nis* (§4.2.2.6), the Hanniu dialect has *daen*,[62] which goes well with these Bodish forms (especially with Khaling), and seems to be an obvious loan from Bodish into Rgyalrong.

Monpa (Motuo) *zum* and Sharchop *zon* have some resemblance to these forms, but can equally well be considered monosyllabicised versions of the normal root **s-nis* (§4.2.2.4d).

Likewise, Kaman *nun[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 (§4.2.2.7).

That leaves as the only conceivable candidates for cognacy with WT *bdun* the following AMD forms:

Taraon	*wẽ*
Deng Darang	*weng[54]*
Idu (Talukdar et al. 1962)	*ĩū* (alongside *inyū* '8')
Idu (Sun 1983)	*i[55]ɦong[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 palatalisation to **b-d(y)un* (compare 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.[63]

4.2.2.9 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.[64]

Since Mikir also expresses EIGHT and NINE subtractively on the basis of TEN (§4.2.3; §4.2.4), it has no monomorphemic inherited numerals between SIX and TEN.

4.2.2.10 ISOLATES?

There remain a few forms for SEVEN which cannot be related to anything else in the current state of our knowledge:

---

62 Data from Nagano (1984).
63 For a somewhat analogous sound-change, compare WT *dbus* 'head; central' > Lhasa *üü*.
64 See §4.2.0.3.

(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 *mja?*; MIJI (Simon 1979) *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 compare also
     Lotha *ti-ingya* '50' (§3.5.2.2).

4.2.3 PROFILE OF NUMBER EIGHT

<div align="center">EIGHT *b-r-gyat ⋊ *b-g-ryat[65]</div>

Key forms for justifying this doubly prefixed reconstruction are WT *brgyad* and Jingpho
mətsát.[66] To these we may add Rgyalrong (Zida) *warzhe(t)*, with *wa-* also < *b-.[67] 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 neighbours in linear order.[68]

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 (§3.2.3.2; §4.2.0.3); and in many Abor-Miri-Dafla languages it is expressed
multiplicatively as '4 x 2' (§4.2.3.7). In Lepcha, '8' and '9' have undergone an etymological
flipflop (§4.0.2; §4.2.4).

The prefixal behaviour 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 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 (note 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 §5.2ff.).

---

65   See STC #163, and pp.35, 45, 54, 57, 74, 88, 95, 96, 131, 141, 144, 161-162, 179 and 191. I have
     reconstructed this etymon for Proto Lolo-Burmese as *ʔrit^L (TSR #171).
66   Other examples of Jingpho *mə-* < *b-* include *məlī* 'four' <*b-ləy, *məŋā* '5' < *b-ŋa.
67   EIGHT is the only Rgyalrong numeral from 2 to 9 not to have the prefix *ke-* (see §5.2). Undoubtedly its
     double prefix was weighty enough to allow it to escape the steamroller of the velar prefix run.
68   See §3.5.4, and STC note 148 (p.45).

4.2.3.1 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*

4.2.3.1.1 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; compare Jg. *mətsát*)

4.2.3.2 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* (§5.5.1.2)

(b)   *p-* > Hmar *pariet*; Lushai *pariat*

This *pa-* is an innovative prefix that goes with all the numerals in a few Chin languages (§5.4.3), and has nothing to do with the PST/PTB **b-* posited for EIGHT in particular (which is reflected, for example, in the labial initial of the preemptive Chinese cognate).[69]

(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', *kăkŭ* '8', *kăkyót* '9' and *kăti* '10', wherein '8' and '9' seem to be reversed etymologically (§4.0.2; §4.2.4).

(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*?).

---

[69]   Forms from certain Bai dialects, e.g. Jianchuan and Dali *pia⁴⁴*, seem clearly to be loans from Chinese. The Bijiang dialect of Bai has a form with initial affricate (§4.2.3.5), which looks like an independent reflex of the PST/PTB etymon.

4.2.3.3 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', *tiiku* '9', *thüre* '10').

Qiang (Mawo) *kha¹* (ZMYYC #1291) looks like an apocopated and metathesised form, and resembles several other reflexes with final *-r* or a rhotacised vowel (for example, Nusu *ṣa¹⁵³* and Kokborok *char*, §4.2.3.5).

4.2.3.4 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 (Agami Singh Rai 1944) *let* (prob. < **ryat*); Khaling *ri*; Kanawari *rai*

Proto Lolo-Burmese **ʔrit* ≠ **ʔryat* > WB *hrac*; Lahu *hí*; Akha *yeh*.; Ahi *xi⁴⁴*; Sani *he²²*; Hani (Gao Huanian 1955) *xae²¹*; Hani (Hu and Dai 1964) *še²¹*; Lisu (Fraser) *h'i⁶*; Luquan *ʔhən⁵⁵*; Nasu (Gao Huanian 1958) *xɛn³⁴*; Achang *çet⁵⁵*; Zaiwa *ʃit⁵⁵*; Maru *ʃɛʔ⁵⁵*; Anong *çen⁵⁵*; Naxi (Lijiang) *xo⁵⁵*, (Yongning) *xu¹³* (see TSR #171)[70]

Jinuo *xɛ⁴⁴*; Tujia *jie²¹*

The aberrant and isolated Sulong language of Arunachal Pradesh has a form with liquid initial, *la³³* (ZMYYC #1291).

4.2.3.5 FORMS WITH AFFRICATE/SIBILANT INITIALS THAT COULD REFLECT EITHER **gy-* OR **ry-*

Ao (Mongsen) *tsit*; Chang *sat*; Maring *chot*; Phom *šət* (alongside *pa-šet* '18', *shə* '9'); Wancho *achat*; Nocte *i-sat*; Tangsa (Yogli) *təchat*, (Moshang) *tachat* (with dental prefix: §4.2.3.1); Konyak *tet* (with preemption by the dental prefix)[71]

Newari *cya:-gu:*

Most Qiangic languages have sibilant spirants or affricates: Ergong *ʒyiɛ* (< **r-gy*); Ersu *ʒ ɿ⁵⁵*; Muya *çyʙ⁵³*; Pumi (Jinghua) *ʂuɛ²³*, (Taoba) *çyʙ³⁵*; Qiang (Taoping) *tʂʰe³³*; Queyu *çyɛ⁵⁵*; Shixing *çyi⁵⁵*. A couple of Qiangic languages have forms with initial semivowel or *h-*: Namuyi *hɿ³³* (with rhinoglottophilia); Guiqiong *je⁵⁵*.

Bai (Bijiang) *tçua⁴⁴*. Other Bai dialects have apparent loans from Chinese; cf. Jianchuan and Dali *pia⁴⁴*, §4.2.3.2.

Dulung *çăt⁵⁵*; Trung *šiat⁴⁴*; Nusu *ṣa¹⁵³*

Garo *chet*; Dimasa *jai*; Kokborok *char* (with the final *-r* apparently due to metathesis: **g-ryat* > **gyar* > *char*); compare the Nusu form just cited, as well as Qiang (Mawo) *kha¹*, §4.2.3.3 above).

---

70 Note the rhinoglottophiliac nasalisation in Luquan, Nasu, and Anong.
71 W.T. French (1983:482) reconstructs Proto Northern Naga **C/V-gyat*, with unspecified vocalic or consonantal prefix.

Karenic: PA-O *sɔ́t*; PWO *xò?*; Palaychi *xó*; Sgaw *xɔ̀?*. Kayah (=Karenni) has a multiplicative form for EIGHT (§1.2.1; §4.2.3.8).

4.2.3.6 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)

Compare also Sunwar *yaan* 'NINE', which is perhaps a transvalued ('upstepped') reflex of this allofam for EIGHT.[72]

4.2.3.7 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 §4.2.3.6, and that we tentatively refer to a new root like **lyoŋ*:

Deng Darang *liɯm³⁵*; Idu (Sun 1983) *i⁵⁵liong³⁵* (alongside *i⁵⁵ɦoŋ⁵³* '7'); Idu (Talukdar 1962) *inyū* (*-ny-* apparently < earlier *-ly-*; compare also Idu *iū* '7'); Mishmi *ili*; Chulikata (LSI) *ilu:*

4.2.3.8 MULTIPLICATIVE FORMS IN ABOR-MIRI-DAFLA AND ELSEWHERE: 8 = 4 x 2

Many AMD languages have multiplicative compound forms for EIGHT of the structure '4 x 2'):[73]

Abor-Miri	*a-pi* '4', *a-nyi* '2' > *pi-nyi ~ pui-nyi* '8'	
Apatani	*a-pi* '4', *nyi* '2' > *pɯ?-nyi ~ pryɯ?-nyi* '8'	
Lhopa	*api:* '4', *anyi* '2' > *pi:-nyi* '8'	

Similarly: Dafla, Gallong, Padam, Tagin *pine*, Yano *plə-ne*; Minyong *pini*[74]; Nyisu *plin*; Nishi *pin, piin*.[75]

Other multiplicative formations for EIGHT in TB include Boro *zokkay nəy* (§3.3.2[C]; §4.2.0.1), and Kayah (= Karenni = Red Karen) *lwīswá?* (§1.2.1; §4.2.0.1).

---

[72]   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 §4.2.4.6.

[73]   J.T. Sun (1993:125) reconstructs a Proto Tani multiplicative compound, **pri-ñi*.

[74]   Compare 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.

[75]   Nishi and Nyisu show apocope of the final vowel, as also in Nishi *ken* '7' (cf. Padam *kane*).

4.2.3.9 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 §4.2.4.5.)

Limbu *phang-si* (Gvozdanović 1985:162) is also a puzzlement.

4.2.4 PROFILE OF NUMBER NINE

NINE **d-kəw* (=**d-kuw*) ✘ **s-gəw* ✘ **d-gaw*

4.2.4.1 WITH DENTAL OR SIBILANT PREFIX

STC reconstructs only the prefix **d-* for the PTB level, relying especially on WT *dgu* and Nung *tegö*.[76] 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'[77], 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 *toku*; Sangtam *tüku*; Yacham-Tengsa *thaku*; Yimchungru *tuku* and 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, for example, **s-*. The same uncertainty attaches to the origin of the *affricate* prefixes in KCN languages: for example, 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, for example, Kokborok (Barish) *chuku*, Jingpho *jəkhû*. I have already observed in a previous analysis of this etymon[78] 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) *sgiu⁵⁵* and Qiang (Taoping) *xguə³³*.[79] None of these sibilant-prefixed words for NINE is participating in a 'prefix run' – the neighbouring numerals lack such a prefix. We should also mention Aka

---

76   See STC #13, and pp.19, 23, 45, 61, 94-95,116, 131, 134, 154, 162, 185, 188 and 196. "**d-gew*" in the Appendix I (p.202) is a typo for **d-gaw*. STC also recognises a Kuki-Naga variant **d-kwa*, probably reflecting a secondary suffix (< **d-kuw-a*). Similar to the Nung form cited in STC are Anong *duɯ³¹gɯ³¹* and Dulong *duɯ³¹gɯ⁵³* (ZMYYC).

77   See §5.4.4, 'Innovative runs in the higher numerals'.

78   Matisoff (1980:15-17).

79   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.

(= 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 neighbouring numerals, as in Maram *soki* '9' (but also *saruk* '6', *sina* '7', *sachat* '8'); or Zeme *sekui* '9', but also *seruk* '6' and *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 compare also Jinuo *tçy³³*.)

> Bai (Dali and Jianchuan) *tçɯ³³*, Bijiang *tçi³³*

4.2.4.2 WITH VELAR PREFIX

A secondary velar prefix is occasionally found with NINE: Puiron *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; §4.0.2); and Rgyalrong *kəngu* (all Rgyalrongnumerals from 2 to 7 also have a velar prefix).

4.2.4.3 WITH PRENASALISED INITIAL

Several Qiangic languages have forms with prenasalised initials (Ergong *ngiɛ*, Muya *ŋguɯ³⁵*, Ersu *ngɛ³³* and Namuyi *ŋgu³³*), as does the Lijiang dialect of Naxi (*ŋgv³³*).

4.2.4.4 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* and 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 *gɦu*

Neither Lolo-Burmese nor Karenic show any evidence of a prefix:

Lolo-Burmese

> PLB *gəw² > WB *kûi*; Lahu *qɔ̂*; Akha *ɣɤ̀*; Hani (Shuikui) *ɣu³¹*; 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 §1.2.1)

4.2.4.5 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 symbolise by setting up a dummy vowel (*kV-n(y)a-ŋ).[80] 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 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	kinya:ng	
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	(with loss of nasal element)
Dafla	kéya	(with loss of nasal element)
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/German Deng	nən⁵⁵mu⁵³	(with dental rather than velar final)
Miju	nat-mo	(with final stop homorganic to the Kaman nasal)

---

80   J.T. Sun (1993:186) has just independently reconstructed this etymon as Proto Tani *kV-(n)aŋ.

97

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 1972) *thai-ne-nia* 'ninth'.[81]

W. Kayah (Karenic) *nɯə*'[82]

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' (§4.2.3.6).˘

4.2.4.6 SUBTRACTIVE FORMS

In two important languages the word for NINE is formed subtractively on the basis of TEN: Meithei *ma-pan* and Mikir *sirkep* (§4.2.0).

---

81 Compare thai-*daŋ*-nya 'eighth', §4.2.3.8. 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 (§1.2.4).

82 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).

PREFIXAL BEHAVIOUR 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 (for example, 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 §4.2.2). Some languages preserve an overt trace of the prefix (Qiang (Mawo) *stə*, Pumi (Taoba) *ŋī³⁵*, Ergong *snie/sŋie* and Queyu *ŋā⁵⁵*); others have lost the prefix entirely (Guixiong *nī⁵⁵* and Muya *n̥yi³⁵*); while still others show preemption by the prefix of the root-initial (Namuyi *sʅ³³*, Shixing *s̜ɛ̄⁵⁵*, Qiang (Taoping) *ɕiŋ³³* and Pumi (Jinghua) *xiɛ̄¹³*).

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 generalised for all the numerals. Similarly, in cases of reprefixation, the more outer prefix is clearly younger than the more inner one (§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.

5.1.1 VOICING AND VOCALISATION 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.[1] 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, for example, *kə-* versus *gə-*, but there is much variability here. LaRaw Maran (forthcoming) and others (Dai Qingxia et al. 1983; Lon Diehl, pers.comm.) 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, for example, Central Monpa *b(i)ci ~ p(i)ci* '4' (Das Gupta 1968).

5.1.1.1 THE VOWELS OF PREFIXAL SYLLABLES

Some languages show considerable fluctuation in the vocalisation and/or aspiration of the same prefix from numeral to numeral. Thus, Sema *kini* '2', but *küthu* '3'; *tsogho* '6', but *tsini* '7' and *thache* '8', but *toku* '9'. Cross-linguistically, the same prefix may be differently vocalised when attached to the same numeral (compare the fluctuation of the vowel from language to language in the new AMD root for NINE, §4.2.4.5).

5.1.2 CONTAMINATION BY PREFIXES OF NON-CONTIGUOUS NUMERALS

As STC observes (note 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 §3.5.4; §4.2.3.)

5.1.3 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 (§4.1.3, §4.2.1, §4.2.2.4). 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 survival for a prefix, most strikingly in isolated forms in subgroups where prefixes do not generally persist, for example, the labial prefix in FOUR in the Burmish language known as Maru or Langsu (Maru *bìt*, Langsu *pjik³¹* (ZMYYC 1287)). With respect to EIGHT, STC recognises 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?

---

[1]   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.

5.1.3.1 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 reorganisation 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 in §4.1.4.1.1); Chulikata *kāsh* '3' (< **g-sum*; §5.5.1.1); Nishi *ken* '7' (< **k-nit*; §4.2.2.5); Nishi *ax* and Nyisu *a-kr* '6' (< PTani **a-krə*; §4.2.1.2); Qiang (Taoping) *ɕiŋ³³* and Ersu *ʃi⁵⁵n̩⁵⁵* '7' (< **s-nis*; §4.2.2.4(c)).

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: Evans 1993) *han* '12' (< *ha* '10' + *ni:ʔ* '2'); Nyisu *plin* and Nishi *pin* '8' (< **pri-ñi* '4 x 2'; §4.2.3.7); 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.[2] Such prefix runs were already a feature of PTB, as can be deduced even if one strictly adheres to the reconstructions in STC:

> 1 <--> 2 <-->3 **g-tyik, *g-nis, *g-sum*
> 4 <--> 5 **b-ləy, *b-ŋa* (⪰ **l-ŋa*)

Consider the numerals of Rgyalrong (= Jiarong):[3]

ONE	*tšek*
TWO	*kenes*

---

2 See the discussion of 'Mutual influence of numerals', §4.0.1. 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 (Gérard Diffloth pers.comm.). Ives Goddard (pers.comm. 1994) observes that the Proto Algonkian numerals from 1 to 5 all have initial *ny-*, pointing to an original quinary system.

3 My thanks to Professor Chang Kun for these data from the Zida (= Tzuta) dialect. The additional forms offered for comparison are from Nagano (1984).

THREE	*kesoṃ, kesam*
FOUR	*kewdži*
FIVE	*kemŋa*
SIX	*keṭa*
SEVEN	*keṣñiṭ, keṣñis, keṣñes*
EIGHT	*warže(t)*
NINE	*keŋgu*
TEN	*ṣtṣi*
ELEVEN	*ṣətšek*
TWELVE	*ṣənes*
TWENTY	*keneṣ-tṣi*

Thus with reference to PTB, Rgyalrong has not opted to save the prefix in ONE, but has retained it in TWO and THREE; not only that, it has generalised its use all the way up to NINE (broken only by EIGHT). The case of Rgyalrong 10–12 is somewhat different: Here the prefix *ṣ-*, (usually vocalised 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 (§3.4.3).

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*), and so on.

We may now introduce a further terminological distinction: *perfect runs* versus *broken runs*. A *perfect run* is an unbroken sequence of numerals with the same prefix. It may be long (for example, the Lushai *pa-* run) or short (for example, 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!). For example:

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 (for example, the long Rgyalrong *ke-* run above is broken by EIGHT which has a different prefix (*wa-*). If EIGHT were some day to succumb to the analogical pressure of its neighbours, the enlarged *ke-* run would extend all the way from 2 to 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 (§3.2).

Changing from a synchronic to a diachronic perspective, we may speak of *secondary prefixation* or *reprefixation*. To return to our Rgyalrong example, the numerals from FOUR to SEVEN have had a secondary velar prefix superadded to their 'original' ones inherited from PTB:

	PTB	Proto Rgyalrong	Zida
FOUR	**b-ləy*	**k-b-liy*	*ke-w-dži*
FIVE	**b-ŋa*	**k-m-ŋa*	*ke-m-nga*
SIX	**d-ruk*	**k-d-ruk*	*ke-ṭa (ṭ < *-d-r-)*
SEVEN	**s-nis*	**k-s-nis*	*ke-ṣnyes, 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 §3.1.5).

For Old Chinese, Baxter (1985) has suggested that the initial *s-* in *sjəd* '4' arose through the influence of the sibilant in '3' *(t)səm*. (See §1.2.6.)

5.2.1 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, for example, 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:

> *ləŋâi, ləkhôŋ – məsūm, məlī, məŋā*
> *krúʔ*
> *sənìt, mətsát, jəkhû*
> *šī*

The monosyllabic intruders serve to demarcate the string of numerals into manageable units, without making it necessary to homogenise 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 §5.4.4.5.)

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, for example, 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.[4/5]

5.3.1 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 §4.2.2.8):

	Written Tibetan	Lhasa Tibetan	Sherpa	Jirel	Kaike	Dzongkha[6]
ONE	*gćig*	*ci:q*	*cikq*	*dokpei*	*ti*	*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*	*nga:q*	*nga:q*	*nga:*	*nga*
SIX	*drug*	*ṭhuu*	*ṭuk*	*ṭhuk*	*ru*	*ṭuk*
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-ba:q*	*cyuta:m-ba:q*	*chyu-tamba*	*chu-*

WT preserves the PTB velar prefix run in 1–3, reminiscent of what we find in many AMD languages (§5.5.1), 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*.

5.3.2 KHALING

The Khaling language of Nepal has several interesting features in its numeral system:

ONE	*tu*	SIX	*ra:*
TWO	*sa:hpu*	SEVEN	*ta:er*
THREE	*suhpu*	EIGHT	*ri*
FOUR	*bha:el*	NINE	*ghu*
FIVE	*bho:m*	TEN	*tadam*

---

4   Rgyalrong, 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 §1.2.3.

5   Many TB languages of Nepal have lost their higher numerals, replacing them with Indo-European ones from Nepali. These are discussed in §2.1.

6   These forms are from Sandberg's (1895) 'Sikkim Bhutia', now the national language of Bhutan under the name Dzongkha (or Danjongka). They are closely confirmed by the tape of Chhewang Rinzin (pers.comm. 1984); on this tape the vowel of SEVEN sounds like barred-i: *dɨn*.

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 (§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 (§5.1.3.1). 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.

5.3.3 GURUNG-TAMANG-THAKALI AND NEWARI

	Gurung	Tamang	Thakali	Newari
ONE	grihq	ki:h	ṭih	cha-gu(-li)
TWO	ngīhq	nyi:h	ngih	ni-gu(-li)
THREE	sõq	som	som	swā-gu:
FOUR	plihq	plih	plih	pẽ-gu:
FIVE	nga:hq	nga:h	nga:h	nya:-gu
SIX	ṭuhq	ṭu:h	ṭuh	khu-gu:
SEVEN	ngiq	nyis	ngis	nhae-gu:
EIGHT	prehq	preht	preh	cya:-gu:
NINE	kuq	ku	ku	gū-gu:
TEN	cyuq	ci	cyu	jhi-gu:

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 *ṭ* 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 (§4.2.3.5).

All of these languages reflect the 'normal' root for SEVEN, *(s-)nis*, realised as virtually identical to TWO in Gurung-Tamang-Thakali.

5.3.4 KANAWARI (= KANAURI) AND LEPCHA (= RONG)

	Kanawari	Lepcha
ONE	id	kat
TWO	nish	nyăt; nyi
THREE	shum	sam
FOUR	pü	fəli
FIVE	nga	fəngo
SIX	ṭuk	tərək
SEVEN	stish	kə-kyək
EIGHT	rai	kə-kŭ
NINE	zgui	kə-kyót
TEN	sai	kə-ti

The Kanawari numerals, characterised 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*; §3.1.1, §3.1.2.1); 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 (§4.0.2). The form *kyək* for SEVEN remains a complete mystery.

5.3.5 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'e?⁵³*	*thur*	*t'or*
TWO	*nai*	*nʌi²³*	*n(y)itsing*	*ñiktsing*
THREE	*sum*	*sum⁵³*	*sam*	*sam*
FOUR	*blee*	*pli⁵³*	*b(i)si/p(i)si*	*p'i*
FIVE	*lenga*	*le²¹nge⁵³*	*nga*	*nga*
SIX	*gro*	*kro?²³/⁵⁴*	*khung*	*khung*
SEVEN	*nis*	*nis⁵⁵*	*zum*	*zum*
EIGHT	*giet*	*cen ¹³*	*yen*	*jen*
NINE	*dugu*	*tu²¹ku⁵⁴*	*gu*	*gu*
TEN	*chi*	*tçi⁵⁴*	*se*	*se*

- Among the Monpa dialects must be included the language known as Sharchop 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, note 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 (§4.2.1.2). Central and Motuo Monpa have an idiosyncratic word for SEVEN (*zum*), which clearly belongs with Sharchop *zon* (§4.0.2), but whose further affiliations are very much in doubt (§4.2.2.4(d), §4.2.2.8). Dubey's Monpa reflects a simple velar initial in EIGHT (§4.2.3.3), but the other dialects have forms (as in SIX) that may reflect a nasal-finalled allofam of the general root (§4.2.3.6).

5.4 NUMERAL PREFIXES IN KUKI-CHIN NAGA

5.4.1 RETENTION OF THE PROTO-SYSTEM (AS CONCEIVED IN STC)

The only runs recognised 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,[7] 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.

5.4.2 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*

---

[7]   Maring *khat* '1' is non-prefixed, a different etymon from *g-tyik. See §3.1.2, §3.1.4.

TWENTY	*sau-chie* 20 x 1	*ha*		*ta*
HUNDRED	*sau-ngau* 20 x 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).

- 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 (§5.4.3).

[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', and so on.

- Note the distinctive **kat* for ONE (§3.1.2) and **som* for TEN (§3.2.3.1).

- Final *-k* in '6' becomes labial *-p* in Gangte, Kuki and Thado, probably via assimilation to the rounded vowel.

- The only numeral to preserve a prefix is SEVEN.[8]

- EIGHT is from **gyat* (§4.2.3.3).

- NINE has peculiar vocalism, pointing to an **-a* suffix. (See also Lushai, §5.4.3.)

- There are distinctive roots for THOUSAND and MYRIAD (fortunately available from contexts in the *Book of Revelation*). See §3.5.4.7(b).

---

8   By coincidence, SEVEN is also the only dissyllabic English numeral!

5.4.2.1 ADDITIVE AND SUBTRACTIVE DEGENERATION OF THE PROTO-SYSTEM

Mikir

ONE	*isi*
TWO	*hini*
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 (§4.1.1.3), 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 §4.2.0.4.

- 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 §4.2.0.3. "*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).

5.4.3 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
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 generalised a prefix *pa-* to all the numerals from 1 to 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 generalisation 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 (§5.4.2(a)). Like Lushai, however, it has generalised a prefix (this time *a-*) for all the numerals 1–9.[9]

5.4.4 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*; that is whether they completely or only partially segment the numerals into consecutive sets. As always, however, some systems are idiosyncratic and resist classification (for example, Kom Rem, §5.4.4.3).

A. *Non-exhaustive segmentation*

5.4.4.1 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-

---

9 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 and 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 §5.4.3) which have generalised a single prefix for all the numerals: in the latter the original prefixes have been *replaced* (except in SEVEN), while in Lakher the new generalised prefixes are *superadded* to the 'inner' prefix (for example, *sa-pangaw*, *sa-charu*, *sa-sari*, *sa-chari*).

5.4.4.2 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 reanalysed as the initial consonant of the resulting monosyllable (see §5.3.1).

- Angami and Chokri generalise a dental prefix for 7–9; Puiron retains the old **s-* in SEVEN, but develops a velar run for 8–9.

5.4.4.3 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 §5.4.4.6 below.[10] 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 *lŋa*.[11] 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 three-way (ternary, tripartite) or four-way (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.

5.4.4.4 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 (§3.2.3.3, §3.2.3.4).

[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; taʔa*

- SIX has the *t-* prefix.
- The highest run includes 6–10.

---

10   This prefix is reminiscent of the favourite Mikir prefix *ing-*, which occurs with many dozen common nouns (but only with one numeral, *ingkoi* TWENTY). See §5.4.2.1.

11   It will be remembered that STC sets up *l-ŋa as a PTB allofam of *b-ŋa (§4.1.4).

112

[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 by, for example, Kezhama (*kele*) (§3.1.5.5).

[C] ZEME-KHOIRAO

	Mzieme	Zeme	Zeliang[12]	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*.

5.4.4.5 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.

---

12   Zeliang is a kind of composite dialect or lingua franca, an acronym for *Ze-me* + *Liang-mai*.

	Lotha	Yacham-Tengsa	Sangtam	Ao Mongsen	Ao 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. (Compare the concept of the 'monosyllabic breather' introduced above, §5.2.1.) 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*; §3.2.3.4).
- Note the apocope in Ao Chungli TEN. This form bears no relationship to the similar looking AMD root *tel* 'ONE' (§3.1.5.1).

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	Tangsa 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

5.4.4.6 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 (§5.4.4.3).

5.4.4.7 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 to 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

5.5.1 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 generalise 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 generalised backwards to include 1 and/or forwards to include 6: that is a single run from 1 to 5 or from 1 to 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.

5.5.1.1 WHERE THE RUN HAS A VELAR PREFIX

	Idu (Talukdar)	Mishmi (Dubey)	Chulikata[13] (LSI)	Taraon[14] (Chakravarty et al. 1963)	Deng Darang (Sun et al. 1980)
ONE	*khe(ng)ge*	*khege*	*e:khe:*	*khing*	*k'ɯn^{55}*
TWO	*kanyi*	*kani*	*ka:ni*	*ka:ing*	*kə^{21}n^{55}*
THREE	*kasõ*	*kaso*	*ka:sh*	*ka:sa:ng*	*kə^{21}sɯng^{45}*
FOUR	*kapri*	*kapri*	*ka:ppi*	*ka:pra:i*	*kə^{21}pʁəi^{55}*
FIVE	*manga*	*manga*	*ma:nga:*	*ma:nga:*	*mə^{21}ngə^{45}*

In these languages the velar run is only from 2 to 4 (the velar in ONE seems to be the root initial – §3.1.2). 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	*kɯo:/komo:*	*kumo*	*kɯ^{21}mu^{53}*
TWO	*ka:ning/kinnin*	*kinin*	*kɯ^{21}jin^{53}*
THREE	*ka:-sa:m*	*ksam*	*kɯ^{21}săm^{53}*
FOUR	*kambrin*	*kambran*	*kɯ^{21}bRɯn^{53}*
FIVE	*ka-li:n*	*klin*	*kɯ^{21}len^{55}*
SIX	*ka:ta:m*	*katam*	*kɯ^{21}təm^{53}*

Note the characteristic nasal-finalled forms for FOUR (§4.1.3.6), as well as the totally idiosyncratic forms for FIVE and SIX (§4.1.4.7; §4.2.1.8).

[13] = 'Taying Mishmi'.
[14] Virtually identical to 'Digaru Mishmi' (LSI III.1, 623), which has *e:khing, ka:-ying, ka:-sâng, ka:-prei* for 1-4.

5.5.1.2 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*[15]
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*	*anī*
THREE	*aum*	*aɦum*
FOUR	*epí*	*api:*
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*	*a:ngo*
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).

---

[15]   See §4.1.3.6.

For THREE, Nishi (like some other Dafla dialects: see below) has a monosyllabic form with *o*- vocalism, which clearly derives from a disyllabic 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.[16] 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	*a:-kin*	*akkin*	*akhin*	*aking/aku*
TWO	*a:-ni*	*anyi*	*anyi*	*anyi/ain*[17]
THREE	*a:-a:m*	*a-om*	*um*	*om*
FOUR	*a:-pli*	*a-pl*	*apli; appi*	*api*
FIVE	*a:-ngo:*	*a:-ng*	*ango*	*ango*
SIX	*akple*	*a:-kr*	*akke*	*aké*

• Note the apocope in the E. Dafla forms FOUR, FIVE, SIX, paralleled also in its word for EIGHT *pli:n* (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: *ka:nag* '7', *plag-nag* '8' (see §4.2.0).

5.5.2 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 characterised by highly idiosyncratic sets of *higher* numerals, though this trait is shared by many other AMD languages (§4.2).

	Aka (= Hruso) (LSI)	Apatani (Simon 1972)	Milang (Das Gupta 1980)	Serdukpen (Dubey 1983)
ONE	*a*	*kū*	*akan; atel*	*han*
TWO	*kshi*	*nĩ*	*ne*	*n(y)ik*
THREE	*zu*	*hĩ*	*ham*	*ung*
FOUR	*fi-ri*	*pe*	*pe*	*bi:si*
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 denasalised Serdukpen form perhaps reflects a lost *p*- prefix, preserved in Milang; in Aka *phum* the labial stop prefix has become the root-initial via apocope of the root-vowel – that is the final -*m* seems to reflect the original root-initial (§5.1.3.1).

---

16   See Matisoff (1978a:277-278) (note 258).
17   Several AMD languages show a tendency to metathesise the initial consonant and vowel of TWO. An intermediate stage is represented by the syllabic nasal in Deng Darang (§4.1.1.1).

Apatani has a special set of numerals used in counting humans, which seems more conservative with respect to prefix preservation; for example, '3' *hingi*, '4' *pilye* and '5' *yango*.

CHAPTER 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, for example, -i- ⋈ -ya (ONE *tik ⋈ tyak (§3.1.4); TEN *gip ⋈ gyap (§3.2.1); *tsiy ⋈ tsyay (§3.2.2); TEN/HUNDRED *liŋ ⋈ *lyaŋ (§3.2.3)); -ay ⋈ -an (ONE *tay ⋈ tan (§3.1.4)); -u- ⋈ -a- (THREE *-sum ⋈ *-sam (§4.1.2); TWENTY *m-kul ⋈ *kal (§3.5.1.1). We have emphasised the inter-influence of numerals in sequence, first in the context of general variational patterns in ST word families (§1.1.1), then in more specific morphophonemic and semantic terms (§4.0.1–§4.0.2.), finally focussing on prefixal behaviour in numeral sets (§5.2ff.). We have seen how some languages express their higher unit numerals (6–9) in additive, subtractive, or multiplicative formations (§4.2.0).

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* (§1.1.2), 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, for example, 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 (for example, Sherpa, Lepcha, Dzongkha: §3.5.3.4), and even between TWENTY and TWELVE (Chepang: §3.5.3.5). We have noted traces of FOUR- (Boro, Kubhinde Dumi: §3.3.2(c)) and FIVE-based (Bantawa: §4.1.5) systems, which may well prove to represent ancient types of numerical organisation 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, for example, between FIVE and *hand* (§4.1.4–§4.1.5), and FIFTY and *ridgepole* (§3.5.2.2). 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 *lê*, 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 *mà*, functionally equivalent to Mandarin *gè* or Thai *ʔan*.) Roop (1970:62-63) reports a similar Lisu general classifier *lyö[11]*, occurring only after *lyi[55]* '4', which he characterises as a 'suppletive allomorph' of the ordinary general classifier *ma[33]*. In the Qiangic
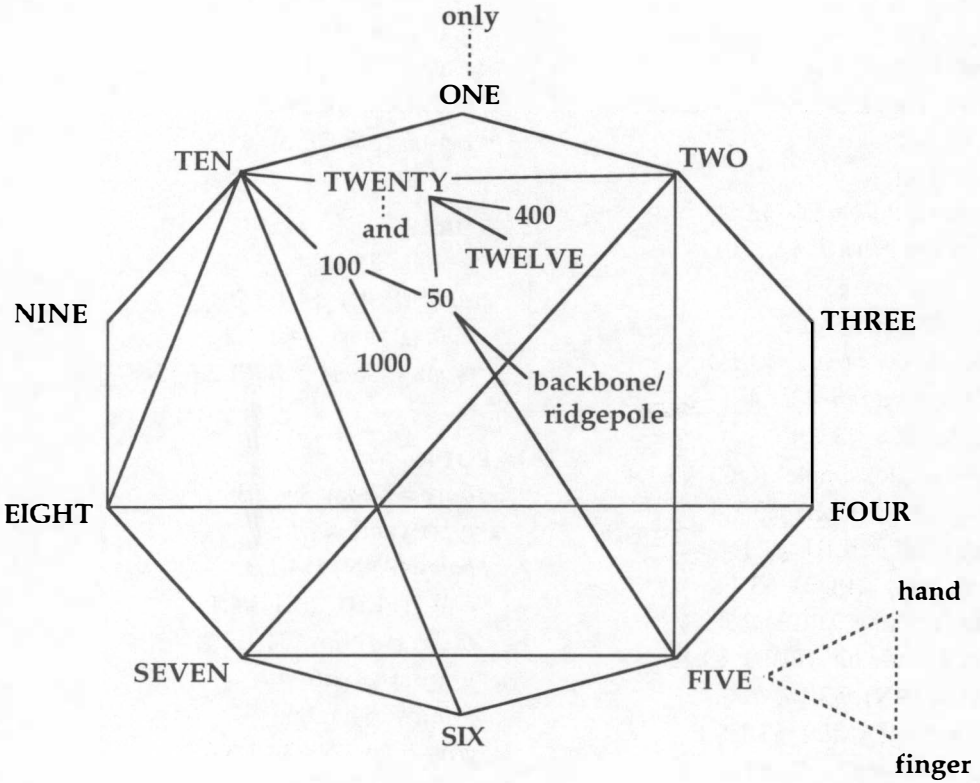
language Muya, a possibly cognate form $l\phi^{55}$ occurs as a citation classifier, but after *all* the numerals from 1 to 10 (for example so^{55}-$l\phi^{55}$ '3', $z\underline{u}^{35}$-$l\phi^{55}$ '4', ηa^{55}-$l\phi^{55}$ '5', and so on). Another Qiangic language, Shixing, has a different etymon for its ordinary general classifier, ko^{33}, but with a special allomorph ko^{35}, 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 monograph should be viewed as part of the *Sino-Tibetan Etymological Dictionary and Thesaurus* project (STEDT), a long-term 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.[1]

---

[1] As this monograph 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).

METASTATIC FLOWCHART OF NUMERICAL SEMANTIC ASSOCIATIONS

only

ONE

TEN   TWENTY   TWO
               and   400
          100   TWELVE
                    50
          1000
                    backbone/
                    ridgepole

NINE   THREE

EIGHT   FOUR

                                    hand

SEVEN   FIVE

        SIX

                            finger

(chart by Jonathan P. Evans)

lak 'hand' §4.1.4
k(r)ut 'hand' §4.2.2

SIX
d-ruk §4.2.1
d-k-ruk §4.2.1.3
d-krok §4.2.1
k-d-ruk §4.2.1.3
ruk or *rok* §4.2.1.4
k-[r]uk §4.2.1.4
d-k-rok or *k-d-rok* §5.1.3
C-krok (PLB) §4.2.1
krə (PTani) §4.2.1.2
krəŋ (AMD) §4.2.1.2
s-ruk (PNN) §4.2.1.5
li̯ôk (OC) §1.2.6
*k-d-ruk < *d-ruk* (Proto Rgyalrong) §5.2
a-krə (PTani) §5.1.3.1

SEVEN
s-nis §4.2.2
g-s-ni-s §4.2.2.6
k-nit §4.2.2.5; §5.1.3.1
*b-dun > *b-[d]yun* §4.2.2.8
s(n)i-t (PLB) §4.2.2
si² (PLB) §4.2.2.4
hnəs (PKaren) §4.2.2.3
kV-nut (PTani) §4.2.2.5
ts'i̯ĕt (OC) §1.2.6
k-s-nis (Proto Rgyalrong) §5.2

EIGHT
b-r-gyat ~ b-g-ryat §4.2.3
s-rit §1.2.1
s-g-ryat §4.2.3.2
pri-ñi (PTani) §4.2.3.8; §5.1.3.1
gyat §4.233; §4.2.3.5
pwăt (OC) §1.2.6; §4.2.3
*d-ryat < *g-ryat* (PKN) §4.2.3
b-ryat (Gurung-Tamang-Thakali) §4.2.3.4
ʔrit ⚹ *ʔryat* (PLB) §3.5.4.4; §4.1.1.4;
   §4.2.3; §4.2.3.4

C/V-gyat (PNN) §4.2.3.5
g-ryan (AMD, HIM) §4.2.3.6
lyoŋ (?) (AMD) §4.2.3.7

NINE
*d-kəw (= *d-kuw)* ⚹ *s-gəw* ⚹ *d-gaw* §4.2.4
d-gəw ⚹ *s-kəw* §5.1
s-d-[k]əw §4.2.4.1
C/V-gə:w (PNN) §4.2.4.1
gəw² (PLB) §4.2.4.4
k-n(y/w)a-ŋ (AMD et al.) §4.2.4.5
kV-(n)aŋ (PTani) §4.2.4.5
ki̯ug (OC) §1.2.6

TEN
gip ⚹ *gyap* §3.2.1
ts(y)iy ⚹ *tsyay* §3.2.2
*sy[sye < *s-tsyiy* (Proto Rgyalrong) §3.2.2
tši¹ (Proto Loloish) §3.2.2
tsyal (Nungish) §3.2.2
*som (< *tsom)* (Proto Kuki-Chin) §3.2.3.1
čam (Proto Tani) §3.2.3.1
rjɯŋ (PTani) §3.2.3.1
pal or *bal* §3.2.3.2
bo:n (PNN) §3.2.3.2
s-r/liŋ ⚹ *s-r/lyaŋ* (AMD et al.) §3.2.3.3
riŋ ⚹ *yiŋ* §3.2.3.3
liŋ ⚹ *lyaŋ* 'ten/hundred' §3.2.3.3
s-ryak ⚹ *s-rwak* §3.2.3.4
ro:k (PNN) §3.2.3.4
d(y)am ⚹ *t(y)am* 'ten; a full decade'
   §3.2.3.5
p/boŋ (Proto Kiranti) §3.2.3.6
ban ⚹ *bal* §3.2.3.2; §4.2.0.3
d̂i̯əp (OC) §1.2.6

TWENTY
m-kul §3.5.1.1
kun 'all' §3.5.1.1
kal 'load; bushel measure; group of twenty'
   (HIM) §3.5.1.1
ja (PNN) §3.5.1.1; §3.5.2.4

124

HUNDRED
*b-r-gya ⚹ *b-g-rya §3.5.4.5
*b-rya §3.5.4.5
*m-rya (Naga) §3.5.4.5
*hra¹ (PLB) §3.5.4.4
*m-lŋya (AMD) §3.5.4.6
*m-li(ŋ) (? < *m-l-ŋya) (AMD) §3.5.4.6

*păk (OC) §1.2.6

THOUSAND
*s-toŋ §3.5.4.7
*s-riŋ ⚹ *s-raŋ §3.5.4.7
*gheslo- (PIE) §3.5.4.7.2

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 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 (pers.comm.) 1991.

Bantawa: Gvozdanović 1985.

Chepang: Hale 1973.

Chourase: Gvozdanović 1985.

Dumi: Gvozdanović 1985.

Dzongkha: Mazaudon 1985. Rinzin (pers.comm.) 1984.

Gurung: Hale 1973.

126

Hayu: Michailovsky 1981:167. LSI III.1
  (Vayu):384-385.

Jirel: Hale 1973.

Kaike: Hale 1973.

Kanawari: Joshi 1909:2-3.

Khaling: Hale 1973.

Kham (Nepal): Hale 1973.

Kulung: Gvozdanović 1985.

Lepcha: Mainwaring/Grünwedel 1898.

Magari: Hale 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 1973.

Sharchop: Rinzin (pers.comm.) 1984.

Sherpa: Hale 1973.

Sikkim Bhutia (Dzongkha): Sandberg
  1895:59.

Sunwar: Hale 1973. Gvozdanović
  1985:143.

Tamang: Hale 1973.

Thakali: Hale 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): Osborne 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.

Anong: ZMYYC.

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 (Jinghua): 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.

Rgyalrong (Jiarong): Nagano 1979. ZMYYC. TBL.

Shixing: ZMYYC. TBL.

Zhábà: ZMYYC. TBL.

Other

Sulong (Sulung): ZMYYC.

Tujia: ZMYYC. TBL.

Nepali (Indo-European): Turner 1931, 1966. Hale 1973.

REFERENCES

Allen, Nicholas J., 1975, *Sketch of Thulung grammar, with three texts and a glossary.* Cornell University East Asia Papers, 6. Ithaca, NY: Cornell.

Barnard, J.T.O., 1934, *A handbook of the Rawang dialect of the Nung language.* Rangoon: Government Printing House.

Baxter, William, 1985, Tibeto-Burman cognates of Old Chinese *-ij* and *-ɨj*. In Thurgood, Matisoff and Bradley, eds., 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. 1710pp. (O.T.) and 533pp. (N.T.), paginated separately.

1971, *Thuthung Thak le Sam.* [New Testament and Psalms in Vaiphei.] Bangalore. 524pp. (N.T) and 137pp. (Psalms), paginated separately.

1972a, *Thuthunthak leh Sam.* [New Testament and Psalms in Gangte.] Bangalore. 513pp. (N.T.) and 189pp. (Psalms), paginated separately.

1972b, *Godan Rodai.* [New Testament in Boro]. Bangalore. 657pp.

1973, *Pathen Lekhabu Theng.* [Holy Bible in Kuki.] Bangalore. 1296pp. (O.T.) and 322pp. (N.T.), paginated separately.

1974, *Laisiangthou.* [Holy Bible in Paite.] Bangalore. 973pp. (O.T.) and 307pp. (N.T.), paginated separately.

1976, *Tongtep Ther leh Lha.* [New Testament and Psalms in Kom Rem.] Bangalore. 568pp.

1979, *Lai Siangtho.* [Holy Bible in Tiddim Chin.] Bangalore. 1072pp. (O.T.) and 350pp. (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: Curzon Press.

Bradley, David, E.J.A. Henderson and Martine Mazaudon, eds, 1989, *Prosodic analysis and Asian linguistics to honour R.K. Sprigg. PL*, C-104.

129

Burling, Robbins, 1961, *A Garo grammar*. Deccan College Postgraduate and Research Institute Monograph Series, 25. Poona: Linguistic Society of India.

1968, *Proto-Lolo-Burmese*. IJAL Special Publication 33.2, Part II. Issued simultaneously as Indiana Publications in Anthropology and Linguistics 43. The Hague: Mouton.

Caughley, Ross C., 1989, Chepang: a Sino-Tibetan language with a duodecimal numeral base? In Bradley et al., eds., 197-199.

1972, A vocabulary of the Chepang language. 40pp. mimeo. Kirtipur, Nepal: Summer Institute of Linguistics.

Chakravarty, L.N. et al., 1963, *A dictionary of the Taraon language*. Philology Section, Research Department, Northeast Frontier Agency, Shillong.

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. 1420pp. Beijing. ('ZMYYC').

Dai Qingxia, 1982, *[Outline grammar of the Achang language.]* In Chinese. Beijing: People's Publishing Co.

Dai Qingxia, et al., 1981, *[Chinese-Jingpho dictionary.]* In Chinese. Kunming: Yunnan People's Publishing Co.

1983, *[Jingpho-Chinese dictionary.]* In Chinese. Kunming: Yunnan People's Publishing Co.

Dai Qingxia and Huang Bufan, 1992, *Zang-Mian Yuzu Yuyan Cíhui [A Tibeto-Burman lexicon]*. In Chinese. Beijing: Central Institute of Nationalities.

Das Gupta, K., 1963, *An introduction to the Gallong language*. Shillong: Northeast Frontier Agency.

1968, *An introduction to Central Monpa*. Northeast Frontier Agency. Shillong.

1969, *Dafla language guide*. Shillong: Northeast Frontier Agency. 114pp.

1975, A few features of the Tagin language. *Resarun* 1/4 (6pp).

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, François, 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 Diffloth and Zide, eds, 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.

Emeneau, Murray B., 1957, Numerals in comparative linguistics (with special reference to Dravidian). BIHP 29:1-10.

Evans, Jonathan P., 1993, Unpublished fieldnotes on the Sanlong dialect of Qiang.

Fraser, J.O., 1922, *Handbook of the Lisu (Yawyin) language*. Rangoon: Government Printing House.

131

French, Walter T., 1983, *Northern Naga: a Tibeto-Burman mesolanguage.* PhD dissertation, City University of New York. 737pp.

Gao Huanian, 1955, [A preliminary investigation of the Hani language of Yang-wu.] In Chinese. Journal of Zhongshan University, 175-231.

1958, *[A Study of the grammar of the Yi (Nasu) language.]* In Chinese. Scientific Publishing Co. Beijing.

Goldstein, Melvyn C. and Nawang Nornang, 1970, *Modern spoken Tibetan: Lhasa dialect.* University of Washington Press. Seattle and London.

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).

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.

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, Ithaca, NY: 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) Hpun. In John McCoy and Timothy Light, eds, 101-134.

1997, *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.

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*. Edited by H.A. Rose. Journal of the Asiatic Society of Bengal, n.s., vol.5, extra number.

Judson, Adoniram, 1953, *Judson's Burmese-English Dictionary*. Centenary edition, revised and enlarged by R.C. Stevenson and F.H. Eveleth. Baptist Board of Publications, Rangoon. 2nd Printing 1966 (1123pp.).

Jui I-fu, 1948, [Notes on the sounds of the Lisu language, with remarks on the Lisu script.] In Chinese. BIHP 17:303-327.

Karapurkar, Pushpa Pai, 1976, *Kokborok (Tripuri) grammar*. Central Institute of Indian Languages Grammar Series, 3. Mysore.

Karlgren, Bernhard, 1957. *Grammata Serica Recensa*. BMFEA 29/1:1-332. ('GSR').

Kitamura, Hajime, T. Nishida and Y. Nagano, eds, 1994, *Current issues in Sino-Tibetan linguistics*. Organizing Committee, 26th International Conference on Sino-Tibetan Languages and Linguistics. Osaka.

Knowlton, Edgar C., Jr., 1976, Serial enumeration in Malayan Semai. In Diffloth and Zide, eds, *Austroasiatic number systems*, p.99. (Special issue of Linguistics: an International Review, 174).

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, 1942, 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: Eastern Bengal and Assam Secretariat Press.

Lu Shaozun, 1983, *[An outline grammar of the Pumi language.]* In Chinese. Beijing: People's Publishing Co.

   1986, *[An outline frammar 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.

McCoy, John and Timothy Light, eds, 1986, *Contributions to Sino-Tibetan studies*. Leiden: E.J. Brill.

Mainwaring, G.B. and A. Grünwedel, 1898, *Dictionary of the Lepcha language*. Berlin: Unger Brothers.

Maran, LaRaw, [in prep.], The Jingpho dictionary: a dictionary of modern Jingpho. [Revised and enlarged version of Hanson 1906, with tones indicated.]

Marcos-Marín, Francisco, 1993, Numerals and typology in minority languages of China. Proceedings of the XVth International Congress of Linguists, vol.4:169-172.

Marrison, Geoffrey E., 1967, *The classification of the Naga languages of North East India*. 2 vols. 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. University of California Press, Berkeley and Los Angeles. Reprinted 1982.

1973b, Tonogenesis in Southeast Asia. In Larry M. Hyman, ed., *Consonant types and tone*, pp.71-95. Los Angeles: Southern California Occasional papers in Linguistics, 1.

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*, 265-287. Stanford, CA: Stanford University.

1978a, *Variational semantics in Tibeto-Burman: the 'organic' approach to linguistic comparison*. F.K. Lehman, ed. Occasional papers of the Wolfenden Society on Tibeto-Burman linguistics, vol.6. 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 Thurgood, Matisoff, and Bradley, eds, 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 J. McCoy and T. Light, eds, 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*, 3-14. Kyoto: Shokado.

1989, Palatal suffixes in Sino-Tibetan. Paper originally presented at 22nd International Conference on Sino-Tibetan Languages and Linguistics, University of Hawaii.

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 Yasuhiko Nagano et al., eds *Current Issues in Sino-Tibetan linguistics*, 36-58. Osaka: National Museum of Ethnology.

1994b, Sangkong of Yunnan: secondary 'verb pronominalization' in Southern Loloish. In Yasuhiko Nagano et al., eds *Current issues in Sino-Tibetan linguistics*, 588-607. Osaka: National Museum of Ethnology.

1994c, Watch out for number ONE: Jingpho ŋāi 'I' and ləŋâi 'one' (with some speculations about Jingpho number TWO). LTBA 17/1:155-165.

1995, Sino-Tibetan palatal suffixes revisited. In Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano, eds, *New horizons in Tibeto-Burman morphosyntax*, 35-91. Osaka: National Museum of Ethnology.

1996, *Languages and dialects of Tibeto-Burman*. STEDT Monograph Series 2. Berkeley, CA: Center for Southeast Asia Studies.

In preparation, *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*. 124-157. Nakon Pathom: Mahidol University.

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, 1988, *La langue Hayu. Paris: Editions du Centre National de la Recherche Scientifique*.

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 [Garo] dictionary*. Calcutta.

Nagaland Bhasha Parishad [Linguistic Circle of Nagaland], 1972a, *Hindi-Chakhesang-English dictionary (Chokri dialect)*. Kohima. 138pp.

1972b, *Hindi-Pochury-English dictionary*. Kohima. 100pp.

1973, *Hindi-Zeliang-English dictionary (Zemi dialect)*. Kohima. 95pp.

1974, *Hindi-Kheja-English dictionary*. Kohima. 199pp.

Nagano, Yasuhiko, 1979, A historical study of rGyarong rhymes. LTBA 5/1:37-47.

1984, *A historical study of the rGyarong verb system*. Tokyo: Shiseido.

Nishi, Yoshio, 1982, Swadesh 100-word lists for some languages of the Monpa group. MS.

Nishi, Yoshio, James A. Matisoff and Yasuhiko Nagano, eds, *New horizons in Tibeto-Burman morphoxyntax*. Senri Ethnological Studies, 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*. PhD dissertation, Cornell University.

Phillips, E.G., 1904, *Outline grammar of the Garo language*. Shillong.

Rai, Agami 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*. PhD dissertation, Yale University.

Sandberg, Graham, 1895, *Manual of the Sikkim Bhutia language or Denjongke*. Second and enlarged edition. London: India Office Library.

Schmidt, Ruth L., 1993, *A practical dictionary of modern Nepali*. Delhi: Ratna Sagar P. Ltd.

Simon, I.M., 1972, *An introduction to Apatani*. Philology Section, Department of Research, Arunachal Pradesh. Gangtok, Sikkim.

1976, *Hill Miri language guide*. Published by Director of Information and Public Relations, Arunachal Pradesh.

1979, *Miji language guide*. Philology Section, Department of Research, Arunachal Pradesh. Gangtok, Sikkim.

Solnit, David B., 1979, Proto-Tibeto-Burman *r in Tiddim Chin and Lushai. Linguistics of the Tibeto-Burman Area 4/2:111-121.

1984, Kayah numerals. MS.

In preparation, *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, No.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.

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.

1993, *A historical-comparative study of the Tani (Mirish) branch of Tibeto-Burman*. PhD dissertation, University of California, Berkeley.

Sun, Jackson Tianshin, 1992, Review of ZMYYC (Chinese Academy of Social Sciences, 1991). LTBA 15/2:73-113.

Taid, Tabu, 1993, Mising wordlist. MS.

Talukdar, G.C. et al., 1962, *A phrase book in Idu*. Philological Section, Research Department, Northeast Frontier Agency, Shillong.

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. PL*, C-87.

Toba, Sueyoshi, 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: a comparative account*. Amsterdam/ Philadelphia: John Benjamins.

Wen Yu, 1950, [An abridged Ch'iang vocabulary (Chiu Tzu Ying dialect).] In Chinese. Studia Serica 9:17-54.

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*. PhD 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 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. Beijing: Chinese Academy of Sciences, Linguistics Research Institute Monograph Series, 5.

Zhao Yansun, 1982, [The problem of the genetic affiliation of the Bai language.] In Chinese. In *Minzu Yuwen Yanjiu Wenji*, 150-188. Xining: Qinghai People's Publishing Co.