Sino-Tibetan numeral systems: prefixes, protoforms and problems
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

Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra
# TABLE OF CONTENTS

SYMBOLS AND ABBREVIATIONS viii
PREFACE ix

CHAPTER 1: INTRODUCTION 1
  1.1 Issues in the reconstruction and systematic behaviour of the Tibeto-Burman numerals 1
    1.1.1 Proto-variation 1
    1.1.2 Lexico-semantic issues 3
  1.2 Overview of Sino-Tibetan numeral systems according to subgroup 5
    1.2.1 Lolo-Burmese and Karenic 5
    1.2.2 Kachin-Nung and Qiangic 7
    1.2.3 Himalayish and Rgyalrong 8
    1.2.4 Kuki-Chin-Naga and Bodo-Garo 9
    1.2.5 Abor-Miri-Dafla 10
    1.2.6 Chinese 10

CHAPTER 2: LANGUAGE CONTACT AND THE WEIGHT OF NUMBERS 12
  2.0 Introduction 12
  2.1 Nepali and the TB languages of Nepal 13

CHAPTER 3: ONE AND TEN AND TEENS AND TWENTIES 17
  3.1 Profile of number ONE 17
    3.1.1 *it ≈ yat 17
    3.1.2 *k-(y)at ≈ *k-(y)it ≈ *k-yan ≈ *k-(y)in 18
    3.1.3 *ka and *ko 19
    3.1.4 *g-(t(y)i-k ≈ *tya-k and *d/tay ≈ *d/tan 19
    3.1.5 New roots for ONE in Abor-Miri-Dafla and elsewhere in Tibeto-Burman 21
    3.1.6 ONE as indefinite article or general numeral affix 24
  3.2 Profile of number TEN 24
    3.2.1 *gip ≈ *gyap 25
    3.2.2 *ts(y)i ≈ *tsyay 25
    3.2.3 New roots for TEN in Kuki-Chin-Naga, Abor-Miri-Dafla, and elsewhere 26
3.3 Teen formation: from 10 to 20

3.3.1 Teen formation in Kuki-Chin-Naga

3.3.2 Teen formation elsewhere in Tibeto-Burman
  [A] Himalayish
  [B] Lolo-Burmese
  [C] Bodo-Garo
  [D] Abor-Miri-Dafla

3.4 Interchange and confusion between ONE and TEN

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

3.4.2 mu- in Mishmi

3.4.3 tsek/s(α)- in Rgyalrong

3.4.4 *s(h)e ‘1’ and *ts(y)i ‘10’

3.5 Round Number Formation: decimal and vigesimal systems

3.5.1 Vigesimality, in Tibeto-Burman and elsewhere

3.5.2 Round number formation in Kuki-Chin-Naga

3.5.3 Round number formation in Himalayish
  [A] Cuona Menba
  [B] Tamang and Sharchop/Tsangla
  [C] Kaire and Jirel
  [D] Kanawari
  [E] Magari and Khaling

3.5.4 HUNDRED and THOUSAND

CHAPTER 4: THE PRIMARY NUMERALS: TWO to NINE

4.0 Introduction

4.1 The lower numerals: 2–5
  4.1.1 Profile of number TWO
  4.1.2 Profile of number THREE
  4.1.3 Profile of number FOUR
  4.1.4 Profile of number FIVE
  4.1.5 Traces of quinary numeral systems in Himalayish

4.2 The higher numerals: 6–9
  4.2.0 Additive, subtractive, and multiplicative formations
  4.2.1 Profile of number SIX
  4.2.2 Profile of number SEVEN
  4.2.3 Profile of number EIGHT
  4.2.4 Profile of number NINE

CHAPTER 5: PREFIXAL BEHAVIOUR WITH NUMERALS

5.1 Prefixal variability and replaceability
  5.1.1 Voicing and vocalisation of prefixes
  5.1.2 Contamination by prefixes of non-contiguous numerals
  5.1.3 Prefix preemption of the root-initial of a numeral
5.2 ‘Prefix runs’ and reprefixation
   5.2.1 Mnemonic and rhythmic considerations
5.3 Numeral prefixes in Himalayish
   5.3.1 Bodish languages with distinctive dental-initialled SEVEN
   5.3.2 Khaling
   5.3.3 Gurung-Tamang-Thakali and Newari
   5.3.4 Kanawari (=Kanuri) and Lepcha (=Rong)
   5.3.5 Monpa dialects
5.4 Numeral prefixes in Kuki-Chin-Naga
   5.4.1 Retention of the proto-system (as conceived in STC)
   5.4.2 Degeneration of the proto-system: prefix loss
   5.4.3 Secondary general numeral prefixes
   5.4.4 Innovative runs in the higher numerals
5.5 Numeral prefixes in Abor-Miri-Dafla
   5.5.1 Runs in the lower numerals
   5.5.2 Systems with no secondary prefixation in the lower numerals

CHAPTER 6: SUMMARY AND AFTERWORD

APPENDIX 1: METASTATIC FLOWCHART OF NUMERICAL SEMANTIC ASSOCIATIONS
APPENDIX 2: INDEX OF RECONSTRUCTED ETYMA
APPENDIX 3: INDEX OF LANGUAGES AND SOURCES
REFERENCES
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A ⋆ B</td>
<td>A and B are co-allofams; A and B belong to the same word-family</td>
</tr>
<tr>
<td>AMD</td>
<td>Abor-Miri-Dafla</td>
</tr>
<tr>
<td>BIHP</td>
<td>Bulletin of the Institute of History and Philology (Beijing; Taipei)</td>
</tr>
<tr>
<td>BMFEA</td>
<td>Bulletin of the Museum of Far Eastern Antiquities (Stockholm)</td>
</tr>
<tr>
<td>BSI</td>
<td>Bible Society of India</td>
</tr>
<tr>
<td>CSDPN</td>
<td><em>Clause, sentence, and discourse patterns in selected languages of Nepal</em> (Hale, ed. 1973)</td>
</tr>
<tr>
<td>GEM</td>
<td>Geoffrey E. Marrison (1967)</td>
</tr>
<tr>
<td>GSR</td>
<td><em>Grammata Serica Recensa</em> (Karlgren 1957)</td>
</tr>
<tr>
<td>GSTC</td>
<td><em>God and the Sino-Tibetan copula</em> (Matisoff 1985b)</td>
</tr>
<tr>
<td>Him.</td>
<td>Himalayish</td>
</tr>
<tr>
<td>HJAS</td>
<td>Harvard Journal of Asiatic Studies</td>
</tr>
<tr>
<td>JASB</td>
<td>Journal of the Asiatic Society of Bengal</td>
</tr>
<tr>
<td>Jg.</td>
<td>Jingpho</td>
</tr>
<tr>
<td>JRASB</td>
<td>Journal of the Royal Asiatic Society of Bengal</td>
</tr>
<tr>
<td>KCN</td>
<td>Kuki-Chin-Naga</td>
</tr>
<tr>
<td>LSI</td>
<td><em>Linguistic survey of India</em> (Grierson and Konow 1903-28)</td>
</tr>
<tr>
<td>LTBA</td>
<td>Linguistics of the Tibeto-Burman area (Berkeley)</td>
</tr>
<tr>
<td>MC</td>
<td>Middle Chinese (= Karlgren’s ‘Ancient Chinese’)</td>
</tr>
<tr>
<td>NBP</td>
<td>Nagaland Bhasa Parishad (Linguistic Circle of Nagaland, Kohima)</td>
</tr>
<tr>
<td>OC</td>
<td>Old Chinese (= Karlgren’s ‘Archaic Chinese’)</td>
</tr>
<tr>
<td>PGmc</td>
<td>Proto Germanic</td>
</tr>
<tr>
<td>PIE</td>
<td>Proto Indo-European</td>
</tr>
<tr>
<td>PLB</td>
<td>Proto Lolo-Burmese (=Proto Burmese-Yipho)</td>
</tr>
<tr>
<td>PNN</td>
<td>Proto Northern-Naga</td>
</tr>
<tr>
<td>PST</td>
<td>Proto Sino-Tibetan</td>
</tr>
<tr>
<td>PTB</td>
<td>Proto Tibeto-Burman</td>
</tr>
<tr>
<td>STC</td>
<td><em>Sino-Tibetan: a conspectus</em> (Benedict 1972)</td>
</tr>
<tr>
<td>STEDT</td>
<td>Sino-Tibetan etymological dictionary and thesaurus project (Berkeley)</td>
</tr>
<tr>
<td>TB</td>
<td>Tibeto-Burman</td>
</tr>
<tr>
<td>TBL</td>
<td><em>A Tibeto-Burman lexicon</em> (Dai/Huang 1992)</td>
</tr>
<tr>
<td>TSR</td>
<td><em>The Loloish tonal split revisited</em> (Matisoff 1972a)</td>
</tr>
<tr>
<td>VSTB</td>
<td><em>Variational semantics in Tibeto-Burman</em> (Matisoff 1978a)</td>
</tr>
<tr>
<td>WB</td>
<td>Written Burmese</td>
</tr>
<tr>
<td>WT</td>
<td>Written Tibetan</td>
</tr>
<tr>
<td>ZMYYC</td>
<td><em>Zang-Mianyu Yuyin he Cihui</em> (Chinese Academy of Social Sciences 1991)</td>
</tr>
</tbody>
</table>
This study was originally prepared for the Seventeenth International Conference on Sino-Tibetan Languages and Linguistics at the University of Oregon (September 1984), 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 permission 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, 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’). In any event, there are vast differences,
both qualitative and quantitative, in the data available for particular TB languages or language-groups.

![Diagram of the branches of Tibeto-Burman](image)

**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). 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* (<Research Arunachal>). 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 as in Figure 2, raises a variety of interesting morphophonemic and lexico-semantic issues.

![Figure 2: Proto Tibeto-Burman Numerals](image)

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 *-ay and *-aw. 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).
around a basic phonological shape and a core of meaning. Variability is observable in all parts
of the TB syllable: rhymes, initials, prefixes, tones.\footnote{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), \textit{Variational semantics in Tibeto-Burman} ("VSTB"). The symbol
\textquoteleft \text{\s} \textquoteleft\text{ is there introduced to stand for the allofamic relationship: X \text{\s} Y \text{"X and Y are
co-allofams; X and Y both belong to the same word-family".}}\footnote{See the discussion in Matisoff (1980:15-17), and \S 4.2.4.}
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 (\textit{NINE}, \textit{TEN}) or in the prefix (\textit{FIVE}, \textit{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 \textbf{VARIATION OF PROTO-RHYME}

The STC recognises a PTB alternation *-uw \textasciitilde*-*aw in \textit{NINE}, on slender evidence, rejecting
as secondary the better-attested variant in -wa (Lushai pakua, Angami thepf\textasciitilde).\footnote{Matisoff (1985b:5, 32), and \S 3.2.2.}
For \textit{TEN}, the first version of STC recognised ‘vowel gradation’ between *-ai and *-i to account for Written
Burmese (WB) \textquoteleft \text{achai 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.}}\footnote{See \S 3.1.1, \S 3.1.2, \S 3.1.4, \S 3.2.1, \S 3.2.2, \S 3.2.3.3, etc.}

\textit{Two} furnishes an example of variation of final consonant. Alongside the principal
allofam in -s, a variant in *-k (underlying, for example, WB \textit{hnac} \textltens{*s-nik}) is also attested
independently in several branches of TB, but is denied PTB status in STC. (See \S 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.\footnote{See \S 3.1.1, \S 3.1.2, \S 3.1.4, \S 3.2.1, \S 3.2.2, \S 3.2.3.3, etc.}

1.1.1.2 \textbf{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 \textit{FIVE}, STC does admit proto-variation,
posing alternation between the *b- and *f- prefixes at the PTB level.

\textit{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, \textit{FIVE} and \textit{EIGHT},
conceiving of this phenomenon in a paradigmatic sense for \textit{FIVE} but in a syntagmatic sense for
\textit{EIGHT}.\footnote{See the discussion in Matisoff (1980:15-17), and \S 4.2.4.}
However, proto-variation in prefix may with equal justice be imputed to at least two other numerals, SIX and NINE. Besides the *d- prefix for NINE, at least four branches of TB point to a sibilant prefix *s- which STC does not recognize (Garo sku, Kanauri zgui, Jingpho džakhū, 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>k/. Yet tonal developments in Loloish, where the word appears in the low-stopped tone despite its voiceless initial (for example, Lahu khɔ?), clearly point to a doubly-prefixed prototype */d-k-rok/. 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, pa-(written pa-):

<table>
<thead>
<tr>
<th>Lushai</th>
<th>1 pakhat</th>
<th>2 pahnih</th>
<th>3 pathum</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 pali</td>
<td>5 panga</td>
<td>6 paruk</td>
<td></td>
</tr>
<tr>
<td>7 pasarih</td>
<td>8 pariat</td>
<td>9 pakuas</td>
<td></td>
</tr>
</tbody>
</table>

Yet STC has to recognize two shorter prefix runs already at the PTB level: the *g- in TWO *g-nis <= THREE *g-sum, and the *b- in FOUR *b-lay <= FIVE *b-ga. 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’.\(^\text{10}\)

Although wholesale borrowing of numerals is by no means unheard of,\(^\text{11}\) 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.\(^\text{12}\)

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.\(^\text{13}\) 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.\(^\text{14}\) 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.\(^\text{15}\)

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.)\(^\text{16}\) 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 (nyŋ), using the Chinese loan òt only in compound numerals (II, 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 (loŋbû) and TWO (lakhông), that until recently have not been relatable to anything else. In Matisoff 1994c I identify the former with the Jg. first person pronoun ngâi ‘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 hnae ‘2’, hrac ‘8’ (< PLB *s-ni-t and *s-riit, respectively).
Another route by which a prefix could survive was by ‘pre-empting’ or driving out a weak (non-obstruental) root-initial\(^\text{19}\), as in SEVEN \(*\text{s-nit} > \text{Lahu} \, \text{šf}^\text{17}\), where the root-initial \(*n-\) has fallen victim to the sibilant prefix. Another famous example is the Maru (Burmish) word for FOUR, \(\text{bit} (\thicksim \,*\text{b-liy})\), an isolated instance of the survival of the \(*b-\) prefix in Lolo-Burmese.\(^\text{20}\)

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.\(^\text{21}\) 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 \(\text{swa}^\text{22}\) as ‘double’ (It also occurs in compounds with the meaning ‘companion’, as in \(\text{khō-bē-swā ‘friend’, bē-swā-rå ‘be companions with’}\).\(^\text{22}\) The Kayah numeral \(\text{ta- ‘one’}^\text{18}\) 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, \(\text{pb sō swā? ‘six round objects’}\), while with all the other numerals, including SEVEN and NINE, the classifier must follow (sō \(\text{swā ta-pb}^\text{22} ‘seven round objects’).

---

\(^{17}\) \(\text{šf}^2\) is the variant that occurs in counting, while the ‘etymologically correct’ alloform \(\text{šf}\) 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 \(\text{five} \, \text{fi-f.}, \) or \(\text{ten} \, \text{-teen} \, \text{-ty}^\text{§3.3}\).

\(^{18}\) One characteristic type of morphophonemic change in numerals is \(\text{destressing}^\text{§3.3.2(B)}\) 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 \(\text{ONE}^\text{Figure 4}\).

\(^{19}\) For the first use of the term \(\text{prefix preemption}^\text{sec Matisoff (1972b)}\).

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

\(^{21}\) 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. \text{shuāng}) ‘pair’, which also underlies the Thai numeral \(\text{sōŋ ‘two’}\).
Sgaw (Jones 1961)  |  E. Kayah (Solnit 1984)
---|---
ONE  |  tā  |  ta-
TWO  |  khă23  |  nā
THREE |  θā  |  sō
FOUR  |  lwī  |  lwī
FIVE  |  jē  |  ŋē ~ ŋē
SIX  |  xē  |  sō swā?
SEVEN  |  mwī  |  sō swā? tə-
EIGHT  |  xēʔ  |  lwī swā?
NINE  |  khwī  |  lwī swā? tə-
TEN  |  ḥi  |  chā ~ chă24

**FIGURE 4: SOME KAREN NUMERALS**

Other Karen dialects, especially Pa-O (Taungthu) have developed secondary dental suffixes with certain numerals: Pa-O lit ‘4’, ngāt ‘5’, kūt ‘9’. In the case of nūt ‘7’ and sōt ‘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. 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, p-. The old causative prefix *s- has been preserved and generalised as ša ~ dža-. Younger strata of prefixation are much in evidence, with a tendency to create fully syllabic prefixes out of sub-syllabic ones, for example, *m-raIJ ‘horse’ > Jg. gum-rāg. With respect to numeral prefixation, Jingpho is relatively conservative, preserving the protoprefixes 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) vu35 < *b-lay ‘four’.

---

23 Note the preemption of the root-initial by the velar prefix, *g-nis > khē.
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, chá sō swā ‘16’, i.e. 10 + (3 x 2), where TEN is in an additive relationship to the following numeral, with chá sō swā ‘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 | tso55 | zau
TWO | m55 | wne
THREE | so35 | wsu
FOUR | v1i35 | wze
FIVE | ηa55 | wque
SIX | kru5 | khuu53 | wtchau
SEVEN | soi5 | sjie
EIGHT | mösat | sja53 | yie
NINE | Ḗkha5 | gtu35 | ūge
TEN | sī | tshe35 | za/sqha

Note the impressive run of the prefix w- (< *b-) in the Ergong numerals from 2 to 6, even longer than the Jingpho run of ma- 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 | goi5 | id | ki:h
TWO | gnis | ni5 | pi:h
THREE | gsum | /um | som
FOUR | bži | pā | plīh
FIVE | lña | ηa | lña:h
SIX | drug | ūk | ū:h
SEVEN | bdun | stif | niś
EIGHT | brgyad | rai | preht
NINE | dgu | zgui | ku
TEN | bçu | 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 ūa-, but has also innovated a ka- 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-Dafila (see Nagano 1984),\(^{28}\) and/or Qiangic (as maintained by Sun Hongkai 1985).\(^{29}\) See Figure 6b.

<table>
<thead>
<tr>
<th></th>
<th>Lepcha</th>
<th>Rgyalrong (Zida dialect)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>kat</td>
<td>tšek</td>
</tr>
<tr>
<td>TWO</td>
<td>nət</td>
<td>kenes</td>
</tr>
<tr>
<td>THREE</td>
<td>sam</td>
<td>kesom • kesam</td>
</tr>
<tr>
<td>FOUR</td>
<td>tali</td>
<td>kewdži</td>
</tr>
<tr>
<td>FIVE</td>
<td>təŋo</td>
<td>kemŋa</td>
</tr>
<tr>
<td>SIX</td>
<td>tārak</td>
<td>kētā</td>
</tr>
<tr>
<td>SEVEN</td>
<td>kakyək</td>
<td>kešnit • kešnis • kešnēs</td>
</tr>
<tr>
<td>EIGHT</td>
<td>kəku</td>
<td>warzē(t)</td>
</tr>
<tr>
<td>NINE</td>
<td>kakyot</td>
<td>kēŋgu</td>
</tr>
<tr>
<td>TEN</td>
<td>kəti</td>
<td>ʃtši</td>
</tr>
</tbody>
</table>

**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).
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.

<table>
<thead>
<tr>
<th></th>
<th>Aka</th>
<th>Miju</th>
<th>Milang</th>
<th>Serdukpen</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>a</td>
<td>kumo</td>
<td>akan; atel</td>
<td>han</td>
</tr>
<tr>
<td>TWO</td>
<td>kshi</td>
<td>kinin</td>
<td>ne</td>
<td>n(y)ik</td>
</tr>
<tr>
<td>THREE</td>
<td>zu</td>
<td>ksam</td>
<td>ham</td>
<td>ung</td>
</tr>
<tr>
<td>FOUR</td>
<td>fi-ri</td>
<td>kambran</td>
<td>pe</td>
<td>bissi</td>
</tr>
<tr>
<td>FIVE</td>
<td>phum</td>
<td>klin</td>
<td>pangu</td>
<td>khu³¹</td>
</tr>
<tr>
<td>SIX</td>
<td>rieh</td>
<td>katam</td>
<td>sap</td>
<td>khit</td>
</tr>
<tr>
<td>SEVEN</td>
<td>mulh</td>
<td>nin</td>
<td>rangal</td>
<td>sit</td>
</tr>
<tr>
<td>EIGHT</td>
<td>sikzi</td>
<td>grin</td>
<td>rayeng</td>
<td>sargiat</td>
</tr>
<tr>
<td>NINE</td>
<td>sthö</td>
<td>natmo</td>
<td>kanyem</td>
<td>diki</td>
</tr>
<tr>
<td>TEN</td>
<td>rhi</td>
<td>kyapmo</td>
<td>hangtak</td>
<td>dokche</td>
</tr>
</tbody>
</table>

**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 | *?iêt (GSR #394) | *it
 | *tsjâk (GSR #1260) | *g-t(y)ik
TWO | *njar (GSR #564) | *g-ni-s
THREE | *ts'am - *sam (GSR #647, 648) | *g-sum
FOUR | *siâd³⁴ (GSR #518) | *b-lôy
FIVE | *ngo³⁵ (GSR #58) | *l-/b-qa
SIX | *liôk³⁶ (GSR #1032) | *d-ruk
SEVEN | *ts'êt (GSR #400) | *s-nis
EIGHT | *pwać³⁷ (GSR #281) | *b-g-ryat/*b-r-gyat
NINE | *kiug > MC *kiôw (GSR #992) | *d-kôw
TEN | *dïap (GSR #686) | *g(y)ip
HUNDRED | *pâk³⁸ (GSR #781) | *r-gya (but WT brgya)

STC derives this from pre-Archaic *p-sây.

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

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").

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

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”).
CHAPTER 2
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’ ḥōk, ‘7’ kīt, ‘8’ pet, ‘9’ kāw, ‘10’ sōp.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.

![Figure 9: Nepali Numerals](image)

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. 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.

![Figure 10: Preserved TB Numerals in Some Languages of Nepal](image)

---

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 kat and Chepang yat 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 numéreaux 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. 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. 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.

<table>
<thead>
<tr>
<th>ONE</th>
<th>ko</th>
<th>SIX</th>
<th>ru</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>nā</td>
<td>SEVEN</td>
<td>yet</td>
</tr>
<tr>
<td>THREE</td>
<td>sium</td>
<td>EIGHT</td>
<td>let11</td>
</tr>
<tr>
<td>FOUR</td>
<td>bā</td>
<td>NINE</td>
<td>gu</td>
</tr>
<tr>
<td>FIVE</td>
<td>pō</td>
<td>TEN</td>
<td>kodium</td>
</tr>
</tbody>
</table>

**Figure 11: The moribund TB-derived numerals of Thulung Rai**

---

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)

<table>
<thead>
<tr>
<th>ONE</th>
<th>kolu</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>nakpu (human) / na?ung (non-human)</td>
</tr>
<tr>
<td>THREE</td>
<td>tshukpu (human) / tshu?ung (non-human)</td>
</tr>
<tr>
<td>FOUR</td>
<td>b(l)i?ung</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>ONE</th>
<th>kom-pu/kwong-pu</th>
<th>kwo-mi/kwong-mi</th>
<th>ko-lu</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>nak-pu</td>
<td>na:ng-mi</td>
<td>na:-yung</td>
</tr>
<tr>
<td>THREE</td>
<td>chhuk-pu</td>
<td>chhung-mi</td>
<td>chhu-yung</td>
</tr>
<tr>
<td>FOUR</td>
<td>blik-pu</td>
<td>blig-mi</td>
<td>bli-ning</td>
</tr>
</tbody>
</table>

**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 (tikpo ‘1’, sakpo ‘2’ ... smo ‘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 (ikpok ‘1’, hiapok ‘2’, sumkapok ‘3’, rețkapok ‘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: tokpu ‘1’, sakpu ‘2’, bhlokpu ‘3’, rõkpu ‘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.

(d) **-cī/-ji.** Several languages have this numeral suffix, including Mewahang hicci ‘2’, sumji ‘3’; Yakkha hitci ‘2’, sumci ‘3’; and Lohorong nicci ‘2’, sumci ‘3’, ricci ‘4’, naci ‘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’ ..., nokci ‘9’) and Kulung (nicci ‘2’ ..., nuci ‘9’, boci ‘10’).

(e) **-sī/-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’, liissi ‘4’, nggaasi ‘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) pik-tsing ‘2’.

Other suffixes, for example, Chepang -joś (Figure 10) and Hayu -gung, 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 yā-mi ‘daughter’, b-mi-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.
CHAPTER 3
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 alloforms 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 นิ้ว ‘1’, sip ‘10’, but sip-ें ‘11’, ร์ละ ‘101’. (This -ें, 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’, rिच-ें ‘101’. (This -ें, like the rest of the Thai numerals from 2 to 9, is of Sino-Tibetan origin (§3.1.1).)

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 *ʔět (p.162).

To these I would now like to add Chepang यात(-जो?), 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):

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mehawang</td>
<td>ek-ku³</td>
</tr>
<tr>
<td>Yakkhaba</td>
<td>ik-ko</td>
</tr>
<tr>
<td>Khulung</td>
<td>i-bum ~ i-bim</td>
</tr>
<tr>
<td>Bantawa</td>
<td>uk- (as in uk-tai, uk-tak, uk-pok, uk-ta, all meaning ‘ONE’ in various dialects: Gvozdanović 1985:188)</td>
</tr>
</tbody>
</table>

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 xat; Nruanghmei khüt.

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

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 khürû); Pochury khe; Meluri ke (also kesû); Sema khe (also lakî); Mishmi (Dubey 1983) khege; Chulikata eikhe; 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 atef); 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/l: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 rkyat-pa, WB khyat ‘single’.

A similar variational pattern in TWO is suggested by the Lepcha doublet ni ≠ nät. 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-Tenga 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’; Lhoppa a-ko (also aken); Gallong a-ko (also aken); Dafla aku (also 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 thik-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 akan (also atir, ayirr); Darang Deng k’un\(^5\); Apatani kun (non-humans), kon (humans); Idu khun\(^5\) (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)i-k ≠ *tya-k ‘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-ken.

\(^6\) 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 \textit{tk-po}, \textit{tak-pu}, Lohorong \textit{thik-ko} (for the second element see §3.1.3); Athpare \textit{thik}; Limbu \textit{lot-thik}; Dzongkha \textit{ci}; Kaike \textit{ti}; and the second element of Sikkim Bhutia \textit{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 \textit{katek} (ZMYYC #911)), but other prefixes are attested elsewhere (for example, Qiang \textit{pets}^\textit{bi} (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 \textit{gcig} and WB \textit{tac}) and at much more recent time-depths (for example, in Nungish, where Rawang has \textit{hti} (= \textit{thi}) (Barnard 1934) but a Nuijiang dialect has \textit{tw}^55).

(c) The Himalayish languages of the Tamang-Gurung-Thakali nucleus seem to point to medial \textit{*r-} rather than \textit{*y-}: Gurung \textit{grihq}, Tamang \textit{kiih} (with preemption), Thakali \textit{thi}.

(d) The vocalic nucleus also shows proto-variation between \textit{*i-} and \textit{*ya-}, a mysterious property of several other numerical roots as well (§3.1.1). The variant with \textit{*ya-} vocalism, \textit{*tya-k} (STC, note 271, p.94), is reflected by Chinese ึ*tsjäk 'one, single' (GSR #1260c), to which we may add a number of putative TB cognates: Bumthang \textit{thek}, tek (Nishi 1982); Monpa (Cuona) \textit{te}^254 (Sun et al., 1980); Bai tia (Dell 1981:61).\footnote{Several forms with affricate initials probably also descend from the allofam \textit{*ya-}, with no direct reflex of a final stop: Newari \textit{cha}; Konyak Naga \textit{ja}; Chang Naga \textit{chie}.} Several forms with affricate initials probably also descend from the allofam \textit{*ya-}, with no direct reflex of a final stop: Newari \textit{cha}; Konyak Naga \textit{ja}; Chang Naga \textit{chie}.

(e) Many daughter languages show no trace of an original final stop. Sometimes this is undoubtedly the regular fate of the \textit{*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 after the \textit{*k}, the resultant \textit{t(y)i-k} then looks a lot like one of the main TB roots for TEN \textit{*ts(y)iy} \textit{*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 \textit{*k} stops, others do not. The vocalism appears to vary among \textit{*i-}, \textit{*ay}, and \textit{*ey}, suggesting that this etymon was often unstressed and hence of unstable vowel quality.\footnote{In Modern Burmese the fully stressed form \textit{ti} (the regular reflex of WB \textit{tac}) appears only in isolation, while the unstressed variant \textit{to} occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form \textit{t}, with constricted vowel, is used in counting, while a low-tone open syllable \textit{ti} occurs otherwise. See also the unstressed form for \textit{ONE} in Kayah (§1.2.1), as well as Dulung \textit{ti}^55 (Sun 1982:244-245), and Karenic forms like Pa-O \textit{tä}^59-ha, Palaychi \textit{ta}-, Sgaw \textit{ta}-.

\footnote{In Modern Burmese the fully stressed form \textit{ti} (the regular reflex of WB \textit{tac}) appears only in isolation, while the unstressed variant \textit{to} occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form \textit{t}, with constricted vowel, is used in counting, while a low-tone open syllable \textit{ti} occurs otherwise. See also the unstressed form for \textit{ONE} in Kayah (§1.2.1), as well as Dulung \textit{ti}^55 (Sun 1982:244-245), and Karenic forms like Pa-O \textit{tä}^59-ha, Palaychi \textit{ta}-, Sgaw \textit{ta}-.

forms like Karok \textit{tik} 'hand; finger', Yagua \textit{tiki} 'one', and Eskimo \textit{tik-iq} 'index finger', as well as Nilo-Saharan forms like Maba \textit{tek}, Fur \textit{dik} 'one'. For a critique of Greenberg's unfettered approach to linguistic comparison, see Matossif (1990a).

\footnote{Other possible reflexes of \textit{*tyak} in languages of Nagaland are Phom \textit{huk}, and the 2nd syllable of Wancho \textit{tu-ta}. For the first syllable of this Wancho form, see below.}

\footnote{French (1983:529) sets up a Proto Northern-Naga root \textit{*kla}, to which he assigns the Konyak and Wancho forms, as well as the second syllable of Yogli \textit{sa kha} '100' ('100 x 1'), suggesting a connection with Jg. \textit{ma}^55 \textit{kha}^31 'all; whole' (Dai et al., 1983:455). (An obvious alternative source for the Yogli syllable would be \textit{*ka} (§3.1.3).}

\footnote{In Modern Burmese the fully stressed form \textit{ti} (the regular reflex of WB \textit{tac}) appears only in isolation, while the unstressed variant \textit{to} occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form \textit{t}, with constricted vowel, is used in counting, while a low-tone open syllable \textit{ti} occurs otherwise. See also the unstressed form for \textit{ONE} in Kayah (§1.2.1), as well as Dulung \textit{ti}^55 (Sun 1982:244-245), and Karenic forms like Pa-O \textit{tä}^59-ha, Palaychi \textit{ta}-, Sgaw \textit{ta}-.

\footnote{In Modern Burmese the fully stressed form \textit{ti} (the regular reflex of WB \textit{tac}) appears only in isolation, while the unstressed variant \textit{to} occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form \textit{t}, with constricted vowel, is used in counting, while a low-tone open syllable \textit{ti} occurs otherwise. See also the unstressed form for \textit{ONE} in Kayah (§1.2.1), as well as Dulung \textit{ti}^55 (Sun 1982:244-245), and Karenic forms like Pa-O \textit{tä}^59-ha, Palaychi \textit{ta}-, Sgaw \textit{ta}-.

\footnote{In Modern Burmese the fully stressed form \textit{ti} (the regular reflex of WB \textit{tac}) appears only in isolation, while the unstressed variant \textit{to} occurs in the stream of speech (e.g. before classifiers). Similarly, in Akha the stressed form \textit{t}, with constricted vowel, is used in counting, while a low-tone open syllable \textit{ti} occurs otherwise. See also the unstressed form for \textit{ONE} in Kayah (§1.2.1), as well as Dulung \textit{ti}^55 (Sun 1982:244-245), and Karenic forms like Pa-O \textit{tä}^59-ha, Palaychi \textit{ta}-, Sgaw \textit{ta}-.}
developed several co-existent variants (much like English one, an, only, and so on; see §3.1), for example, Lahu ṭe ‘one’, ḍē-Śē ‘all’, ū ‘only’, ūʔ-Śhī ‘nothing’, a-Śf ‘little bit’ < PLB *da-y, *nday, *ʔdík, *de-k, and *ʔgyik, respectively.\(^{10,11}\)

It is actually far from certain that Lahu ṭe ‘one; whole; alan’ and ḍē-Śē ‘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 *da-y *tay, underlying forms like Jingpho tai ‘single’, atai ‘one, as of a pair’, guntai ‘single’, shingtai ‘only’; Boro ota-y ‘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 *tàn ʬ ‘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  
atir ~ 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 *ʔgyik ‘little bit’ offered in TSR #70 include WT cíg ‘a little, few, some’, WB kyac ‘be diminutive, smaller than ordinary’, Lahu a-ci. Akha á-cyq, and Moso tsi.\(^ {55}\)

\(^{11}\) Several PLB variants are reconstructed in Matisoff (1972a) (‘TSR’) #31/#48 and #70: *-tik ≠ *ti ≠ *ʔdík ≠ *ʔ-gyik ≠ *kyik. A revised analysis is offered in Matisoff (1985b) (‘GSTC’) #148, where there is reconstructed a new PST word-family of the shape *da-y ≠ *tay ≠ *da-n ≠ *tan. Still another alloform *ty is recognised in GSTC to underlie forms like WB th * ‘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 *da-y ≠ *tay group include Monpa (Dubey) thee, Nocte wan-the (Dubey, Das Gupta), van-the (GEM) (for the first syllable see *hag, §3.1.5.3), Ersu teF, Pumi ti.\(^ {13}\) (Qinghua); ti\(^ {13}\) (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  kesi (with prefixal k-)
Dimasa  se16
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  sa ‘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

Serdrukpen  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 sā) is “high-pitched”, while the nearly homophonous Lakher word for ‘thousand’ is “low-pitched” (sā). Undoubtedly cognate to these Kamarupan forms is Jingpho sā ‘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', *van-yi* '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*\(^2\) (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 \(*lrig \approx lryag\) (§3.2.3.3; §3.4):

- **Pwo Karen**  
  *lån* (Jones 1961:618)
- **Yimchunguru**  
  *khū-lang* (first syllable is a prefix)
- **Sangtam**  
  *khūrū* (also *khe*)
- **Kezhama**  
  *kele* (*ke* is a prefix, part of a 1–3 run)
- **Mao**  
  *kali* (*ka* is a prefix, part of a 1–3 run)
- **Sema**  
  *la ki* (also *khe*)
- **Maram**  
  *hang-li-ne* (cf. *hang, §3.1.5.3*)

It is 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 *ti*\(^{31}\), Mawo *ti* (Sun 1981:217), clearly from the *t(y)i-* family (§3.1.4).
Deng Geman \( ku^3/mu^3 \) (Sun et al. 1980:252)
Kaman (Miju Mishmi) \( ku-mo \) (Das Gupta 1977a)
\( kmo \sim kômo \) (LSI 3.1:623)

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 \( ni-pal \sim ni-pan \) ‘8’ (‘2 from 10’); cf. *ban *bal ‘ten’, §4.2.0.3.

(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, 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. Often a language will maintain several

---

20 \( ku-/ku-/k- \) 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’, \( mà \) (Matusoff 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 of. 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 alloforms -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 tšip, 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- variation24, as witness these forms from an AMD language: Kaman (Miju Mishmi) kyap-mo (LSI has kap), Deng Geman kiap55 mu54.

The obvious Chinese cognate is +, reconstructed as OC *djəp 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 tši ~ sī, Namsang (= Nocte) i-tši, Moshang rok-ši, Garo tši, Dimasa dži, Miju si (in composition), Karen (Taungthu) tši, (Pwo and Sgaw) shi.

The vocalism of WB ačhai 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-thambaq, Tsangla/Sharchop s(h)e, Kanawari săi.27 Here belong several other Himalayish forms with -u vocalism:28 WT bēu (Lhasa cu), Kailke chyu, Gurung cyuq, Thakali cyu, Jirel cyu-tambaq, Sikkim Bhutia chu-tamba, Dzongkha (Mazaudon) cu-thām.29

(Kamarupan) Monpa (Cuona) tci53, (Dubey) chi; Monpa (Motuo) se; Garo chi-kung (Phillips), ci-king (Burling); Kokborok ĉi

(Bai) Bai (= Minchia) tsw8 (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’, aši ‘one’, anī ‘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-lay (STC 463), but WT gəl), 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 tshe\textsuperscript{55}, Proto Rgyalrong *sytsye (Nagano 1984) \(< *\text{-}t\text{si}y\text{iy}\) (JAM)

(Loloish) Proto Loloish *tši (> Lahu chi, Akha tsé, Lisu htsì\textsuperscript{4}, Phunoi tásé (ta- ‘one’), and so on.

(Nungish) Dulong tsái\textsuperscript{55}, Rawang hti sel, Nujiang ts’i\textsuperscript{55} tshān\textsuperscript{530} (first syllables mean ‘one’). Have these curious Nungish forms with final -1 developed from *-y, or do they point to an allofam *tsyl?\textsuperscript{30}

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,\textsuperscript{31} both as the independent numeral for TEN, and as the first constituent in higher multiples thereof:


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’, čan-ŋo ‘fifty’, čem-piŋ ‘eighty’).\textsuperscript{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 *bon, though he suggests that this might be a “loan from Austroasiatic into Northern Naga”, citing Khasi ši pōn ‘ten’ (ši means ‘one’).\textsuperscript{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 *-I.\textsuperscript{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).

\textsuperscript{30} Undoubtedly this -n is from an earlier lateral *-I.

\textsuperscript{31} See Ono (1965).

\textsuperscript{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.

\textsuperscript{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 *rjwŋ), but it follows the unit. See §3.2.3.3.

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

\textsuperscript{35} The independent word for TEN in Meithei is tara, whose affiliations are elsewhere (§3.2.3.3(a)).
3.2.3.3 *s-rl/lig and *s-rl/lyag

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 erîŋ ~ erîaj (Das Gupta; note the intralingual variation of rhyme), reng-cheng (Yano), raŋ (Robinson), il-li (E. Dafla), Apatani alyâ (for humans) ≠ lyâ (for non-humans; < *ly-ŋ); Gallong ɬri~ ɬyi~; Padam (Dubey) ɬyi, ɬi; Aka (Hruso) rhi, ru; Taraon ha:long, Darang Deng xa55lun55; Idu hû (Talukdar et al. 1962), hong55hûng55 (Luoba: Sun 1983); Chulikata hush (< *hu-shV, with vowel of second syllable apocopated36)

J.T. Sun (1993:144) sets up Proto Tani *rjwI on the basis of Bengni ɯ-rjwŋ, Lhopa/Bokar ujwìng, and the above Abor-Miri (=Padam-Mising) forms, also citing Dhammai lîn, Bangru røŋ53, and Idu/Luoba (ZMYYC) hJoI55 (used in multiples of ten, for example, ni55hJoI55 ‘20’, a35hJoI55 ‘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 hû ≠ hûŋ55), 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 *lhyäŋ ‘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 høŋ- (for example, høŋ-phanga ‘50’), Southern Rengma hê (for example, hem-pfû ‘50’), Angami (Kohima) he- (for example, hie-pengou ‘50’), and Chokrû (=Chokri) he- (for example, hie-pûngu ‘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 ≠ *ryâŋ, or rather to the stop-finalled PNN etymon *rok discussed below (§3.2.3.4):

(a) With velar prefix:

Angami kerû, Chokri kûri, Chakesang keri, Liangmai kariu, Maram kero, Mzieme/Zeliang kerei, Zeme kereu, Mikir kre-39

36 This is an apocopating language. Compare Chulikata kâsh ‘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).
(b) With dental prefix:

Ao (Mongsen) tera, Ao (Chungli) ter (with apocope)\(^40\); Khoirao sara, Lotha taro, Meithei tara, Meluri tara, Ntenyi dagha, ta’a\(^41\); Pochury tūra, Rengma tsarū, Sangtam thūre, Tangkhul thara, Yacham-Tengsa thelu, Yimchungru thirū

(c) With palatal affricate word-initially:

Kezhama chiro, Mao chūro, Sema chiūghi

The first syllables in these last three forms require some comment. On the one hand, they bear a superficial resemblance to reflexes of *tsiy *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’, chūro ‘10’). The second syllable of Sema chiūghi (where the ‘gh’ presumably stands for [y]) 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 pherāṇg, 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\(^2\) ‘10’ (‘10 x 1’), rok-tachat ‘80’;
- Tangsa (Yogli) rauk-shi ‘10’, rauk-tūchat ‘80’;
- Kimsing ro-shi ‘10’, ro-bangi ‘50’ (Das Gupta 1978);
- Nocte i-chi ‘10’ (< *tsiy), 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-cūruk ‘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 -hra:w ‘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-rj *s-ryag *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-ba:q, Jirel cyu-ta:m-ba:q 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) ka55tu55, Pumi (Qinghua) qa55stie55, Qiang (Taoping) xa21dy33 and Qiang (Mawo) hadiu.

3.2.3.6 *p/bop

Several Kiranti languages (East Nepal) have multiplicative morphemes that occur in compounds for the multiples of ten, reflecting Proto Kiranti *pog or *bog:


Limbu also has an interesting form i-boog ‘NINE’, which looks as if it may have been transvalued or ‘downstepped’ from an earlier meaning of TEN (compare Yakkhaba ip-pog). 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/bog 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 Itams-pa 'be full', tham-pa them-pa 'complete, full', Idem-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 alofamically related to this root is Lepcha t'ap (< *tap, with homorganic final stop), used in teen-formation, e.g. sam-t'ap ‘13’, târäk-t'ap ‘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’)

(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 *pat* ‘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 /fi:f/ but fifteen /fi:f-*/; three /ˈtrei/ but thirteen /ˈθɜːr-*/)? 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 *duodēviginti* ‘18’, *undēviginti* ‘19’)?

---

• Strictly speaking the concept of TEENS is only relevant to decimal systems of numerals. In vigesimal systems, the numbers 1–19 correspond to the ‘units’ 1–9 of decimal systems. In vigesimal systems, it is advantageous to be able to refer to the numbers from 21 to 39 as a group – I suggest the term TWENTEENS. The twenties 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

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 MORPHOPHONEMICALLY RELATED TO THE INDEPENDENT NUMERAL TEN

<table>
<thead>
<tr>
<th>Language</th>
<th>TEN</th>
<th>-TEEN</th>
<th>3</th>
<th>13</th>
<th>5</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angami (Kho.)</td>
<td>keru</td>
<td>kero-</td>
<td>se</td>
<td>kerose</td>
<td>pengu</td>
<td>keropengu</td>
</tr>
<tr>
<td>Angami (Koh.)</td>
<td>keru</td>
<td>kere-</td>
<td>se</td>
<td>kerose</td>
<td>pengou</td>
<td>kerepengou</td>
</tr>
<tr>
<td>Ao (Chungli)</td>
<td>ter</td>
<td>ter(i)-</td>
<td>asem</td>
<td>terasem</td>
<td>pungu</td>
<td>teripungu</td>
</tr>
<tr>
<td>Ao (Mongsen)</td>
<td>tera</td>
<td>tera-</td>
<td>asam</td>
<td>terasam</td>
<td>phanga</td>
<td>teraphanga</td>
</tr>
<tr>
<td>Chokri</td>
<td>kuri</td>
<td>kuri-</td>
<td>su</td>
<td>kurisu</td>
<td>pungu</td>
<td>kuri-pungu</td>
</tr>
<tr>
<td>Meluri</td>
<td>ter</td>
<td>tera-</td>
<td>keche</td>
<td>terakeche</td>
<td>manga</td>
<td>teramanga</td>
</tr>
<tr>
<td>Nocte</td>
<td>ichi</td>
<td>ichi-</td>
<td>vanram</td>
<td>ichivanram</td>
<td>bangga</td>
<td>ichibanga</td>
</tr>
<tr>
<td>Ntenyi</td>
<td>ta’a</td>
<td>ta’-</td>
<td>keching</td>
<td>ta’akecham</td>
<td>munga</td>
<td>ta’amanga</td>
</tr>
<tr>
<td>Sangtam</td>
<td>thure</td>
<td>thure-</td>
<td>asang</td>
<td>thureasang</td>
<td>munga</td>
<td>thuremunga</td>
</tr>
<tr>
<td>Sema</td>
<td>chughi</td>
<td>chughi-</td>
<td>kuthu</td>
<td>chughikuthu</td>
<td>pongu</td>
<td>chughipongu</td>
</tr>
</tbody>
</table>

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 Morphe Me Occurs Between the TEN and the UNIT

<table>
<thead>
<tr>
<th>TEN</th>
<th>LINKER</th>
<th>3</th>
<th>13</th>
<th>5</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang</td>
<td>an -tak-</td>
<td>sam</td>
<td>antaksam</td>
<td>ngau</td>
<td>antakngau</td>
</tr>
<tr>
<td>Khoirao</td>
<td>sara -na-</td>
<td>katham</td>
<td>charanakasum</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Konyak</td>
<td>pen -me-</td>
<td>lem</td>
<td>penmelem</td>
<td>nga</td>
<td>penmenga</td>
</tr>
<tr>
<td>Lotha</td>
<td>taro -si-</td>
<td>etham</td>
<td>tarosietham</td>
<td>mungo</td>
<td>tarosimungo</td>
</tr>
<tr>
<td>Moshang</td>
<td>rokshi -ra-</td>
<td>atum</td>
<td>rokshiraatum</td>
<td>bangang</td>
<td>rokshirabangga</td>
</tr>
<tr>
<td>Nruanghmei</td>
<td>ruh -na-</td>
<td>katham</td>
<td>ruhnakatham</td>
<td>pangu</td>
<td>ruhnapanung</td>
</tr>
<tr>
<td>Phom</td>
<td>an -pu-</td>
<td>jam</td>
<td>anpujam</td>
<td>nga</td>
<td>anpunung</td>
</tr>
<tr>
<td>Tangkhul</td>
<td>thara -da-</td>
<td>katham</td>
<td>tharakatham</td>
<td>phangang</td>
<td>tharakaphangga</td>
</tr>
<tr>
<td>Wancho</td>
<td>ban -ba-</td>
<td>ajam</td>
<td>banbajam</td>
<td>aga</td>
<td>banbaga</td>
</tr>
<tr>
<td>Y-Tengsa</td>
<td>thelu -le-</td>
<td>asam</td>
<td>talulesam</td>
<td>phungug</td>
<td>talulephungu</td>
</tr>
<tr>
<td>Yimchungru</td>
<td>thu -kheak-</td>
<td>asam</td>
<td>thurukheakasum</td>
<td>phungug</td>
<td>thurukheakphungu</td>
</tr>
<tr>
<td>Zeme</td>
<td>kereu -ze-</td>
<td>kechum</td>
<td>kereuzekechum</td>
<td>mengeg</td>
<td>kereuzemengeu</td>
</tr>
</tbody>
</table>

### 3.3.1.3 Where the Linking Morphe Me Comes After the UNIT

<table>
<thead>
<tr>
<th>TEN</th>
<th>LINKER</th>
<th>3</th>
<th>13</th>
<th>5</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Liangmai</td>
<td>kariu -kiu</td>
<td>shum</td>
<td>karushumkiu</td>
<td>mangiu</td>
<td>kariumangiukiu</td>
</tr>
<tr>
<td>Mao</td>
<td>churo -o</td>
<td>kosu</td>
<td>churokosu-o</td>
<td>pongo</td>
<td>churopongo-o</td>
</tr>
<tr>
<td>Meithei</td>
<td>ara -thoi</td>
<td>ahum</td>
<td>tarahumthoi</td>
<td>manga</td>
<td>taramanga</td>
</tr>
<tr>
<td>Mzieme</td>
<td>kerei -ngkei</td>
<td>ketsum</td>
<td>kereketsumngkei</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Puiron</td>
<td>som -to</td>
<td>thum</td>
<td>somthumo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rengma</td>
<td>tsaru -chu</td>
<td>keshan</td>
<td>tsarukeshanchu</td>
<td>pfu</td>
<td>tsarupfuchu</td>
</tr>
</tbody>
</table>

### 3.3.1.4 Where the Combining Form (’-TEEN’) Is a Different Etymon From the Independent Numeral TEN

<table>
<thead>
<tr>
<th>Mikir</th>
<th>TEN</th>
<th>3</th>
<th>13</th>
<th>5</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>kep</td>
<td>-TEEN</td>
<td>3</td>
<td>13</td>
<td>5</td>
<td>15</td>
</tr>
</tbody>
</table>

50 ‘13’ and ‘12’ are the only Khoirao teens to be found in Marrison. Note the change in the consonant of the UNIT morpheme (katham > -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

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>2</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>hangline</td>
<td>kerui-kaniko</td>
<td>hang-na</td>
<td>nangko</td>
<td>hang-tum</td>
</tr>
</tbody>
</table>

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 allotform of the -ne of hangline. 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.

<table>
<thead>
<tr>
<th>Angami (Khonoma)</th>
<th>Ao (Mongsen)</th>
<th>Meluri</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIXTEEN</td>
<td>(kerosuru)</td>
<td>mukyimupenterok</td>
</tr>
<tr>
<td>SEVENTEEN</td>
<td>mekupomothena</td>
<td>mukyimupenti</td>
</tr>
<tr>
<td>EIGHTEEN</td>
<td>mekupomothetha</td>
<td>mukyimupentsit</td>
</tr>
<tr>
<td>NINETEEN</td>
<td>mekupomothekeu</td>
<td>mukyimupentuku</td>
</tr>
<tr>
<td>TWENTY</td>
<td>muku</td>
<td>mukyi</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ntenyi</th>
<th>Pochury</th>
<th>Rengma</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIXTEEN</td>
<td>kwushetuo</td>
<td>mkeshuntoro</td>
</tr>
<tr>
<td>SEVENTEEN</td>
<td>kwushetughu</td>
<td>mkeshunturu</td>
</tr>
<tr>
<td>EIGHTEEN</td>
<td>kwushetuzza</td>
<td>mkeshuntuze</td>
</tr>
<tr>
<td>NINETEEN</td>
<td>kwushetukhu</td>
<td>mkeshuntoku</td>
</tr>
<tr>
<td>TWENTY</td>
<td>mekweru/mukwung</td>
<td>mke</td>
</tr>
</tbody>
</table>

These formations are subtractive in a different sense from, for example, Latin duodēviginti ‘18’ and undēviginti ‘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’, somkhana ‘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

<table>
<thead>
<tr>
<th>ONE</th>
<th>kat</th>
<th>THIRTEEN</th>
<th>sam-t’ap</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEN</td>
<td>kati</td>
<td>FOURTEEN</td>
<td>fali-t’ap</td>
</tr>
<tr>
<td>ELEVEN</td>
<td>kati-kat-t’ap</td>
<td>SIXTEEN</td>
<td>ṯaṟak-t’ap</td>
</tr>
<tr>
<td>TWELVE</td>
<td>kati-nyat-t’ap</td>
<td>NINETEEN</td>
<td>dakyot-t’ap</td>
</tr>
</tbody>
</table>

Lepcha ‘11’ and ‘12’ are formed with the linking morpheme t’ap added to the independent numeral for TEN plus the units ONE (kat) and TWO (nyat). However, in the higher teens the independent morpheme for TEN disappears, and the t’ap 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 te ‘1’, te-chi ‘10’), in the same way as the other round numbers (for example, Lahu ni ‘2’, ni-chi ‘20’). Teens are then formed by adding the UNIT morpheme after the classifier for TEN, for example, Lahu:

\[
\begin{align*}
ṯe\ chi\ ṯe & \quad ‘11’ \ ‘(1 \times 10) + 1’ \\
\ \ ṯe\ chi\ kh3 & \quad ‘16’ \ ‘(1 \times 10) + 6’.
\end{align*}
\]

In Burmese, TEN is also a classifier with respect to the round numbers: WB tac ‘1’, tə-chay ‘10’; hnae ‘2’, hnae-chay ‘20’ (tə- and hnae- 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-hnae ‘12’, chay-khrok ‘16’.

Alternatively, TEN may function as an ordinary numeral, as in Akha tshe ‘10’ (with ONE not expressed), tshe-ti? ‘11’, tshe-kɔ̀ ‘16’.

58 I suspect that this t’ap 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)

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


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”:

<table>
<thead>
<tr>
<th>Number</th>
<th>ONE</th>
<th>TWO</th>
<th>THREE</th>
<th>FOUR</th>
<th>FIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>se</td>
<td>nay</td>
<td>tam</td>
<td>bra</td>
<td>ba</td>
</tr>
<tr>
<td>SIX</td>
<td>zokkay-se kanay</td>
<td>zokkay-se katam</td>
<td>zokkay-nay</td>
<td>zokkay-nay kase</td>
<td>dos (&lt; Indo-Aryan)</td>
</tr>
<tr>
<td>SEVEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EIGHT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NINE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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-nay 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’):

<table>
<thead>
<tr>
<th>TEEN</th>
<th>zokkay-tam kase</th>
<th>‘(4 x 3) + 1’</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIFTEEN</td>
<td>zokkay-tam kanay</td>
<td>‘(4 x 3) + 2’</td>
</tr>
<tr>
<td>SIXTEEN</td>
<td>zokkay-brə</td>
<td>‘4 x 4’</td>
</tr>
</tbody>
</table>

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


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 katay ‘be extra’, khamatay ‘increase, multiply’, akatay ‘remnant’; Wancho a-tai ‘far’, tai-hu ‘many’, and so on.

[D] ABOR-MIRI-DALFA

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-):


Idu hū ‘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)i;k; §3.1.5.2 *s(h)e; §3.1.5.5 *(k-)IV(N); §3.2.3.3 *s-riŋ ≠ *s-ryan; 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 kofka IN ABOB-MIRI-DAFLA, HIMALAYISH, AND NAPA

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-na ‘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 kyp-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 tsek/ s(a) IN RGYALRONG

In the Zida dialect of Rgyalrong(data from Chang Kun), ‘10’ is stshi; what seems to be the same initial element occurs in satshek ‘11’, stanes ‘12’, and presumably all the higher teens as well. It could well be that this prefix is a reduced form of the independent numeral tsek ‘1’, so that stshi meant ‘1 x 10’ (i.e. < *tsek-tshi); in the teens, however, where it co-occurs with the UNIT morphemes, the s- then came to mean ‘10’ (-TEEN). (Note that this analysis implies that an older form of ‘11’ was something like *tsek-tsek, with subsequently greater and greater destressing of the first syllable.)
3.4.4 *s(h)e ‘1’ and *ts(y)i ‘10’

Given the high degree of phonetic similarity between *s(h)e ‘1’ (§3.1.5.2) and *ts(y)i ‘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, Quebec) 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:

Danish68

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>TEN</td>
<td>ti</td>
</tr>
<tr>
<td>TWENTY</td>
<td>tyve</td>
</tr>
<tr>
<td>THIRTY</td>
<td>tredive</td>
</tr>
<tr>
<td>FORTY</td>
<td>fyrre(tyve)</td>
</tr>
<tr>
<td>FIFTY</td>
<td>halvtreds(indstyve)</td>
</tr>
<tr>
<td>SIXTY</td>
<td>tres(indstyve)</td>
</tr>
<tr>
<td>SEVENTY</td>
<td>halvfjerds(indstyve)</td>
</tr>
<tr>
<td>EIGHTY</td>
<td>firs(indstyve)</td>
</tr>
<tr>
<td>NINETY</td>
<td>halvfems(indstyve)</td>
</tr>
<tr>
<td>HUNDRED</td>
<td>hundrede</td>
</tr>
</tbody>
</table>

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, firs-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.


68 Data from Koefoed (1958). My thanks to Gary Holland for some further elucidation.
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-ku?\(^71\), 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):

<table>
<thead>
<tr>
<th>Angami (Khonoma)</th>
<th>meku</th>
<th>Meluri</th>
<th>mukwe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angami (Kohima)</td>
<td>mepfu</td>
<td>Nruanghmei</td>
<td>ncuix</td>
</tr>
<tr>
<td>Ao (Chungli)</td>
<td>metsu</td>
<td>Ntenyi</td>
<td>makweru/mukwung</td>
</tr>
<tr>
<td>Ao (Mongsen)</td>
<td>mukyi</td>
<td>Pochury</td>
<td>mke</td>
</tr>
<tr>
<td>Chokri</td>
<td>mechi</td>
<td>Rengma</td>
<td>nki</td>
</tr>
<tr>
<td>Khoirao</td>
<td>machi</td>
<td>Sangtam</td>
<td>mukyu</td>
</tr>
<tr>
<td>Liangmai</td>
<td>makai</td>
<td>Sema</td>
<td>muku</td>
</tr>
</tbody>
</table>

\(^69\) 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...’
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’ and *kun ‘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 khye75.

To these we should probably add Monpa (Motuo) (AMD group) k’ai, as in k’ai-ga ‘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-sap ‘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 khe, 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 pi- ~ bi- ~ pi- ~ bet-;

[B] Yacham-Tenga 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 ni’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’:

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:

Nocte vanyi ichi ruak- ruaknyi ruakram ruakbeli ruakbanga
Maring khani chip som- somni somthun 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 maku (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 Otomangean *kala! (Pers.comm. Terrence S. Kaufman, April 1994.)

78 The independent Wancho word for ‘20’ is caor 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 ‘somnga’ for ‘20’, though this certainly seems to be an error, since sompa 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'

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>30</td>
<td>40</td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td></td>
</tr>
<tr>
<td>Yimchungru</td>
<td>thürü</td>
<td>samrü</td>
<td>yirü</td>
<td>rukrü</td>
<td>nierü</td>
<td>zharü</td>
<td>kurü</td>
</tr>
</tbody>
</table>

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 $3 \times 10$, 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 $10 \times 4, 10 \times 5...10 \times 9$, 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-ley* ≠ *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').
<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>3</th>
<th>30</th>
<th>'TY'</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angami (Khonoma)</td>
<td>keru</td>
<td>se</td>
<td>serü</td>
<td>lhi-</td>
</tr>
<tr>
<td>Angami (Kohima)</td>
<td>keru</td>
<td>se</td>
<td>serü</td>
<td>hie-</td>
</tr>
<tr>
<td>Chokri</td>
<td>kūr</td>
<td>i</td>
<td>sii</td>
<td>—</td>
</tr>
<tr>
<td>Kezhama</td>
<td>chiro</td>
<td>katsu</td>
<td>—</td>
<td>lha-</td>
</tr>
<tr>
<td>Khoirao</td>
<td>sara</td>
<td>kathum</td>
<td>thumra</td>
<td>ra-/re(k)</td>
</tr>
<tr>
<td>Liangmai(^1)</td>
<td>kariu</td>
<td>shum</td>
<td>samriu</td>
<td>ri(a)-</td>
</tr>
<tr>
<td>Mao</td>
<td>chūro</td>
<td>kosü</td>
<td>shüro</td>
<td>ri-</td>
</tr>
<tr>
<td>Maram</td>
<td>kerei</td>
<td>ketsum</td>
<td>tsamrei</td>
<td>riak-</td>
</tr>
<tr>
<td>Mzieme</td>
<td>kerei</td>
<td>kethum</td>
<td>tumru</td>
<td>rek-</td>
</tr>
<tr>
<td>Nruanghmei</td>
<td>ruh</td>
<td>kathum</td>
<td>tūmru</td>
<td>rek-</td>
</tr>
<tr>
<td>Rengma</td>
<td>tsarü</td>
<td>keshan</td>
<td>shenrü</td>
<td>en-</td>
</tr>
<tr>
<td>Sema(^2)</td>
<td>chūghi</td>
<td>kithu</td>
<td>sheghi</td>
<td>lho-</td>
</tr>
<tr>
<td>Tangkhul</td>
<td>thara</td>
<td>kathum</td>
<td>thumra</td>
<td>hang-</td>
</tr>
<tr>
<td>Zeme</td>
<td>kereu</td>
<td>kechum</td>
<td>himreu</td>
<td>he-/re/-riak(^3)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>Angami (Khonoma)</td>
<td>lhida</td>
<td>lhipengu</td>
<td>lhisuru</td>
</tr>
<tr>
<td>Angami (Kohima)</td>
<td>hiede</td>
<td>hiepengou</td>
<td>hiesorou</td>
</tr>
<tr>
<td>Chokri</td>
<td>hieda</td>
<td>hiepengu</td>
<td>hieshwuru</td>
</tr>
<tr>
<td>Kezhama</td>
<td>—</td>
<td>lhapangu</td>
<td>—</td>
</tr>
<tr>
<td>Khoirao</td>
<td>ramri</td>
<td>renga</td>
<td>reksaruk</td>
</tr>
<tr>
<td>Liangmai</td>
<td>atai</td>
<td>ringiu</td>
<td>riacharuk</td>
</tr>
<tr>
<td>Mao</td>
<td>ridei</td>
<td>ripongo</td>
<td>richoro</td>
</tr>
<tr>
<td>Maram</td>
<td>ragdai</td>
<td>rengo</td>
<td>reksaruk</td>
</tr>
<tr>
<td>Mzieme</td>
<td>riakdai</td>
<td>riangnge(^4)</td>
<td>riakheruk</td>
</tr>
<tr>
<td>Nruanghmei</td>
<td>rekdai</td>
<td>rekngu</td>
<td>rekčuruk</td>
</tr>
<tr>
<td>Rengma</td>
<td>henzi</td>
<td>hempfü</td>
<td>hentsaro</td>
</tr>
<tr>
<td>Sema</td>
<td>lhobdhi</td>
<td>lhopengu</td>
<td>lhotsogho</td>
</tr>
<tr>
<td>Tangkhul</td>
<td>hangmati</td>
<td>hangphanga</td>
<td>hangtharuk</td>
</tr>
<tr>
<td>Zeme</td>
<td>hedai</td>
<td>rengeu</td>
<td>riakseruk</td>
</tr>
</tbody>
</table>

---

\(^1\) Liangmai '40' does not exactly fit the pattern; it has the prefix a- instead of ri(a)- ('4' is madaJ); '70, 80, 90' all have ria- (riachania, riatachad, riachakiu).

\(^2\) Note küthu 'three' ≠ she- 'thir-'.

\(^3\) Zeme '70, 80, 90' have riak: riaksena, riakdesat, riaksekui.

\(^4\) 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

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>3</th>
<th>‘.TY’</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nocte</td>
<td>ichi</td>
<td>vanram</td>
<td>ruak-</td>
<td>ruakram</td>
<td>ruakbeli</td>
<td>ruakbangla</td>
</tr>
</tbody>
</table>

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\(^\text{85}\) (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]):

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>10</th>
<th>2</th>
<th>20</th>
<th>3</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lahu</td>
<td>të</td>
<td>të chi</td>
<td>nî</td>
<td>nî chi</td>
<td>šë?</td>
<td>šë? chi</td>
</tr>
</tbody>
</table>

\[
të\chi\ '10'\ (\text{Num} + \text{Clf}) \quad '1 \times 10' \\
nî\chi\ '20' \quad '2 \times 10' \\
šë?\chi\ '30' \quad '3 \times 10'...
\]

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

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>10</th>
<th>2</th>
<th>20</th>
<th>3</th>
<th>30</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tangsa</td>
<td>ashi</td>
<td>rokshi</td>
<td>ani</td>
<td>rokni</td>
<td>atum</td>
<td>roktum</td>
</tr>
<tr>
<td>(Moshang)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tangsa</td>
<td>ashi</td>
<td>raukshi</td>
<td>anei</td>
<td>rauknei</td>
<td>adim</td>
<td>raukdim</td>
</tr>
<tr>
<td>(Yogli)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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).
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, thiinyang).

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. However, these eight languages have generalised a dental prefix with the numerals all the way from 6 to 10, 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 alloform of the ordinary numeral for FIVE *b-oa or *m-oa. This is not implausible phonologically, since *m-oa 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) 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).

---

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-inglya) 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.
<table>
<thead>
<tr>
<th>Numbers</th>
<th>Words</th>
<th>Numbers</th>
<th>Words</th>
<th>Numbers</th>
<th>Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 (3 x 10)</td>
<td>Ao (Chungli): <em>semer</em></td>
<td>40 (4 x 10)</td>
<td>Ao (Mongsen): <em>lira</em></td>
<td>50 (10 x 5?)</td>
<td>Lotha: <em>tunam</em></td>
</tr>
<tr>
<td>40 (4 x 10)</td>
<td>Ao (Chungli): <em>lir</em></td>
<td>50 (10 x 5?)</td>
<td>Ao (Mongsen): <em>tenem</em></td>
<td>60 (6 x 10)</td>
<td>Meluri: <em>rokra</em></td>
</tr>
<tr>
<td>50 (10 x 5?)</td>
<td>Ao (Chungli): <em>tenem</em></td>
<td>60 (6 x 10)</td>
<td>Ao (Mongsen): <em>rokra</em></td>
<td>70 (7 x 10)</td>
<td>Ntenyi: <em>sangre</em></td>
</tr>
<tr>
<td>60 (6 x 10)</td>
<td>Ao (Chungli): <em>roker</em></td>
<td>70 (7 x 10)</td>
<td>Lotha: <em>tiingya</em></td>
<td>80 (8 x 10)</td>
<td>Pochury: <em>nyure</em></td>
</tr>
<tr>
<td>70 (7 x 10)</td>
<td>Ao (Chungli): <em>roker</em></td>
<td>80 (8 x 10)</td>
<td>Ao (Mongsen): <em>lira</em></td>
<td>90 (9 x 10)</td>
<td>Sangtam: <em>zyure</em></td>
</tr>
<tr>
<td>80 (8 x 10)</td>
<td>Ao (Chungli): <em>roker</em></td>
<td>90 (9 x 10)</td>
<td>Lotha: <em>rura</em></td>
<td></td>
<td>Yimchungru: <em>kuru</em></td>
</tr>
<tr>
<td>90 (9 x 10)</td>
<td>Ao (Chungli): <em>roker</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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). 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 *yangkhay* ‘50’ ‘might have some affiliation with *yaj* ‘backbone; middle of the back’ or ‘middle of the roof in houses’ (i.e. ‘ridgepole”). 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.'
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 s -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’ (zaroo) 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 tera (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 pezi; a more ancient-looking form (< PTB *b-lay) 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:

<table>
<thead>
<tr>
<th></th>
<th>isi</th>
<th>6</th>
<th>throk</th>
<th>60</th>
<th>throk-kep</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>hini</td>
<td>7</td>
<td>throksi</td>
<td>(‘6 + 1’)</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>kethom</td>
<td>8</td>
<td>nirkep</td>
<td>(‘10 – 2’)</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>phir/phli</td>
<td>9</td>
<td>sirkep</td>
<td>(‘10 – 1’)</td>
<td>90</td>
</tr>
<tr>
<td>5</td>
<td>phongo</td>
<td>10</td>
<td>kep</td>
<td>100</td>
<td>pharo</td>
</tr>
</tbody>
</table>

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’.

### 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).

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.

<table>
<thead>
<tr>
<th></th>
<th>Wancho</th>
<th>Phom</th>
<th>Konyak</th>
<th>Chang</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEN</td>
<td>ban</td>
<td>an</td>
<td>pen</td>
<td>an</td>
</tr>
<tr>
<td>TWENTY (indep.)</td>
<td>tsa/ca</td>
<td>ha</td>
<td>ta</td>
<td>sauchie</td>
</tr>
<tr>
<td>TWENTY (in comp.)</td>
<td>pu-</td>
<td>pū- ~ bī-</td>
<td>ta- ~ te-</td>
<td>sau-</td>
</tr>
<tr>
<td>FORTY</td>
<td>punyi</td>
<td>pinnyi</td>
<td>teija</td>
<td>saunyi</td>
</tr>
<tr>
<td>SIXTY</td>
<td>puram; hujam</td>
<td>pūjam</td>
<td>telemja</td>
<td>sausam</td>
</tr>
<tr>
<td>EIGHTY</td>
<td>puli</td>
<td>būali</td>
<td>tepelija</td>
<td>saulei</td>
</tr>
<tr>
<td>THIRTY</td>
<td>ca-ban</td>
<td>ha-pūan</td>
<td>ta-pen</td>
<td>kujih</td>
</tr>
</tbody>
</table>

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”.

---

<sup>95</sup> 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.

<sup>96</sup> Compare French quatre vingts ‘80’, quatre-vingt-dix ‘90’ (‘(4 x 20) + 10’), and the Danish system discussed above (§3.5.1).

<sup>97</sup> 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”.
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’). The final elements -sam, -lei, -(i)ngau are the morphemes ‘3, 4, 5’; the first syllable *an- ‘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):

<table>
<thead>
<tr>
<th></th>
<th>Meithei</th>
<th>Yacham-Tengsa</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEN</td>
<td>tara</td>
<td>thelu</td>
</tr>
<tr>
<td>TWENTY (independent)</td>
<td>kul</td>
<td>machi/tamong</td>
</tr>
<tr>
<td>TWENTY (in composition)</td>
<td>-phu</td>
<td>machi/-tamong/-mesung-</td>
</tr>
<tr>
<td>FORTY</td>
<td>niphu</td>
<td>mesung-anat</td>
</tr>
<tr>
<td>SIXTY</td>
<td>humphu</td>
<td></td>
</tr>
<tr>
<td>EIGHTY</td>
<td>mariphu</td>
<td>tamong-phule</td>
</tr>
<tr>
<td>THIRTY</td>
<td>kun-thra</td>
<td>machi-li-thelu</td>
</tr>
<tr>
<td>FIFTY</td>
<td>yangkhei</td>
<td>tamong-anat-tule-thelo</td>
</tr>
<tr>
<td>SEVENTY</td>
<td>humphu-tara</td>
<td>tamong-asam-tule-thelo</td>
</tr>
<tr>
<td>NINETY</td>
<td>mariphu-tara</td>
<td>tamong-phuicu-le-thelu</td>
</tr>
<tr>
<td>HUNDRED</td>
<td>cha</td>
<td>mesung-phung</td>
</tr>
</tbody>
</table>

---

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 (pū-), 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/yaŋkhay ‘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

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:th, a:th, nau ‘6–9’, but sa:thi, sattari, asi, nabbe ‘60–90’. Although these were

---

99 The data on Chepang, Gurung, Jirel, Kaire, Khainge, 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/Grunwedel (1898) (Lepcha), Mazaudon (1985) (Dzongkha), Chhwang Rinzin (pers.comm. 1984) (Dzongkha, Champor), 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:

<table>
<thead>
<tr>
<th>Wr. Tibetan</th>
<th>Thakali</th>
<th>Gurung</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 āču</td>
<td>cyu</td>
<td>cyuq</td>
</tr>
<tr>
<td>3</td>
<td>gsum</td>
<td>soq</td>
</tr>
<tr>
<td>30</td>
<td>sum-ču</td>
<td>soq-jyu</td>
</tr>
<tr>
<td>4</td>
<td>bži</td>
<td>plih</td>
</tr>
<tr>
<td>40</td>
<td>bži-āču</td>
<td>plih-cyu</td>
</tr>
<tr>
<td>5</td>
<td>lga</td>
<td>ngah</td>
</tr>
<tr>
<td>50</td>
<td>lga-āču</td>
<td>ngah-cyu</td>
</tr>
</tbody>
</table>

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):

\[
\begin{align*}
2 & \text{ ni-gu:} & 20 & \text{ni:-gu:} & / < \text{ni + i}/ \\
3 & \text{ swa-gu:} & 30 & \text{swi:-gu:} & / < \text{swa + i}/ \\
4 & \text{ pe-gu:} & 40 & \text{pi:-gu:} & / < \text{pe + i}/ \\
5 & \text{ nya:-gu:} & 50 & \text{nyae-gu:} & / < \text{nya: + i}/ \\
6 & \text{ khu-gu:} & 60 & \text{khwi:-gu:} & / < \text{khu + i}/ \\
9 & \text{ gu-gu:} & 90 & \text{gwi:-gu:} & / < \text{gu + i}/ \\
\end{align*}
\]

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:


<table>
<thead>
<tr>
<th></th>
<th>tc(53)</th>
<th>tc(55)</th>
<th>tc(53)</th>
<th>tc(53)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ten</td>
<td>tc(53)</td>
<td>tc(53)</td>
<td>tc(53)</td>
<td>tc(53)</td>
</tr>
<tr>
<td>twenty</td>
<td>kh(55)l(55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>forty</td>
<td>che(53)-n(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sixty</td>
<td>che(53)-sum(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>eighty</td>
<td>che(53)-pli(53)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^{100}\)-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: tc\(53\)the\(53\)-11’, tc\(55\)n\(53\)i\(35\)-12’, tc\(53\)sum\(53\)-13’, tc\(53\)pli\(53\)-14’.
In this dialect the word for ‘20’, khA\textsubscript{55}I\textsubscript{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, che\textsuperscript{53}, though this may actually represent a co-allofam of the same etymon, perhaps *kyaI (< *kal + i ?). (In the other dialect treated by Lu Shaozun, Wenlang (Northern Cuona), the simple form for ‘20’ is also kha\textsubscript{55}I\textsubscript{55}, but the variant that occurs with 40–100 is khaI\textsubscript{55}.) The odd round numbers (30–90) are additive formations based on the next lower multiple of 20.

[B] TAMANG AND SHARCHOP/TSANGLA

<table>
<thead>
<tr>
<th>Tamang</th>
<th>Sharchop</th>
</tr>
</thead>
<tbody>
<tr>
<td>TEN</td>
<td>ci</td>
</tr>
<tr>
<td>TWENTY</td>
<td>khaI-ki:h</td>
</tr>
<tr>
<td>FORTY</td>
<td>khaI-nyi:h</td>
</tr>
<tr>
<td>SIXTY</td>
<td>haI-som</td>
</tr>
<tr>
<td>EIGHTY</td>
<td>khaI-pli</td>
</tr>
<tr>
<td>HUNDRED</td>
<td>khaI-nga:h</td>
</tr>
<tr>
<td>THIRTY</td>
<td>khaI-ki:h-syi-ci</td>
</tr>
<tr>
<td>FIFTY</td>
<td>khaI-nyi:h-syi-ci</td>
</tr>
<tr>
<td>SEVENTY</td>
<td>khaI-som-syi-ci</td>
</tr>
<tr>
<td>NINETY</td>
<td>khaI-pli:syi-ci</td>
</tr>
</tbody>
</table>

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 khaI:pacs (‘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

<table>
<thead>
<tr>
<th>Jirel</th>
<th>Kaike</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>cyuta:m:baq</td>
</tr>
<tr>
<td>20</td>
<td>(nye: syu)</td>
</tr>
<tr>
<td>40</td>
<td>khalq-nyiq</td>
</tr>
<tr>
<td>60</td>
<td>khalq-sumq</td>
</tr>
<tr>
<td>80</td>
<td>khalq-syi</td>
</tr>
<tr>
<td>100</td>
<td>(sei-jyiq)</td>
</tr>
</tbody>
</table>
Jirel and Kaise 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 Kaise uses nhi-chyu throughout the twenties (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’.

Kaise 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...’), Kaise 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 Kaise, 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’.

[KANAWARI

Kanawari is a rather well-behaved vigesimal language, though like Kaise 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 pii-nija ‘80’.

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

30  dje'-nija
50  dai-nija
70  sa'e-shum-nija
90  sa'e-pü-nija

According to Joshi/Rose, the dje- in ‘50’ is a loan from Hindi meaning ‘two and a half’: 50 = 2 1/2 x 20! THIRTY seems to be additive, with dje- ‘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’, cahr-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 cahr 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)hazel (tu-hazel '20', sah-khazel '40', suk-khazel '60', hazel-khazel '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: pacas '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):

<table>
<thead>
<tr>
<th>Higher Number</th>
<th>Vigesimal</th>
<th>Decimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWENTY</td>
<td>k’a; k’a-kat</td>
<td>20 (x 1)’</td>
</tr>
<tr>
<td>FORTY</td>
<td>k’a-nyət</td>
<td>20 x 2’</td>
</tr>
<tr>
<td>SIXTY</td>
<td>k’a-sam</td>
<td>20 x 3’</td>
</tr>
<tr>
<td>EIGHTY</td>
<td>k’a-fəli</td>
<td>20 x 4’</td>
</tr>
<tr>
<td>HUNDRED</td>
<td>k’a-fango</td>
<td>20 x 5’</td>
</tr>
</tbody>
</table>

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’:

<table>
<thead>
<tr>
<th>Higher Number</th>
<th>Vigesimal</th>
<th>Decimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>THIRTY</td>
<td>k’a-kat sə kəti</td>
<td>(20 x 1) + 10’</td>
</tr>
<tr>
<td>FIFTY</td>
<td>k’a-nyət sə kəti</td>
<td>(20 x 2) + 10’</td>
</tr>
<tr>
<td>SEVENTY</td>
<td>k’a-sam sə kəti</td>
<td>(20 x 3) + 10’</td>
</tr>
<tr>
<td>NINETY</td>
<td>k’a-fəli sə kəti</td>
<td>(20 x 4) + 10’</td>
</tr>
</tbody>
</table>

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-fango ‘100’.
We should note that although the independent numeral *kati* ‘10’ has the *ka-* prefix, and though this has been carried over into the decimal higher round numbers, *ka-* may originally have had no connection with TEN at all. The *ka-* in *kati* 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 *khal-jik-tang-curkhu* ‘39’ (‘20 + 19’), implying that one counts in an unbroken string from 20 to 40, so that ‘30’ must be *khal-jik-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 TEN:

<table>
<thead>
<tr>
<th>FORTY</th>
<th><em>khal-ji</em></th>
<th>SEVENTY</th>
<th><em>khal-din</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(ji ‘4’)</td>
<td></td>
<td>(din ‘7’)</td>
</tr>
<tr>
<td>FIFTY</td>
<td><em>khal-ngaq</em></td>
<td>EIGHTY</td>
<td><em>khal-ge</em></td>
</tr>
<tr>
<td></td>
<td>(ngaq ‘5’)</td>
<td></td>
<td>(ge ‘8’)</td>
</tr>
<tr>
<td>SIXTY</td>
<td><em>khal-Tuk</em></td>
<td>NINETY</td>
<td><em>khal-gu</em></td>
</tr>
<tr>
<td></td>
<td>(Tuk ‘6’)</td>
<td></td>
<td>(gu ‘9’)</td>
</tr>
</tbody>
</table>

This obviously secondary ‘decimalised’ system continues through and beyond HUNDRED: *khal-citambaq* ‘100’ (‘10 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’). 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 *bçu* ‘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’ \((khe < *m-kul)\), but it also has lexemes for the next three powers of twenty:

<table>
<thead>
<tr>
<th>Power</th>
<th>Lexeme</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>(khe)</td>
<td>20</td>
</tr>
<tr>
<td>20(^2)</td>
<td>(jicu^{105})</td>
<td>400</td>
</tr>
<tr>
<td>20(^3)</td>
<td>(kheche^{106})</td>
<td>8,000</td>
</tr>
<tr>
<td>20(^4)</td>
<td>(jâche)</td>
<td>160,000</td>
</tr>
</tbody>
</table>

Even so, the language has a normal decimal system of teen-formation (\(\text{TEN} + \text{UNIT}\)). 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.

<table>
<thead>
<tr>
<th>Decimal</th>
<th>Vigesimal</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 (ji-i-cu) (\text{‘2} \times 10)</td>
<td>(khe-ci) (\text{‘20} \times 1)</td>
</tr>
<tr>
<td>21 (tsa-ci) (\text{‘20} + 1)</td>
<td>(khe-ci) ((\text{da}) ci) (\text{‘(20} \times 1) + 1)</td>
</tr>
<tr>
<td>22 (tsa-ni)</td>
<td>(khe-ci) ((\text{da}) ni)</td>
</tr>
<tr>
<td>30 (sum-cu) (\text{‘3} \times 10)</td>
<td>(khe pje da \text{ji}‘20 \times (-1/2 + 2))</td>
</tr>
<tr>
<td>35 (so-ni) (\text{‘thir-} + 1)</td>
<td>(khe \text{ko da } \text{ji}‘20 \times (-1/4 + 2))</td>
</tr>
<tr>
<td>40 (zi-p-cu) (\text{‘4} \times 10)</td>
<td>(khe-pi‘20 \times 2)</td>
</tr>
<tr>
<td>50 (\etaa-p-cu)</td>
<td>(khe pje da sum‘20 \times (-1/2 + 3))</td>
</tr>
<tr>
<td>55 (\etaa-ni)</td>
<td>(khe ko da sum‘20 \times (-1/4 + 3))</td>
</tr>
<tr>
<td>60 (\text{d(h)uk-cu})</td>
<td>(khe-sum‘20 \times 3)</td>
</tr>
<tr>
<td>70 (\text{dyn-cu})</td>
<td>(khe pje da zi‘20 \times (-1/2 + 4))</td>
</tr>
<tr>
<td>80 (\text{ge-p-cu})</td>
<td>(khe-zi‘20 \times 4)</td>
</tr>
<tr>
<td>90 (\text{gu-p-cu})</td>
<td>(khe pje da n(a)‘20 \times (-1/2 + 5))</td>
</tr>
<tr>
<td>100 (cik-ja‘1 \times 100)</td>
<td>(khe-\text{(n)a‘20 \times 5})</td>
</tr>
<tr>
<td>400 (\text{zip-ja‘4} \times 100)</td>
<td>(jicu-ci‘400 \times 1)</td>
</tr>
<tr>
<td>500 (\etaap-ja‘5 \times 100)</td>
<td>(jicu-ci \text{da } khe-\text{(n)a‘(400} \times 1) + (20 \times 5))</td>
</tr>
<tr>
<td>600 (\text{d(h)uk-ja‘6 \times 100})</td>
<td>(jicu-pje da ji‘400 \times 1 \text{ 1/2})</td>
</tr>
</tbody>
</table>

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 \times 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 \textit{une grosse} (> English \textit{gros}) ‘12 dozen; 144’ derives from \textit{une grosse douzaine} ‘a big dozen’). The etymology of the first syllable of \(j\text{\(i\)iche}\) remains obscure.\n
\(^{107}\) Michailovsky (pers. comm. 1995) suggests a connection with Tibetan \(yag\) ‘again’, that is ‘even bigger’.

\(^{108}\) The morpheme \textit{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 \textit{rtsa} used to connect the tens to the units, e.g. \textit{nyi-shu-rtsa-gcig} ‘21’ (‘2 \times 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 \textit{da} (< WT \textit{dag}) then took over connective function in the Dzongkha vigesimal system (e.g. \(khe-ci \text{ da } ci\)).

\(^{109}\) 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).

\(^{110}\) One can also say \(ja-thampa\) (‘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-daṅ gnyis (phyed ‘half’, daṅ ‘with’, gnyis ‘2’), that is ‘which with an additional one-half, would be two’), the odd round numbers are formed subtractivly in Dzongkha: for example, ‘30’ khe phe daṅ, that is ‘twenty times one-half-less-than-two’, or ‘twenty times one-and-a-half’.

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: yatt-jo?, nis-jo?, sum-jo?, play-jo?, ponga-jo?.111 SIX through TEN are expressed by Nepali numerals: cha-gota:, sat-gota:, ?a: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: yatt-ha:l. The first syllable means ONE (§3.1.1), and the second element is clearly the inherited TB word for TWENTY (< *kal x *m-kul; cf. Khaling (k)hael). 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:

\[
\begin{align*}
yatt-ha:l & \text{ ?a:t-gota:} & 20 & ('(1 \times 12) + 8') \\
yatt-ha:l & \text{ das-gota:} & 22 & ('(1 \times 12) + 10') \\
nis-ha:l & & 24 & ('2 \times 12') \\
nis-ha:l & \text{ ponga-jo?} & 29 & ('(2 \times 12) + 5') \\
nis-ha:l & \text{ sat-gota:} & 31 & ('(2 \times 12) + 7') \\
sum-ha:l & \text{ play-jo?} & 40 & ('(3 \times 12) + 4') \\
play-ha:l & \text{ nis-jo?} & 50 & ('(4 \times 12) + 2') \\
ponga-ha:l & & 60 & ('5 \times 12')
\end{align*}
\]

---

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 cetvertogo ‘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.)

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 transvaluated 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’)

Kuki-Naga

<table>
<thead>
<tr>
<th></th>
<th>HUNDRED</th>
<th>TWENTY</th>
<th>FIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chang</td>
<td>sau-ngau</td>
<td>sau</td>
<td>ngau</td>
</tr>
<tr>
<td>Wancho14</td>
<td>pu-ga</td>
<td>pu-</td>
<td>ga</td>
</tr>
<tr>
<td>Yacham-Tengsa</td>
<td>mesung-phung</td>
<td>mesung-</td>
<td>phung</td>
</tr>
</tbody>
</table>

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).
3.5.4.3 WHERE HUNDRED HAS A SPECIAL RELATIONSHIP WITH FIFTY

In Miju Mishmi (Das Gupta 1977a), ‘100’ is waie-mo\textsuperscript{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\textsuperscript{1} or *?ra\textsuperscript{1}

Qiangic
Ersu ta\textsuperscript{55} za\textsuperscript{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)\textsuperscript{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

\textsuperscript{115} Sun Hongkai et al. (1980) transcribe Deng Geman (Kaman) ‘100’ as wo\textsuperscript{55} je\textsuperscript{53} mu\textsuperscript{53}.

\textsuperscript{116} The apparent homophony of these two syllables reminds one of the Lahu number hi hi ‘8,000’, where even the tones are identical. This is pure accident, however: the first syllable is from PLB *\textsuperscript{3}rit ‘eight’, but the second is a loan from Shan hi\textsuperscript{h} ‘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-gya117

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-r-ya with labial prefix
Mikir paro, Nruangmei 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 allopham also belong laryngeal-initialled forms like Zeliang-Zeme-Mzieme and Maram hai.

*m-r-ya with nasal prefix
Lotha nzoa; Maring macha; Meluri and Ntenyi meza; Pochury mza-ke

*r-g-ya with liquid prefix
Jingpho latsal118; Puiron raja; Kom Rem raza; Phom gho (presumably 'gh' is a voiced velar fricative < *r); Kokborok racha; Garo ritcha-sa; Dimasa raja-si

*γ-ya or *-r-ya (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 tshi; Yimchungru chhi
Lushai, Lakher, Hmar, Gangte, Paite, Tiddim za, Vaiphei ja
Boro jou ; Dulung cya
Pumi sha55; Qiang tshi55
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. matsät 'eight'.
3.5.4.6 A SPECIAL ABOR-MIRI-DAFLA ROOT FOR HUNDRED AND ITS RELATIONSHIP TO TEN

\[ *m-li(\eta) (? < *m-l-\text{gya}) \]

Abor-Miri \( li-ko \) – \( ling-ko \)
Minyong \( ling-ko \)
Dafla \( leng-go \)
Nishi \( lunkh \)
Apatani \( lange, laq-e \)
Monpa Cuona \( c'\text{e}^{55} \text{le}^{21} \text{nge}^{53} \)
Lhopa \( lung \)
Darang Deng \( malum^{55} \)
Chulikata \( malu: \)
Mishmi \( malo^{119} \)
Gallong \( hamyi (< *s-m\text{li}?) \)

It will be observed that these forms bear more than a slight resemblance to a set \( *\text{rig} \approx *\text{yi}\)g ‘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 \( uujung \) ‘10’ and \( lung \) ‘100’),\(^{120}\) though they may have ‘contaminated’ each other.\(^{121}\) It is possible that these forms for HUNDRED are ultimately to be derived from \( *m-\text{gya} \) (ult. \(< *b-\text{rgya} \)) via apocope of the root vowel.

3.5.4.7 THOUSAND

(a) \( *s-tog \)

STC #32 (pp.21, 94) sets up the root \( *s-tog \) on the basis of forms from two languages, WT \( stog \) and WB \( thog \). 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 \( stt^{55} \)
Qiang \( xto^{55} \)
Dulung \( tu^{55} \)
Karen (Pwo) \( thon, \) (Sgaw) \( k\text{a}tho \)

\(^{119}\) 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 \( *\text{lu}u\)j ‘hundred’, distinct from PTani \( *\text{rj}\text{u}u\)j ‘ten’ (p.144), citing forms like Bengni and Bokar \( luu\), Bangru \( lau\)j\(^{53}\), Dhammai \( buu\text{-lu}j\), 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 \( \text{nyi-gya} ‘200’ \) has been borrowed as Kanauri \( ni-\text{ja} ‘20’ \).
(b) *s-rēŋ ≠ *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 riijing-si (se ‘1’); Tangkhul thing-kha (kha ‘1’); Jingpho ching; Karen (Taungthu) tāʔ-rèng (tāʔ ‘1’), (Palaychi) hreng

(2) with the rhyme -aŋ

Taraon rejaŋ; 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 pi). 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’

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nocte</td>
<td>cha-ichi</td>
<td>(cha ‘100’, ichi ‘10’)</td>
</tr>
<tr>
<td>Ntenyi</td>
<td>meza-ta’a</td>
<td>(meza ‘100’, ta’a ‘10’)</td>
</tr>
<tr>
<td>Dafla</td>
<td>leng-rengcheng</td>
<td>(leng ‘100’, rengcheng ‘10’)</td>
</tr>
<tr>
<td>Apatani</td>
<td>lā-lyā</td>
<td>(lā ‘100’, lya ‘10’ )</td>
</tr>
<tr>
<td>Abor-Miri</td>
<td>li-ying-ko</td>
<td>(li – ling ‘100’, eying ‘10’)</td>
</tr>
<tr>
<td>Adi(^{123})</td>
<td>ling-ko-ijing</td>
<td>(ling-ko ‘100’, iying ‘10’)</td>
</tr>
</tbody>
</table>

\(^{122}\) Note that the Germanic word for THOUSAND is historically a compound meaning ‘swollen hundred’ (PGmc *bush-hundi, Old Norse bushundrad).

\(^{123}\) 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  kho:l pacas  (kho:l ‘20’, pacas ‘50’ (< Nepali))

(d) ‘20 x 10 x 5’

Chang  sau-an ngaun  (sau ‘20’, an ‘10’, ngaun ‘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’!

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 xilo- < *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 hanzār (apparently with rhinoglottophilia in the first syllable; see Matisoff 1975.)

Konyak, Phom hāja

Sangtam hajar

---

124 Could puban be a typo in Marrison (1967:267) for puga-ban (‘100 x 10’)?
CHAPTER 4

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 lyə ‘10’ (of non-humans) – alyə ‘10’ (of humans) (§3.2.3.3); Dzongkha tsa ‘20’ (for counting objects from 21 to 29, for example, tsa-pi ‘22’) versus per ‘20’ (for dates, for example, per-pi ‘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 masùm ‘3’ (< *g-sum) ←-- mali ‘4’ (< *b-lay).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 shìt ‘8’ --＞ shì ‘9’; Serdukpen khu ‘5’ --＞ khit ‘6’; Khaling sa:hp ‘2’ --＞ suhp ‘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-ŋə), 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-ŋə; 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 tsənī ‘7’ --> tsəsī ‘8’ (Gvozdanović 1985:143; here both the prefix and rhyme of ‘8’ have been affected); Serdukpen khıt ‘6’ (originally with final velar, *d-k-ruk) <-- sit ‘7’.

(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 (lit ‘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 šēp ‘3’, with non-etymological high-stopped tone (written with -ʰ?), doubtless developed the glottal stop as a demarcational feature to set it off from the next higher numeral š ‘4’, that begins with a vowel.

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.

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

In Lepcha an etymological flipflop between 'EIGHT' and 'NINE' seems to have occurred: kākū '8' (but cf. PTB *d-kaw 'nine', §4.2.4), kākyōt '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-pog, Yakkhaba ip-pog '10') (§3.2.3.6; Gvozdanović 1985:162).

The Moli dialect of Bahing has etymologically correct forms for '6' (rukhu) and '7' (cũn), but in the Bigutar dialect the cognate forms ruka and cani mean '5' and '6', respectively (Gvozdanović 1985:135).

In Kulung the word tupci, etymologically '6', has come to mean '5' in the Pawoi dialect, while the word retcí, 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, skpu and bhbkpu, 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 rökpu has come to mean '4' (half of eight), just as in the Ranitar dialect of Bantawa, where the cognate retkapok 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-kaw '9' (Gvozdanović 1985:143). Similarly, Sunwar yaan '9' looks like it derives from *g-ryan '8' (§4.2.3.6).

---

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 lexicemically 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) hhac. Still others seem to derive from open syllables, for example, Lahu nř < 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 suffixed at the PST level, and the etymon is so reconstructed in the Indexes: ‘*\( g\)-nis = *\( g\)-ni-s’.

For Proto Lolo-Burmese (PLB) I have reconstructed *(?)ni-t (TSR #160), but there is no hard evidence to enable us to distinguish between *\(-t\) and *\(-k\) here, and perhaps *(?)ni-? would be a better reflection of this indeterminacy.

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

- Lepcha: \( nyi, ny\)ät
- 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 \( a\)hrat; Ao teret, terat; Yacham-Tengsa telet (STC #6; GEM 34).

The unmistakeable 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). 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-liːf ‘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 lakhõg, which has never been successfully related to anything else. (Curiously, Jingpho also has an isolated form laŋkāi ‘ONE’, with the same prefix and under the same (rare and secondary) falling tone.)

---

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 gāi, see Matisoff (1994) ‘Watch out for number one’, LTBA 17.1.
4.1.1.1 Forms with velar prefix

Angami (Khonoma) kena, (Kohima) keni; Chakhesang and Chokri küna; Kezhama kenhi; Khoi rao kati (with denasalisation of root-initial, maybe under the influence of kathum ‘3’); Mao kahei; Maring khani; Meluri keni; Mzieme kena; Nruanghmei kanei, küei; Ntenyi kenyi; Pochury künü; Puiron kani; Rengma kohüng; 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 kani; Deng Darang ka2n55 (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 kurn Jin53 (’j’ is the palatal semivowel); Digaru Mishmi ka-ying; Idu kanyi; Miju kinin; Mishmi (Dubey) kani; Taron kaying; Aka (= Hruso) kshi

These Abor-Miri-Dafla forms require some comment. Several languages (Deng Geman, Digaru Mishmi, Miju and Taron) 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 Taron -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 a2 ni55; Rawang aʌni

Apatani ni; Abor-Miri, Gallong, Lhopa, Minyong, Tagin a-nyi; Padam a-ni; Nishi anni; Dafla (Das Gupta) anyi ~ aɪn; Milang ne; Monpa (Cuona) nA23

Boro nɔi; Kokborok nuy

Chepang nis-jo?, Gurung ngihq; Jirel nyiq; Kaise nghi; Kanawari nish; Kham neh-plo; Magari nis; Newari ni-gul-li; Sherpa ngyi; Sikkim Bhutia nyi; Tamang nyi:h; Thakali ngih; Thulung Rai nə; Tibetan (Lhasa) nyi

Ersu ne55; Bai (= Minchia) ne2; Pumi n23; Qiang nyi55

11 Something very similar happens in Lahu, where /yl/ acquires local friction before the high front vowels /i,el/, becoming a voiced slit spirant. See Matisoff (1973a:5-6).
Ao (Chungli) anā; Kimsing anāi; Chang nyī; Konyak i (with palatalisation of the initial); Liangmai nīa; Lotha enī, onī; Meithei anī; Phom nyī; Sangtam anyū; Tangsa (Moshang) anī, (Yogli) anēi; Wanchan an(y)i

Gangte nīh; Kom Rem nī; Kuki nī; Lakher no; Paite nīh; Thado nī; Tiddim nīh

We must include here Jingpho nī, a bound form which is never used in isolation, but only in certain set expressions like nī nā?r two nights and round numbers like nī tsā ‘200’. It has also been grammaticalised into a plural or collective suffix, for example gwi nī ‘the dogs’. The independent Jingpho numeral for TWO is the mysterious lekhoa (§4.1.1.5).

4.1.1.3 FORMS WITH OTHER THAN VELAR PREFIXES

Yimchungnu 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-nyī, 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 hnae, which could descend either from PLB *-ik or *-it (cf. WB hrac ‘8’ < PLB *?rit or *?ryat, but also WB chac ‘joint’ < PLB *tsik).

Forms from Abor-Miri-Dafla and Himalayish also attest to the antiquity of a velar suffix with this numeral:13

Hayu nak-pu (for humans), na?ung (for non-humans) (Michailovsky 1988:167); Sunwar nik-syī; Bahing nik-sī; Sharchop (Chhewang Rinzin) nyik-tsing; Monpa (Motuo) nyik-tsing; Serdukpen n(y)ik; Thulung nāk

The second syllable of the Sunwar and Bahing forms certainly seem related to the -tsing in Sharchop and Monpa, though their wider affiliations are still unknown.

4.1.1.5 UNUSUAL FORMS

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

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jingpho</td>
<td>lakhông</td>
<td>(no known etymology)</td>
</tr>
<tr>
<td>Khaling</td>
<td>sa:h-pu</td>
<td>(apparently contaminated by suh-pu ‘3’)</td>
</tr>
<tr>
<td>Bumthang</td>
<td>zon</td>
<td></td>
</tr>
</tbody>
</table>

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-I*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-* variation in this etymon, a pattern already noticed in a number of another etymologies (for example, ‘fragrance’ PTB *b-suq *b-saŋ (STC #405)).15

4.1.2.1 FORMS WITH VELAR PREFIX

Written Tibetan gsum

Chulikata ka⁷ʃ (with preemption via apocope of the rhyme); Deng Darang kə²²ʌmg⁴⁵; Deng Geman ku²²am⁵³; Idu (Sun 1983) ka³¹song⁵⁵; Miju Mishmi ksam; Taraon ka₃saŋ

Kezhama katsi; Khoirao kathum; Mao kosii; Maring khiyum; Meluri keche; Mikir kethom; Mzieme ketsam; Nruanghmei kathum; Nnteniy keching, keshang; Pochury kiche; Rengma keshan; Sema këthu; Tangkhul kathum; Zelhang 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-rig *s-ray ‘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-\textit{qum}; Aka zu;
Apatani \textit{hi}; Dafla (=Nishi) (a-)\textit{om}, \textit{um}; Bokar a-\textit{hum}; Lhopa \textit{afium}\textsuperscript{16}; Milang ham;
Monpa (Motuo) \textit{sam}, (Dubey and Cuona) \textit{sum}; Serdukpen \textit{ung}

Angami \textit{se}; Chakhesang \textit{s\text{\textbar}}; Chang \textit{sam}; Chokri \textit{s\textbar}; Liangmai \textit{shum}; Phom \textit{jam};
Wancho a-\textit{jam}, a-\textit{zam}; Ao (Chungli) asem, (Mongsen) \textit{asam}; Kimsing \textit{acam}; Lotha \textit{etham}; Meithei \textit{ahum}; Sangtam \textit{asang}; Tangsa (Moshang) \textit{atum}, (Yogli) \textit{adim}; Yacham-Tengsa and Yimchungru \textit{asam}; Nocte \textit{van-ram}, \textit{wan-ram}; Konyak \textit{lem};

Lakher \textit{th\text{\textbar}}; Gangte, Kom Rem, Kuki, Thado, Paite, Puiron, and Tiddim \textit{thum}

Rawang (Barnard 1934) \textit{atsum}; Dulong \textit{a}\textsuperscript{21} \textit{sum}\textsuperscript{53}

Boro and Kokborok \textit{tham}

Ersu \textit{si}\textsuperscript{55}; Pumi \textit{s\text{\textbar}}\textsuperscript{23}; Qiang \textit{tshi}\textsuperscript{55}; Bai (= Minchia) \textit{sa}\

Chepang \textit{sum-j\text{\textbar}}; Gurung \textit{soq}; Hayu \textit{tshuk-pu} (for humans), \textit{tshu}\textsuperscript{2}-\textit{ung} (for non-humans); Jirel \textit{sumq}; Kaise \textit{sum}; Kanawari \textit{shum}; Khaling \textit{shuh-pu}; Kham \textit{sohm-lo};

Tibetan (Lhasa) \textit{sum}

Karen (Pa-O) \textit{som}, (Pwo) \textit{c\text{\textbar}n}, (Palaychi) \textit{tyq}, (Sgaw) \textit{c\text{\textbar}n}

Proto Lolo-Burmese *\textit{sum} (Tone *2) > WB \textit{s\textbar}m, Lahu \textit{\textbar\textbar} – \textit{s\textbar}, etc.

The glottal stop in Lahu is secondary, a junctural feature that arose to separate THREE from FOUR in counting, since Lahu 5 ‘4’ has a vocalic onset.\textsuperscript{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, \textit{pa}-, used with all the primary numerals: Hmar, Lushai and Vaiphei \textit{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 \textit{hang-tum}; Nocte \textit{van-ram}.

Most strikingly, Jingpho has developed a nasal prefix with this numeral – \textit{m\textbar\textbar}{}\textit{sum}. This is part of a celebrated Jingpho run of the labial nasal prefix from THREE to \textit{FIVE} (\textit{mas\textbar\textbar}{}\textit{m}, \textit{m\textbar\textbar}{}\textit{l}, \textit{m\textbar\textbar}{}\textit{a\textbar}) Clearly it is not THREE that is influencing the two higher numerals, but vice versa, since both FOUR and \textit{FIVE} have been associated with labial prefixes from PTB times. See §4.1.3; §4.1.4.

\textsuperscript{16} The voicedness of the \textit{hi} 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 \textit{a}- prefix (see §5.5.1.2) might have caused the voicing. J.T. Sun (1993:110) sets up Proto Tani *\textit{fium} for this root.

\textsuperscript{17} See §4.0.1(c).
4.1.3 PROFILE OF NUMBER FOUR

FOUR *b-lyi or *b-lay

The following forms are cited in STC #410:\n
Written Tibetan bzi; Thulung bli; Kanauri po, Magari buli; Digaro kapre; Miri pi; Nung abyi, abali; Jingpho mali; Written Burmese le; Maru bit (with secondary final stop); Mikir phi.

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, puu-lje (also pe, with preemption); Monpa (Cuona) pli;\(^\text{23}\) (Dubey) blee, (Central) b(i)ci ~ p(i)ci; Serdukpen bisi\(^\text{20}\)

Ao (Chungli) pezu; (Mongsen) phiili; Kezhama pedi; Kimsing balai; Konyak peli\(^\text{21}\); Lakher pali; Mao padei; Maring phiili; Mikir phi (also phir (Marrison), with metathesis and/or apocope); Muklom Tangsa balee; Nocte beli; Nruangmei padei; Rengma pezi; Sema bidhi; Tangsa (Moshang) bali; (Yogli) bala; Yacham-Tengsa phale; Yimchunguru phiyi

Dulung a\(^2\) bli\(^3\) (with secondary vocalic prefix)

Boro broi; Dimasa biri; Garo bni; Kokborok bruy

Chepang playjo?; Hayu b(l)i'lung; Gurung plihq; Tamang and Thakali plih; Thulung Rai bla; Sharchop pshi; Dumi balikpi.

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) pi; Nishi and Padam appi; Tagin ep\(^\text{22}\)

Khaling bhael (alongside bhoxm ‘5’, also with preemption); Kanawari pu; Newari pe-gu:

A rare ‘survival via preemption’ of the *b- prefix in Burmish is Maru bit. (The development *-iy (= *ay) > Maru -it is regular, as is the parallel *-uw (= *aw) > Maru -uk.)\(^\text{23}\)

\(^{18}\) 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.


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

\(^{22}\) 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-lay < *b-lay

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 *məlī; Khoirao *malhī; Kom Rem *manli; Liangmai *madai; Lotha *mezū; Maram *madai; Meithei *marī; Meluri *mezū; Mzieme *m(ə)dai; Ntenyi *mez(h)ū; Pochury *mezū; Puiron *məlī; Sangtam *mūzyū; Tangkhul *matī; Zeliang *mdai; Zeme *medai

4.1.3.3 FORMS WITH LABIAL FRICATIVE PREFIX:

Aka (= Hruso) *fi-rī; Lepcha *fəli

4.1.3.4 FORMS THAT DO NOT OVERTLY REFLECT ANY CONSONANTAL PREFIX

Karen: Taungthu *liit (with suffix; §4.0.1.(d)); Pwo *li, *liʔ; Palaychi and Sgaw *lwi
Himalayish: Kaire *li, Bahing *le, Kulung *li-chi, Limbu *li-si
Chin: Gangte, Kuki, Paite, Thado, Tiddim *li
N.Naga: Chang *leî; Phom *a-li; Wancho *a-li

4.1.3.4.1 FORMS THAT INDIRECTLY REFLECT A CONSONANTAL PREFIX

PLB *hlay² > WB *le, Lisu *li şek; YI (Dafang) *ti³; Naxi (Lijiang) *lu³; MPI *li⁶; Lahu *x(n)²⁵; Akha *x, Bisu *ha; Phunoi *hàn (with rhinoglottophilia), and so on.

4.1.3.4.2 WITH REPLACEMENT OF THE ROOT-INITIAL LATERAL BY A STOP²⁶

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) *z̤ *v³³

Jirel *syī; Sheri *ji; Sikkim Bhutia/Dzongkha *zyi – *syī

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 linguist’s dilemma’.
Ersu езжа, Pumi (Taoba) езжа, (Jinghua) 骓, Qiang (Taoping) 塅, (Mawo) צב; Ergong 碨; Muya gebra; Queyu (Yajiang) 鬒, Guiqiong smarty, Namuyi  […], Shixing  […]
Bai (Dell 1981) stw², (Dali, Jianchuan) cî⁴⁴, (Bijiang) si⁴⁴

4.1.3.6 WITH VELAR PREFIX:

*g-lay* >

Minyong aki (with preemption)

*g-b-lay>*

Rgyalrong(Zida) kewdzyi (with lenition of the *-b- to -w-)

Idu Mishmi kapri

Taraon kapra:i

Deng Daïang kə⁴⁴ pəər⁵⁵

Chulikata ka:ppi

Digaro kaprei

With this last allofam also belong Miju Mishmi kambran, Deng Geman kəu²bəu²⁵³, evidently with a fully syllabicised prefix and secondary nasalisation in both syllables: *g(N)-b-lay- (N).²⁷

4.1.4 PROFILE OF NUMBER FIVE

FIVE *l-ŋa ≠ *b-ŋa

In STC #78²⁸ the following forms are cited in support of the reconstruction *l-ŋa ≠ *b-ŋa:

WT ŋa; Jingpho mga; WB ŋa; Garo boŋa; Lushai ŋa ~ pəŋa²⁹.

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. mga < *b-ŋa). Yet an *m- prefix is independently required for PTB beyond any question, both with nominal and verbal root. 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 gum- that Jingpho developed in its word for ‘horse’ gumræ(g) < *m-ray.
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-qa (for example, Rangkhol ringa) (STC, ibid.)
Kom Rem ranga
Monpa (Cuona) le2inge54, (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:-jo?
Milang pangu
Dulung pu2nga53; 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) pengu, (Mongsen) phanga; Kezhama pangu; Kimsing bangi; Lakher pangaw; Mao pongo; Maring phanga; Mikir phongo; Muklom Tangsa and Nocte banga; Nruanghmei pangu; Sema pongo; 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; Chakesang puhung, Chokri püngu

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

*b-(γ)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’, dsk.
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) ma\(^3\)nga\(^{15}\); Mishmi (Dubey) manga; Taraon ma:\(\text{ng}\); Deng Darang ma\(^2\)nga\(^{45}\); Chulikata manga; 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; Meitheli, Meluri manga; Mzieme mengei; Ntenyi m\(\text{ng}\); Pochury m\(\text{ng}\); Sangtam m\(\text{ng}\); Zeliang mengei; Zeme mengeu

4.1.4.2.1 WITH PREEMPTION OF THE ROOT-INITIAL BY m-

Bai (Dali) mu\(^3\)

4.1.4.3 FORMS WITH LABIAL SPIRANT PREFIX:

Lepcha f\(\text{ng}\)o (alongside f\(\text{li} \) ‘4’)

4.1.4.4 DOUBLY PREFIXED FORMS:

*b-l\(\text{ng}\)a > Padam/Abor pil\(\text{ng}\)o (see LSI III.1, p.622); Shimong Adi pi-ri-\(\text{ng}\)

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

Dzongkha/Sikkim Bhutia n\(\text{ga}\); Gurung nga:ha; Jirel nga:q; Kaire nga:; Kanawari nga; Sharchop nga; Newari nya:-gu; Sherpa, Tamang and Thakali nga:q; Thulung Rai ngo; Dumi \(\text{ng}\); Kulung \(\text{ng}\)-chi; Limbu \(\text{ng}\)-si

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

Gangte, Kuki, Paiete, Tiddim nga

Monpa (Motuo) nga; Apatani nga; in this category we may include other AMD forms with vocalic prefix, for example, Abor-Miri, Dafla, Gallong, Minyong, Tagin a\(\text{ng}\); Nishi a\(\text{ng}\); Lhopa ongo; Bangni u-\(\text{ng}\)

Ersu nguar\(^33\); Muya \(\text{ng}\)\(^1\); Queyu nga-te\(^3\); Guiqiong \(\text{ng}\)\(^3\); Namuyi \(\text{ng}\)\(^3\)

Proto Lolo-Burmese *\(\text{ng}\)\(^2\) > WB \(\text{ng}\); Lahu \(\text{ng}\); Lisu \(\text{ng}\); MPI \(\text{ng}\)\(^2\); BISU \(\text{ng}\) ~ \(\text{h}\); Phunoi ?\(\text{an}\) (with rhinoglottophilia), and so on.
Bai (Jianchuan, Bijiang) ŋv
Karen (Pa-O) ngat (with suffix); other Karen dialects have lenited the nasal to a palatal semivowel: Pwo je, jai?, Palaychi and Sgaw je

4.1.4.6 WITH SPIRANTISED INITIALS:
In several Qiangic languages the velar nasal has become a velar or ‘uvular’ fricative:
Pumi yuā; Qiang súa33, Shixing ŋā

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-g(a):
Kaman kw21-len55; 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:
Bantawa (Ranițar dialect): chuk ‘HAND’, ükchuk ‘FIVE’ (‘1 × 5’), hūachuk ‘TEN’ (‘2 × 5’)
Mewahang: huk ‘HAND’, ihuk ‘FIVE’
Chourase: brem ‘FINGER’, kollabremci ‘FIVE’ (kolo ‘one’)
Yakkha: muktā ‘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 soixantedix-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’ (zokkany-nay), with the special morpheme zokkay ‘group of four’ (rather than with the unrelated cardinal numeral bra 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, retkapok 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 (=Karen The Karen) the numerals SIX and EIGHT are expressed as doubles of THREE and FOUR (sō ‘three’, sō swā ‘six’; lwī ‘four’, lwīswā ‘eight’), while SEVEN and NINE are in turn additively formed from SIX and EIGHT (sō swā to- ‘seven’ ‘(3 x 2) + 1’; lwīswā to- ‘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 u-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
79

from 20’). 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

<table>
<thead>
<tr>
<th>Number</th>
<th>Sanskrit</th>
</tr>
</thead>
<tbody>
<tr>
<td>17</td>
<td>saptadaśa</td>
</tr>
<tr>
<td>18</td>
<td>aśtadaśa</td>
</tr>
<tr>
<td>19</td>
<td>navadaśa (‘9 + 10’) or ūnaviṃśati &lt; ekaviṃśati (‘one diminished 20’ (pers.comm. Robert P. Goldman 1994))</td>
</tr>
<tr>
<td>20</td>
<td>viṃśati</td>
</tr>
<tr>
<td>40</td>
<td>catvāriṃśat</td>
</tr>
<tr>
<td>49</td>
<td>navacatvārīṃśat (‘9 + 40’) or ūnapaṇcāśat (‘one-dimensional 50’)</td>
</tr>
<tr>
<td>50</td>
<td>paṇcāśat</td>
</tr>
</tbody>
</table>

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):

<table>
<thead>
<tr>
<th>Mikir</th>
<th>Meithei</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>isi</td>
</tr>
<tr>
<td>TWO</td>
<td>hini</td>
</tr>
<tr>
<td>EIGHT</td>
<td>nirkep</td>
</tr>
<tr>
<td>NINE</td>
<td>sirkep</td>
</tr>
<tr>
<td>TEN</td>
<td>kep</td>
</tr>
</tbody>
</table>

4.2.0.4 ADDITIVE FORMATIONS OF THE BASIC UNIT NUMERALS38

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:

Khmer39

<table>
<thead>
<tr>
<th>ONE</th>
<th>SIX</th>
<th>FIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>muay</td>
<td>prammuay</td>
<td>pram</td>
</tr>
<tr>
<td>pii</td>
<td>prampii</td>
<td>dop</td>
</tr>
<tr>
<td>bøy</td>
<td>prambøy</td>
<td></td>
</tr>
<tr>
<td>buän</td>
<td>prambuän</td>
<td></td>
</tr>
<tr>
<td>pram</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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’, ‘VIII’, 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 üks‘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).

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

<table>
<thead>
<tr>
<th>ONE</th>
<th>se</th>
<th>SIX</th>
<th>zokkay-se kanay</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>nay</td>
<td>SEVEN</td>
<td>zokkay-se katam</td>
</tr>
<tr>
<td>THREE</td>
<td>tam</td>
<td>EIGHT</td>
<td>zokkay-nay</td>
</tr>
<tr>
<td>FOUR</td>
<td>bra</td>
<td>NINE</td>
<td>zokkay-nay kase</td>
</tr>
<tr>
<td>FIVE</td>
<td>ba</td>
<td>TEN</td>
<td>dos (&lt; Indo-Aryan)</td>
</tr>
</tbody>
</table>

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 Ranîtar dialect of Bantawa:

Bantawa (Ranîtar dialect)

<table>
<thead>
<tr>
<th>ONE</th>
<th>ük-ta</th>
<th>SIX</th>
<th>bhan-ka-chuk</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>hûa-ta</td>
<td>SEVEN</td>
<td>bhan-hû-chuk</td>
</tr>
<tr>
<td>THREE</td>
<td>sum-ka</td>
<td>EIGHT</td>
<td>bhan-sum-chuk</td>
</tr>
<tr>
<td>FOUR</td>
<td>ret-kat</td>
<td>NINE</td>
<td>bhan-ret-chuk</td>
</tr>
<tr>
<td>FIVE</td>
<td>ük-chuk</td>
<td>TEN</td>
<td>hûa-chuk</td>
</tr>
</tbody>
</table>

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 ük ‘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)

<table>
<thead>
<tr>
<th>ONE</th>
<th>kolok</th>
<th>SIX</th>
<th>muktapi usongbi kolok</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>hitci</td>
<td>SEVEN</td>
<td>muktapi usongbi hitci</td>
</tr>
<tr>
<td>THREE</td>
<td>sumji</td>
<td>EIGHT</td>
<td>muktapi usongbi sumci</td>
</tr>
<tr>
<td>FOUR</td>
<td>sumcibi usongbi kolok</td>
<td>NINE</td>
<td>mukcurukbi kolok hongbi</td>
</tr>
<tr>
<td>FIVE</td>
<td>muktapi</td>
<td>TEN</td>
<td>muktapi hita</td>
</tr>
</tbody>
</table>

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

SIX *d-ruk (STC)/*d-krok (JAM)

In TSR #35 I reconstruct this etymon as Proto Lolo-Burmese *C-krok, where ‘C-’ stands for a voiced prefix for which there is evidence on tonal grounds, and which I interpret as pointing to PTB *d-krok.

The problem is that many TB languages reflect a VELAR prefix (or at any rate a velar component in the prevocalic part of the syllable), instead of – or in addition to – a dental one. STC (note 321, p.115) attempts to account for this in terms of regular sound change of an initial *dental group to a velar one, that is ‘prefixal *d-r-‘ > Written Burmese khr- (versus ‘cluster *dr-‘ > WB khy-), but this is not very convincing. 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 tug; Lepcha tarok; Digaro thar; 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 dru?; Jirel thuk; Sherpa tung; Sikkim Bhotia tuk; Gurung tugh; Tamang tuhl; Thakali tuhl; Kulung tuk-chi; Limbu tuk-si; Chamling tukara

Ao t(े)rok; Kimsing tarok; Lotha tirok; Meitei tark; Meluri tabor; Mikir throk; Ntenyi togo, tōo; Pochury toro; Sangtam thuro; Tangkhul tharuk; Tangsa (Moshang) taruk; (Yogli) tiiruk; 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; Nruangmei ciunei; Rengma tsaro; Sema tsogho

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

\[
\begin{array}{lll}
\text{*kruk} & \text{Written Burmese} & \text{khrok} \\
\text{Achang} & \text{Xzo}^{55} & \\
\text{Langsu} & \text{khjauk}^{55} & \\
\text{Zaiwa} & \text{khju}^{55} & \\
\text{Nusu} & \text{kh}'u^{53} & \\
\text{Naxi (Yongning)} & \text{kho}^{13} & \\
\text{Jingpho} & \text{kru}^{55} & \\
\text{Trung} & \text{k'lu}^{44} & \\
\end{array}
\]

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.
J.T. Sun (1993:132) reconstructs Proto Tani *kr̂a, on the basis of Apatani x safeguard, Bengni a-kju, Bokar (=Adi=Lhoba) a-ku, Padam (=Abor) a-ke, Gallong ak-ka, 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 -η: 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 *kr̂aŋ.

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 tə41k' o54
   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 tʃu54; Sani khu22; Hani (Gao Huanian 1955) khu2; Hani (Hu and Dai 1964) ku2; Lisu (Fraser 1922) hchaw6; Luquan (Ma Xueliang 1949) tʃ' u55; Nasu (Gao Huanian 1958) tʃu54; Moso tʃ'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: keta (Zida dialect); katruk, truk, keto, ki-trog, ka-tshuo, koco, ktru, ku-tok (cited in Nagano 1984); and katšok (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 tʃu55, Pumi (Taoba) tʃu55, Muya teỹi55, Queyu tʃ' o, Shixing tʃ' o55) but Qiang

42 Kom Rem and Puiron have a velar prefix with other high numerals as well: Kom Rem and Puiron karet '8'; Puiron kak'aw '9' (see §5.4.4.2, §5.4.4.3).
(Mawo) $xt\check{s}a$, (Taoping) $xt\check{su}$ (ZMYYC #916) seem clearly to point to a complex proto-consonant group where the first element was a velar, $^*k-d$-$ruk$.43

Garo (Mom in) 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 $^*li\check{ok}$ in GSR #1032 (see §1.2.6).

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

- (Himalayish) Kaire ru; Khaling ra; Thulung Rai ru
- (Northern Naga) Chang lak; Konyak wok; Phom vok
- (Abor-Miri-Dafla) Aka (LSI) rieh; Chulikata ahe;
  Sulong (ZMYYC #916) yok
- (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 gup44

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 shühüh; Chokri shwüri; 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) wtuhau 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 qhu and Guiqiong khó).
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) \( f^{44} \), 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 \( ku^{21}tam^{53} \). The stop/nasal interchange is paralleled in EIGHT (Miju §grim; 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 \( *-\text{sum} \times *-\text{sam} \) (‘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:

\[ \text{khu '5'} / \text{khit '6'} / \text{sit '7'}. \]

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 \( *s\text{-}nis^{47} \)/ PLB \( *s(n)i\text{-}t \) (TSR #128)

As we have already pointed out several times,\(^{48}\) the similarity of this reconstruction to that for TWO, \( *g\text{-}nis \), 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 \( *s\text{-}nis \) as meaning FIVE. The prefix \( *s\text{-} \) is of no help in this connection, since the TB root for FIVE either took labial (\( b-, m- \)) or lateral (\( l\text{-} \)) prefixes, not \( s\text{-} \).

WB has \( hnac '2', khu'\text{-}hnac '7', \) 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 \( *s\text{-} \) PREFIX, WITH RETENTION OF THE ROOT-INITIAL

STC lists Kanawari \( stis^{50} \) (with denasalisation of initial), Rgyalrong \( kēnēs \sim kēsēnēs \), Garo \( sni \), and Jingpho \( sən̂it \). To these we may add:

\(^{46}\) Lahu \( f\text{-} \), for example, comes from PLB \( *hw\text{-} \) or \( *?w\text{-} \). 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) sur\textsuperscript{21} nyit\textsuperscript{55}
Boro shini; Dimasa sini; Kokborok chini
Newari nhæ-gu:

Written Burmese khu'-hnac (alongside hnac ‘2’); Atsi n?yit; Maru n?at (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) sta (with denasalisation of the root-initial); Pumi (Taoba) ɺi\textsuperscript{25}; Ergong snie/snie; Queyu ɺa\textsuperscript{55}

4.2.2.1 WITH DEVELOPMENT OF PREFIXAL *s- TO A DENTAL STOP OR AFFRICATE

Angami thena, thenie; Ao (Chungli) tenet, (Mongsen) teni; Chakhesang thena; Chokri thûna; Yimchungru thûnie; Yacham-Tengsa thanyet; Sangtam thûnye
Liangmai chania; Mao chanî; Nruanghmei cûnei; Rengma tsanû; 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 sagî; Lakher sari; Lushai pasarih; Paite sagih; Puiron sari; Thado sAGî; Tiddim sâgî; Vaiphei sagî

Manipur/Naga

Meithei taret (cf. taruk ‘6’); Meluri terü (cf. taro ‘6’); Ntenyi tûghû (cf. togho, tûo ‘6’); Pochury tûrû (cf. toro ‘6’).

The STC regards these forms as reflecting a quite separate root from *s-nis,\textsuperscript{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.

\textsuperscript{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\(^{53}\); Kaise ne
Karen\(^{54}\) (Pa-O) nət, (Pwo) nwè, (Palaychi) nwiq, (Sgaw) nwí
Konyak nyit; Phom and Chang nyet
Guiqiong ni\(^{55}\); Muya nyi\(^{35}\)

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/shi; Hani (Gao Huanian 1955) sj\(^{21}\)

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

Lahu ši; Lisu (Fraser) shi\(^{5}\); (Jui 1958) sj\(^{11}\); 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 eta\-m 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.

---


\(^{53}\) Cited as “ngis” in CSDPN, an obvious typo.

\(^{54}\) Benedict (1979:13) sets up Proto Karen *hnos, 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 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 ʃʃi, Shixing ʂe, Pumi (Jinghua) ʐi

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 ʃʃi 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, ɕi, such that the original prefix is now the root-initial, and the original root-initial is now the syllable-final consonant.

(d) Sharchop and Monpa

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

Sharchop (=Tsangla) zon; 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:

---

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.

Abor-Miri *ki-nit*; Padam-Mising (Tabu Taid) *kwnut*; Lhopa (=Bokar) *kwnur*; Apatani *kanu*; Bengni *ka-ni*; Minyong *kenit*; Dafla (Das Gupta), Gallong, Padam, Tagin *kane*; Nishi (Dubey) *ken*.

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-nyl*), or by innovating a new root for SEVEN (for example, Chulikata *joh*, Taraon *we*, German Deng *nuun*<sup>53</sup>, Idu *iū*, Mi ji *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) *kešnyit ~ kešnyis ~ 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’, *ggni* ‘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) *nuun*<sup>54</sup> (no parallelism with *kuw²jin*<sup>53</sup> ‘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 *grum*<sup>53</sup> ‘8’), even though the latter is to be reconstructed with PTB *₇₆₄.*

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:

---

<sup>60</sup> 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.

<sup>61</sup> 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) duín (the vowel in Chhewang Rinzin’s speech is t); Baima de\(^{13}\) (pers.comm. Sun Hongkai 1991); also probably Khaling tā~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 nuun\(^{53}\) 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
- Deng Darang
- Idu (Talukdar et al. 1962)
- Idu (Sun 1983)
- Mishmi
- Chulikata (LSI)

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’:

\[
\text{throk} \ '6' + \ '1' = \text{throk}-\text{si} \ '7'.
\]

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 be some contamination here from FIVE *l-qa?

(B) Aka (=Hruso) (LSI) mulh.
This slightly resembles the last syllable of the Milang form.

(C) Dhammai mja?, MJII (Simon 1979) myah
These closely resemble the Aka form.

(D) Lepcha ka-kyak.
This is part of a 7-10 prefix run, also including ka-ku ‘8’, ka-kyot ‘9’, ka-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

EIGHT *b-r-gyat ≠ *b-g-ryat

Key forms for justifying this doubly prefixed reconstruction are WT brgyad and Jingpho matsát. To these we may add Rgyalrong (Zida) warzhe(t), with wa- also < *b-. 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 *pwāt (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.

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-ka w ‘9’”. The STC is here groping toward the notion of ‘prefix run’ (see §5.2ff.).
4.2.3.1 KUKI-CHIN-NAGA FORMS WITH DENTAL OR AFFRICATE PREFIX

Lakher *chari*; Mao *chacha*; Tangkhul *chishat*.

Angami and Chakesang *thetha*; Chokri *títha*; Kezhama *tíche*; Kimsing *tecat*; Liangmai *tachat*; Lotha *tiza*; Meluri and Pochury *túze*; Nruangmei *tacút*; Ntenyi *tíza*; Rengma *tíkhü*; Sema *thache*; Tangsa (Moshang) *tachat*, (YogI) *túchat*; Yacham-Tengsa *thesep*, *teset*; Yimchungri *tísha*; Zeliang *tesat*.

The Liangmai and Nruangmei forms break up runs of numerals with affricate prefixes:

<table>
<thead>
<tr>
<th></th>
<th>Liangmai</th>
<th>Nruangmei</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIX</td>
<td>charuk</td>
<td>cúruk</td>
</tr>
<tr>
<td>SEVEN</td>
<td>chania</td>
<td>cúnei</td>
</tr>
<tr>
<td>EIGHT</td>
<td>tachat</td>
<td>tacút</td>
</tr>
<tr>
<td>NINE</td>
<td>chakiuh</td>
<td>cúkùu</td>
</tr>
</tbody>
</table>

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. *matsát*).

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: *kākyak* ‘7’, *kākū* ‘8’, *kākyót* ‘9’ and *kāti* ‘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 *hset* (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 *piao*, 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; Kaire kye; Sherpa ge

Gangte giet; Kuki get; Paite, Tiddim, Vaiphei giat

Sangtams ke is a lone monosyllabic form that breaks up a 6-10 run of dental prefixes (thüro ‘6’, thüyne ‘7’, tiku ‘9’, thüre ‘10’).

Qiang (Mawo) kha lå (ZMYYC #1291) looks like an apocopated and metathesised form, and resembles several other reflexes with final -r or a rhotacised vowel (for example, Nusu ša53 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 *?rit *?ryat > WB hrac; Lahu hi; Akha yeh; Ahi xi44; Sani he22; Hani (Gao Huanian 1955) xae21; Hani (Hu and Dai 1964) še21; Lisu (Fraser) h’6; Luquan ?han55, Nasu (Gao Huanian 1958) xen34; Achang cer55, Zaiwa fıt55, Maru je?55; Anong cer55, Naxi (Lijiang) xo55, (Yongning) xu13 (see TSR #171)70

Jinuo xe44; Tuji jie21

The aberrant and isolated Sulong language of Arunachal Pradesh has a form with liquid initial, la53 (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 šat (alongside pa-šat ‘18’, sha ‘9’); Wancho achat; Nocte i-sat; Tangsa (Yogli) tachat, (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 ʒyi (< *r-γy); Ersu ʒ155; Muya cye53; Pumi (Jinghua) sue23, (Taoba) cyv15; Qiang (Taoping) tʃ55; Queyu cye55; Shixing cyi55. A couple of Qiangic languages have forms with initial semivowel or h-: Namuiyi hi33 (with rhinoglottophilia); Guiqiong je55.

Bai (Bijiang) tcua44. Other Bai dialects have apparent loans from Chinese; cf. Jianchuan and Dali pia44, §4.2.3.2.

Dul ung cät55; Trung šiat44; Nusu ša53

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-γyat, with unspecified vocalic or consonantal prefix.
Karenic: PA-O sóːt; PWO xò²; Palaychi xó; Sgaw xó?. 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) giun⁵¹; 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 lium³⁵, Idu (Sun 1983) i⁵⁵liong³⁵ (alongside i⁵⁵gyűŋ⁵³ ‘7’); Idu (Talukdar 1962) inyû (=-ny- apparently < earlier -ly-; compare also Idu iù ‘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

<table>
<thead>
<tr>
<th>Language</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abor-Miri</td>
<td>a-pi ‘4’, a-nyi ‘2’ &gt; pi-nyi ~ pui-nyi ‘8’</td>
</tr>
<tr>
<td>Apatani</td>
<td>a-pi ‘4’, nyi ‘2’ &gt; pu?-nyi ~ pry?-nyi ‘8’</td>
</tr>
<tr>
<td>Lhopa</td>
<td>api: ‘4’, anyi ‘2’ &gt; pi-nyi ‘8’</td>
</tr>
</tbody>
</table>

Similarly: Dafla, Gallong, Padam, Tagin pine, Yano plæ-ne; Minyong pini⁷⁴; Nyisu plin; Nishi pin, piën.⁷⁵

Other multiplicative formations for EIGHT in TB include Boro zokkay nay (§3.3.2[C]; §4.2.0.1), and Kayah (= Karenni = Red Karen) lwīswā’ (§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-kow). For another possibility, see §4.2.4.6.
⁷⁴ Compare Minyong a-nyi ‘2’, but a-ki ‘4’ (< ?). Note that the inherited PTB root for FOUR, *b-lay, 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*, *tıku*; Chakhesang *thechi*; Chokri *tıčhi*; Kezhama *tepfü*; Kimsing *tak(a)tu*; Konyak *tu* (with preemption of the initial); Lotha and Sema *toku*; Maring *tako*; Meluri *tokhu*; Moshang *takru* (-r- < ?); Ntenyi *tıkıru*; Pochury *toku*; Sangtam *tıku*; Yacham-Tengsa *thakru*; Yimchunguru *tıku* and Yogli *tıka`. To these we may add extra-KCN forms like Monpa (Cuona: Sun et al. 1980) *tu²[^1]ku*[^52], (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 *cıiku*, 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 *jákhu*. I have already observed in a previous analysis of this etymon78 that Jingpho *já-* here may well come from *s-*, given the fact that the Jingpho causative prefix *sá-* (< PTB *s*-) undergoes a predictable morphophonemic change to *já-*, 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) *zung* and Garo *sku* (also Dimasa *sugu*). Similar to the Nung form cited in STC are Anong *duwa*[^31] and Dulong *duwa*[^33] (ZMYYC).

Qiang (Taoping) *xguns*[^31] 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 *duwa*[^31] and Dulong *duwa*[^33] (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-* (*rgua*), which is probably velar in articulation like Taoping *x-*, and could descend from *s-* as well.
(= Hruso: data from LSI) stheu, sthō '9'. This puzzling form may reflect a doubly-prefixed prototype *s-d-(k)w 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 סו (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) tsi3 (This is the only form with an affricate initial in Loloish proper, but compare also Jinuo tcy3.)

Bai (Dali and Jianchuan) tceu3, Bijiang tci3

4.2.4.2 WITH VELAR PREFIX

A secondary velar prefix is occasionally found with NINE: Puiron kakwa (alongside karet '8'); Lepcha kakyót (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 kangu (all Rgyalrong numerals 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 ngie, Muya nguw35; Ersu nge33 and Namuyi ngu33), 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 gur33, Queyu gu3, Shixing gu3, Thulong gu, Khaling gfu

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

Lolo-Burmese

PLB *gəw > WB kūi; Lahu qū; Akha ʔp; Hani (Shuikui) ʔu31; Lisu ku55; Naxi (Yongning) gvr33; Achang kau3; Zaiwa kau2; Langsu (=Maru) kūk3 (the secondary -k is regular for the rhyme *-əw); Nusu gu53; Tujia kwe55

Karenic

Pho (Moulmein), Palaychi, Sgaw khwī; 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 (*k V-n(y)a-ŋ). 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 kinyaŋ
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 ka2ŋn̂unĝ55
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 nən̂55 mʊ51 (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 *k V-(n)aŋ.
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) *nuu*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-da!J-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)).
CHAPTER 5
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 \( \ast d-gaw \) \( \ast s-kaw \)). 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 \( \ast s-nis \) (see §4.2.2). Some languages preserve an overt trace of the prefix (Qiang (Mawo) sta, Pumi (Taoba) \( \phi t \)\(^{35} \), Ergong snie/sjie and Queyu \( p\ddot{a} \)\(^{35} \)); others have lost the prefix entirely (Guixiong \( n\ddot{t} \)\(^{55} \) and Muya \( nyi \)\(^{35} \)); while still others show preemption by the prefix of the root-initial (Namuyi \( sj \)\(^{33} \), Shixing \( s\ddot{g} \)\(^{55} \), Qiang (Taoping) \( ci\ddot{g} \)\(^{33} \) and Pumi (Jinghua) \( xi\ddot{e} \)\(^{13} \).

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. 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, ka- versus ga-, 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, Serna kini ‘2’, but kîlthu ‘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 bît, Langsu pjîk31 (ZMYYC 1287)). With respect to EIGHT, STC recognises metathesis even for the proto-level: *b-r-ryat ≠ *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-ga), the root-vowel -a had disappeared, so that the 
former root-initial consonant m- 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:

<table>
<thead>
<tr>
<th>PREFIX</th>
<th>PREFIXAL VOWEL</th>
<th>ROOT INITIAL</th>
<th>ROOT VOWEL</th>
<th>FINAL CONSONANT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PTB</td>
<td>b</td>
<td>e</td>
<td>0</td>
<td>zero</td>
</tr>
<tr>
<td>Aka</td>
<td>ph</td>
<td>u</td>
<td>m</td>
<td></td>
</tr>
</tbody>
</table>

We have found a number of monosyllabic forms of this apocopated/preemptive type, 
including: Puiron *pang ‘5’, Khaling *bhom ‘5’, Boro and Kokborok *ba ‘5’ (all < *b-ga); 
Boro *do ‘6’ (< *d-ruk; all cited in §4.1.4.1.1); Chulikata *kåsh ‘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) *ciŋ ‘3’ and Ersu *tiŋ ‘7’ (< *s-nis; §4.2.2.4(c)).

This process applies equally well to disyllabic 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: ? ‘2’); Nyisu *pliŋ and Nishi *pin ‘8’ (< *pr-ŋi ‘4 x 2’; §4.2.3.7); Hill Miri *cem-piŋ ‘80’ (< *cam-p(r)i-ŋi).

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. 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 4 5

*3-tyik, *g-nis, *g-sum
*3-b-laŋ, *b-ga (≠ *l-ga)

Consider the numerals of Rgyalrong (= Jiarong):³

<table>
<thead>
<tr>
<th>ONE</th>
<th>TWO</th>
</tr>
</thead>
<tbody>
<tr>
<td>*tšek</td>
<td>*kene</td>
</tr>
</tbody>
</table>

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.com.). Ives Goddard (pers.com. 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 
ofered for comparison are from Nagano (1984).
THREE
FOUR
FIVE
SIX
SEVEN
EIGHT
NINE
TEN
ELEVEN
TWELVE

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 ɡ-, (usually vocalised with shwa as ɡa-) 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 ɡa- in ‘1–2’ (lajj, lakho) and ɡa- in ‘3–5’ (māsām, mālī, maŋā), 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 ɡe- 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 ɡe- 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:

<table>
<thead>
<tr>
<th></th>
<th>PTB</th>
<th>Proto Rgyalrong</th>
<th>Zida</th>
</tr>
</thead>
<tbody>
<tr>
<td>FOUR</td>
<td>*b-lōy</td>
<td>*k-b-liy</td>
<td>ke-w-dzi</td>
</tr>
<tr>
<td>FIVE</td>
<td>*b-ŋa</td>
<td>*k-m-ŋa</td>
<td>ke-m-nga</td>
</tr>
<tr>
<td>SIX</td>
<td>*d-ruk</td>
<td>*k-d-ruk</td>
<td>ke-ṭa (t &lt; *-d-r-)</td>
</tr>
<tr>
<td>SEVEN</td>
<td>*s-nis</td>
<td>*k-s-nis</td>
<td>ke-šnyes, etc.</td>
</tr>
</tbody>
</table>

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 *sjaed ‘4’ arose through the influence of the sibilant in ‘3’ *(t)sam. (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:

\[
\begin{align*}
\text{laŋåi, \ lækʰøŋ – məsûm, məli, məŋå} \\
\text{krú?} \\
\text{sənɨt, mətsát, jəkhû} \\
\text{ṣì}
\end{align*}
\]

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:

\[
1, 2 \ldots 3, 4, 5 \ldots 6 \ldots 7, 8, 9 \ldots 10.
\]

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:

\[
\begin{align*}
\text{ichi, ni} \\
\text{san, shi, go} \\
\text{ roku, shichi, hachi} \\
\text{k uu, j uu}
\end{align*}
\]
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.⁴⁵

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):

<table>
<thead>
<tr>
<th></th>
<th>Written Tibetan</th>
<th>Lhasa Tibetan</th>
<th>Sherpa</th>
<th>Jirel</th>
<th>Kaike</th>
<th>Dzongkha⁵</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>gCig ciq</td>
<td>ciːq</td>
<td>cikq</td>
<td>dokpei</td>
<td>ti</td>
<td>chi</td>
</tr>
<tr>
<td>TWO</td>
<td>gnyis ngyi</td>
<td>nyii</td>
<td>ngyi</td>
<td>nytq</td>
<td>nghyi</td>
<td>nyi</td>
</tr>
<tr>
<td>THREE</td>
<td>gsun sum</td>
<td>sum</td>
<td>sumq</td>
<td>sumq</td>
<td>sum</td>
<td>sum</td>
</tr>
<tr>
<td>FOUR</td>
<td>bži shi</td>
<td>ji</td>
<td>ji</td>
<td>syi</td>
<td>li</td>
<td>zhi</td>
</tr>
<tr>
<td>FIVE</td>
<td>lpha nga:q</td>
<td>nga</td>
<td>nga:q</td>
<td>nga:q</td>
<td>nga:q</td>
<td>nga:</td>
</tr>
<tr>
<td>SIX</td>
<td>drug thuk</td>
<td>thuu</td>
<td>tuk</td>
<td>thuk</td>
<td>ru</td>
<td>tuk</td>
</tr>
<tr>
<td>SEVEN</td>
<td>bdun tüün</td>
<td>din</td>
<td>duin</td>
<td>ne</td>
<td>duin</td>
<td></td>
</tr>
<tr>
<td>EIGHT</td>
<td>brgyad kEE</td>
<td>ge</td>
<td>gyet</td>
<td>kye</td>
<td>gye</td>
<td></td>
</tr>
<tr>
<td>NINE</td>
<td>dgu gu</td>
<td>gu</td>
<td>gu</td>
<td>gu</td>
<td>gu</td>
<td></td>
</tr>
<tr>
<td>TEN</td>
<td>buc cu</td>
<td>citham:baːq</td>
<td>cyutaːm:baːq</td>
<td>chyu-tamba</td>
<td>chu-</td>
<td></td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>SIX</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>tu</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TWO</td>
<td>saːhpu</td>
<td>SEVEN</td>
<td>taeɾ</td>
</tr>
<tr>
<td>THREE</td>
<td>suhpu</td>
<td>EIGHT</td>
<td>ri</td>
</tr>
<tr>
<td>FOUR</td>
<td>bhacel</td>
<td>NINE</td>
<td>ghu</td>
</tr>
<tr>
<td>FIVE</td>
<td>bhom</td>
<td>TEN</td>
<td>tadam</td>
</tr>
</tbody>
</table>

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: din.
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

<table>
<thead>
<tr>
<th></th>
<th>Gurung</th>
<th>Tamang</th>
<th>Thakali</th>
<th>Newari</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>grihq</td>
<td>kih</td>
<td>tih</td>
<td>cha-gu(-li)</td>
</tr>
<tr>
<td>TWO</td>
<td>ngihq</td>
<td>nyi:h</td>
<td>ngih</td>
<td>ni-gu(-li)</td>
</tr>
<tr>
<td>THREE</td>
<td>söq</td>
<td>som</td>
<td>som</td>
<td>swä-gu:</td>
</tr>
<tr>
<td>FOUR</td>
<td>plihq</td>
<td>plih</td>
<td>plih</td>
<td>pē-gu:</td>
</tr>
<tr>
<td>FIVE</td>
<td>nga:hpq</td>
<td>nga:h</td>
<td>nga:h</td>
<td>nya:-gu:</td>
</tr>
<tr>
<td>SIX</td>
<td>tühq</td>
<td>tuh</td>
<td>tuh</td>
<td>khu-gu:</td>
</tr>
<tr>
<td>SEVEN</td>
<td>ngiq</td>
<td>nyis</td>
<td>ngis</td>
<td>nhae-gu:</td>
</tr>
<tr>
<td>EIGHT</td>
<td>prehq</td>
<td>preht</td>
<td>preh</td>
<td>cya:-gu:</td>
</tr>
<tr>
<td>NINE</td>
<td>kuq</td>
<td>ku</td>
<td>ku</td>
<td>gū-gu:</td>
</tr>
<tr>
<td>TEN</td>
<td>cyuq</td>
<td>ci</td>
<td>cyu</td>
<td>jhi-gu:</td>
</tr>
</tbody>
</table>

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 t 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)

<table>
<thead>
<tr>
<th></th>
<th>Kanawari</th>
<th>Lepcha</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>id</td>
<td>kat</td>
</tr>
<tr>
<td>TWO</td>
<td>nish</td>
<td>nyät; nyi</td>
</tr>
<tr>
<td>THREE</td>
<td>shum</td>
<td>sam</td>
</tr>
<tr>
<td>FOUR</td>
<td>pü</td>
<td>fəli</td>
</tr>
<tr>
<td>FIVE</td>
<td>nga</td>
<td>fəngo</td>
</tr>
<tr>
<td>SIX</td>
<td>tuk</td>
<td>tərak</td>
</tr>
<tr>
<td>SEVEN</td>
<td>stish</td>
<td>kə-kyək</td>
</tr>
<tr>
<td>EIGHT</td>
<td>rai</td>
<td>kə-kü</td>
</tr>
<tr>
<td>NINE</td>
<td>zgui</td>
<td>kə-kyöt</td>
</tr>
<tr>
<td>TEN</td>
<td>sai</td>
<td>kə-ti</td>
</tr>
</tbody>
</table>
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)ay); 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 fa-, and the dental prefix ta- 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 -kyot ‘9’ apparently < *gyat EIGHT, and ku ‘8’ apparently < *d-kaw NINE (§4.0.2). The form kyak for SEVEN remains a complete mystery.

5.3.5 MONPA DIALECTS

<table>
<thead>
<tr>
<th>Monpa (Dubey 1983)</th>
<th>M.Cuona (Sun et al.1980)</th>
<th>Central Monpa (Das Gupta 1968)</th>
<th>M.Motuo (Sun et al. 1980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>thee</td>
<td>t’e53</td>
<td>thur</td>
</tr>
<tr>
<td>TWO</td>
<td>nai</td>
<td>nai23</td>
<td>n(y)itsing</td>
</tr>
<tr>
<td>THREE</td>
<td>sum</td>
<td>sum53</td>
<td>sam</td>
</tr>
<tr>
<td>FOUR</td>
<td>blee</td>
<td>pli53</td>
<td>b(i)s/i(p)i</td>
</tr>
<tr>
<td>FIVE</td>
<td>lenga</td>
<td>le21nge53</td>
<td>nga</td>
</tr>
<tr>
<td>SIX</td>
<td>gro</td>
<td>kro72/54</td>
<td>khung</td>
</tr>
<tr>
<td>SEVEN</td>
<td>nis</td>
<td>nis55</td>
<td>zum</td>
</tr>
<tr>
<td>EIGHT</td>
<td>giet</td>
<td>cen13</td>
<td>yen</td>
</tr>
<tr>
<td>NINE</td>
<td>dugu</td>
<td>tu21ku54</td>
<td>gu</td>
</tr>
<tr>
<td>TEN</td>
<td>chi</td>
<td>tci54</td>
<td>se</td>
</tr>
</tbody>
</table>

- 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, khoṅ, 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-lay, *b-ga).

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!)

<table>
<thead>
<tr>
<th>Maring</th>
<th>Khat</th>
<th>Khani</th>
<th>Khiyum</th>
<th>Phili</th>
<th>Phanga</th>
<th>Tharuk</th>
<th>Ani</th>
<th>Chat</th>
<th>Taka</th>
<th>Chip</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>khat</td>
<td>khani</td>
<td>khiyum</td>
<td>phili</td>
<td>phanga</td>
<td>tharuk</td>
<td>ani</td>
<td>chat</td>
<td>taka</td>
<td>chip</td>
</tr>
</tbody>
</table>

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-gaw). 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

<table>
<thead>
<tr>
<th>Chang</th>
<th>Phom</th>
<th>Konyak</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>chie</td>
<td>hük</td>
</tr>
<tr>
<td>TWO</td>
<td>nyi</td>
<td>nyi</td>
</tr>
<tr>
<td>THREE</td>
<td>lei</td>
<td>alì</td>
</tr>
<tr>
<td>FOUR</td>
<td>ngau</td>
<td>nga</td>
</tr>
<tr>
<td>FIVE</td>
<td>lak</td>
<td>vok</td>
</tr>
<tr>
<td>SIX</td>
<td>nyet</td>
<td>nyet</td>
</tr>
<tr>
<td>SEVEN</td>
<td>sat</td>
<td>shut</td>
</tr>
<tr>
<td>EIGHT</td>
<td>guh</td>
<td>shü</td>
</tr>
<tr>
<td>NINE</td>
<td>an</td>
<td>an</td>
</tr>
</tbody>
</table>

7 Maring khat ‘1’ is non-prefixed, a different etymology from *g-tyik. See §3.1.2, §3.1.4.
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

<table>
<thead>
<tr>
<th></th>
<th>Paite</th>
<th>Tiddim</th>
<th>Gangte</th>
<th>Kuki</th>
<th>Thado</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ONE</strong></td>
<td>khat</td>
<td>—</td>
<td>khat</td>
<td>—</td>
<td>xát</td>
</tr>
<tr>
<td><strong>TWO</strong></td>
<td>nih</td>
<td>nih</td>
<td>nih</td>
<td>ni; ba</td>
<td>ní ~ ní</td>
</tr>
<tr>
<td><strong>THREE</strong></td>
<td>thum</td>
<td>thum</td>
<td>thum</td>
<td>thum</td>
<td>thúm; thíng</td>
</tr>
<tr>
<td><strong>FOUR</strong></td>
<td>li</td>
<td>li</td>
<td>li</td>
<td>li</td>
<td>li</td>
</tr>
<tr>
<td><strong>FIVE</strong></td>
<td>nga</td>
<td>nga</td>
<td>ngâ</td>
<td>nga</td>
<td>—</td>
</tr>
<tr>
<td><strong>SIX</strong></td>
<td>guk</td>
<td>guk</td>
<td>gûp</td>
<td>gup</td>
<td>gûp</td>
</tr>
<tr>
<td><strong>SEVEN</strong></td>
<td>sagih</td>
<td>sagi?</td>
<td>sagih</td>
<td>sâgi</td>
<td>sâgi</td>
</tr>
<tr>
<td><strong>EIGHT</strong></td>
<td>giat</td>
<td>giat</td>
<td>giet</td>
<td>get</td>
<td>—</td>
</tr>
<tr>
<td><strong>NINE</strong></td>
<td>kua</td>
<td>kua</td>
<td>kuo</td>
<td>ko</td>
<td>—</td>
</tr>
<tr>
<td><strong>TEN</strong></td>
<td>sawm</td>
<td>sawm</td>
<td>sâwm</td>
<td>som</td>
<td>sòm</td>
</tr>
</tbody>
</table>

- 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**.
- **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).
5.4.2.1 ADDITIVE AND SUBTRACTIVE DEGENERATION OF THE PROTO-SYSTEM

Mikir

<table>
<thead>
<tr>
<th>Number</th>
<th>Mikir</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>isi</td>
</tr>
<tr>
<td>TWO</td>
<td>hini</td>
</tr>
<tr>
<td>THREE</td>
<td>kethom</td>
</tr>
<tr>
<td>FOUR</td>
<td>phir '4'</td>
</tr>
<tr>
<td></td>
<td>(phli-kep '4') (GEM);</td>
</tr>
<tr>
<td></td>
<td>phli '4'</td>
</tr>
<tr>
<td></td>
<td>(phli-kep '40') (Grüssner)</td>
</tr>
<tr>
<td>FIVE</td>
<td>phongo, pho (GEM)</td>
</tr>
<tr>
<td></td>
<td>pho, phonho (Grüssner)</td>
</tr>
<tr>
<td>SIX</td>
<td>throk</td>
</tr>
<tr>
<td>SEVEN</td>
<td>throksi</td>
</tr>
<tr>
<td>EIGHT</td>
<td>nirkep</td>
</tr>
<tr>
<td>NINE</td>
<td>sirkep</td>
</tr>
<tr>
<td>TEN</td>
<td>kep</td>
</tr>
<tr>
<td>ELEVEN</td>
<td>kre-isi</td>
</tr>
</tbody>
</table>

- 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 /ri/ [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:

<table>
<thead>
<tr>
<th>Number</th>
<th>Lushai</th>
<th>Hmar</th>
<th>Vaiphei</th>
<th>Wancho</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>pakhat</td>
<td>pakhat</td>
<td>pakhat</td>
<td>tuta</td>
</tr>
<tr>
<td>TWO</td>
<td>pahnih</td>
<td>pahni</td>
<td>pani</td>
<td>ani, anyi</td>
</tr>
<tr>
<td>THREE</td>
<td>pathum</td>
<td>pathum</td>
<td>pathum</td>
<td>ajam</td>
</tr>
<tr>
<td>FOUR</td>
<td>pali</td>
<td>pali</td>
<td>pali</td>
<td>ali (Das Gupta), li (GEM)</td>
</tr>
</tbody>
</table>
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:**

<table>
<thead>
<tr>
<th></th>
<th>Tangkhul</th>
<th>Liangmai</th>
<th>Yimchungru</th>
<th>Lakher</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>akha/khatkha</td>
<td>khad</td>
<td>khülang</td>
<td>-kha; sa-</td>
</tr>
<tr>
<td>TWO</td>
<td>khani</td>
<td>nia</td>
<td>manie</td>
<td>-no</td>
</tr>
<tr>
<td>THREE</td>
<td>kathum</td>
<td>shum</td>
<td>asam</td>
<td>-thō</td>
</tr>
<tr>
<td>FOUR</td>
<td>mati</td>
<td>madai</td>
<td>phiyi</td>
<td>-pali</td>
</tr>
<tr>
<td>FIVE</td>
<td>phanga</td>
<td>mangiu</td>
<td>phüngü</td>
<td>-pangaw</td>
</tr>
<tr>
<td>SIX</td>
<td>tharuk</td>
<td>charuk</td>
<td>thruruk</td>
<td>-charu</td>
</tr>
<tr>
<td>SEVEN</td>
<td>shini</td>
<td>chania</td>
<td>thünie</td>
<td>-sari</td>
</tr>
<tr>
<td>EIGHT</td>
<td>chishat</td>
<td>(tachat)</td>
<td>tizha</td>
<td>-chari</td>
</tr>
<tr>
<td>NINE</td>
<td>chiko</td>
<td>chakiuh</td>
<td>tuku</td>
<td>-chaki</td>
</tr>
<tr>
<td>TEN</td>
<td>thara</td>
<td>kariu</td>
<td>thürü</td>
<td>-hraw; sy-</td>
</tr>
</tbody>
</table>

---

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-).

Yimchungri 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:

<table>
<thead>
<tr>
<th>Angami</th>
<th>Kohima</th>
<th>Chokri</th>
<th>Puiron</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>puo</td>
<td>po</td>
<td>pü</td>
</tr>
<tr>
<td>TWO</td>
<td>kenie</td>
<td>kena</td>
<td>kiina</td>
</tr>
<tr>
<td>THREE</td>
<td>se</td>
<td>se</td>
<td>sii</td>
</tr>
<tr>
<td>FOUR</td>
<td>die</td>
<td>da</td>
<td>da</td>
</tr>
<tr>
<td>FIVE</td>
<td>pengou</td>
<td>pengu</td>
<td>pingu</td>
</tr>
<tr>
<td>SIX</td>
<td>sorou</td>
<td>suru</td>
<td>swürü</td>
</tr>
<tr>
<td>SEVEN</td>
<td>thenie</td>
<td>thena</td>
<td>thüna</td>
</tr>
<tr>
<td>EIGHT</td>
<td>thetha</td>
<td>thetha</td>
<td>tüthä</td>
</tr>
<tr>
<td>NINE</td>
<td>thepfü</td>
<td>theku</td>
<td>thüchi</td>
</tr>
<tr>
<td>TEN</td>
<td>kerü</td>
<td>kerü</td>
<td>küri</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>ONE</th>
<th>inkhat</th>
<th>SIX</th>
<th>karuk</th>
</tr>
</thead>
<tbody>
<tr>
<td>TWO</td>
<td>inhni</td>
<td>SEVEN</td>
<td>sari</td>
</tr>
<tr>
<td>THREE</td>
<td>inthum</td>
<td>EIGHT</td>
<td>karet</td>
</tr>
<tr>
<td>FOUR</td>
<td>manli</td>
<td>NINE</td>
<td>ko:</td>
</tr>
<tr>
<td>FIVE</td>
<td>ranga</td>
<td>TEN</td>
<td>som</td>
</tr>
</tbody>
</table>
This language has a secondary 1–3 run with in-, similar to the syllabic prefixes of the languages in §5.4.4.6 below. 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 lja. 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 ‘melt’, 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

<table>
<thead>
<tr>
<th></th>
<th>Meluri</th>
<th>Pochury</th>
<th>Ntenyi</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>ke; kesü</td>
<td>khe</td>
<td>kesü</td>
</tr>
<tr>
<td>TWO</td>
<td>keni</td>
<td>küni</td>
<td>kenyi</td>
</tr>
<tr>
<td>THREE</td>
<td>keche</td>
<td>küche</td>
<td>keching; kechang</td>
</tr>
<tr>
<td>FOUR</td>
<td>mezu</td>
<td>mzü</td>
<td>mezhi; mezü</td>
</tr>
<tr>
<td>FIVE</td>
<td>manga</td>
<td>mnga</td>
<td>munga</td>
</tr>
<tr>
<td>SIX</td>
<td>taro</td>
<td>toro</td>
<td>togo; tüo</td>
</tr>
<tr>
<td>SEVEN</td>
<td>terü</td>
<td>türü</td>
<td>tüghü</td>
</tr>
<tr>
<td>EIGHT</td>
<td>tüze</td>
<td>tüze</td>
<td>tüza</td>
</tr>
<tr>
<td>NINE</td>
<td>tokhu</td>
<td>toku</td>
<td>tükhu</td>
</tr>
<tr>
<td>TEN</td>
<td>tera</td>
<td>türa</td>
<td>dagha; ta?a</td>
</tr>
</tbody>
</table>

* 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).
[B] MAO-NRUANGMEI

<table>
<thead>
<tr>
<th>One</th>
<th>Mao</th>
<th>Nruangmei</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>kali</td>
<td>khüt</td>
</tr>
<tr>
<td>Two</td>
<td>kahei</td>
<td>kanei; kiinei</td>
</tr>
<tr>
<td>Three</td>
<td>kosü</td>
<td>kathum</td>
</tr>
<tr>
<td>Four</td>
<td>padei</td>
<td>padei</td>
</tr>
<tr>
<td>Five</td>
<td>pongo</td>
<td>pangu</td>
</tr>
<tr>
<td>Six</td>
<td>choro</td>
<td>cüruk</td>
</tr>
<tr>
<td>Seven</td>
<td>chani</td>
<td>cünei</td>
</tr>
<tr>
<td>Eight</td>
<td>chacha</td>
<td>tacüt</td>
</tr>
<tr>
<td>Nine</td>
<td>choku</td>
<td>cükiu</td>
</tr>
<tr>
<td>Ten</td>
<td>chüro</td>
<td>ruh</td>
</tr>
</tbody>
</table>

- SIX has a palatal prefix, c- or ch-.
- Mao has a neater clumping than Nruangmei. Nruangmei’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-KHORAO

<table>
<thead>
<tr>
<th>One</th>
<th>Mzieme</th>
<th>Zeme</th>
<th>Zeliang</th>
<th>Khoirao</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>ket</td>
<td>kat; hangkat</td>
<td>kat</td>
<td>khat</td>
</tr>
<tr>
<td>Two</td>
<td>kena</td>
<td>kena</td>
<td>kena</td>
<td>kati</td>
</tr>
<tr>
<td>Three</td>
<td>ketsum</td>
<td>kechum</td>
<td>kechum</td>
<td>kathum</td>
</tr>
<tr>
<td>Four</td>
<td>madai; mdai</td>
<td>medai</td>
<td>mdai</td>
<td>malhi</td>
</tr>
<tr>
<td>Five</td>
<td>mengei</td>
<td>mengeu</td>
<td>mengei</td>
<td>manga</td>
</tr>
<tr>
<td>Six</td>
<td>heruk</td>
<td>seruk</td>
<td>heruk</td>
<td>saruk</td>
</tr>
<tr>
<td>Seven</td>
<td>hena</td>
<td>sena</td>
<td>sinna</td>
<td>sini</td>
</tr>
<tr>
<td>Eight</td>
<td>heset</td>
<td>desat</td>
<td>tesat</td>
<td>kachat</td>
</tr>
<tr>
<td>Nine</td>
<td>hekui</td>
<td>sekui</td>
<td>hekui</td>
<td>chaku</td>
</tr>
<tr>
<td>Ten</td>
<td>kerei</td>
<td>kereu</td>
<td>kerei</td>
<td>sara</td>
</tr>
</tbody>
</table>

- 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.

\[
\text{Zeliang is a kind of composite dialect or lingua franca, an acronym for Ze-me + Liang-mai.}
\]
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):

<table>
<thead>
<tr>
<th></th>
<th>Lotha</th>
<th>Yacham-Tengsa</th>
<th>Sangtam</th>
<th>Mongsen</th>
<th>Ao</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>ekha</td>
<td>khatu</td>
<td>khe; khiru</td>
<td>akha/ra</td>
<td>ka</td>
</tr>
<tr>
<td>TWO</td>
<td>eni/oni</td>
<td>anat</td>
<td>anyu</td>
<td>anet</td>
<td>ana</td>
</tr>
<tr>
<td>THREE</td>
<td>etham</td>
<td>asam</td>
<td>asang</td>
<td>asam</td>
<td>asem</td>
</tr>
<tr>
<td>FOUR</td>
<td>mezii</td>
<td>phale</td>
<td>muzyi</td>
<td>phuli</td>
<td>pezii</td>
</tr>
<tr>
<td>FIVE</td>
<td>mungo</td>
<td>phungu</td>
<td>munga</td>
<td>phanga</td>
<td>pengu</td>
</tr>
<tr>
<td>SIX</td>
<td>tirok</td>
<td>thelok</td>
<td>thiro</td>
<td>terok</td>
<td>trok</td>
</tr>
<tr>
<td>SEVEN</td>
<td>ti-ing</td>
<td>thanyet</td>
<td>thunye</td>
<td>teni</td>
<td>tenet</td>
</tr>
<tr>
<td>EIGHT</td>
<td>tiza</td>
<td>thesep, teset</td>
<td>ke</td>
<td>tsit</td>
<td>ti</td>
</tr>
<tr>
<td>NINE</td>
<td>toku</td>
<td>thaku</td>
<td>tuku</td>
<td>tuku</td>
<td>tuku</td>
</tr>
<tr>
<td>TEN</td>
<td>taro</td>
<td>thelu</td>
<td>thure</td>
<td>tera</td>
<td>ter</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Tangsa</th>
<th>Kimsing</th>
<th>Meithei</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>ashi</td>
<td>ashi</td>
<td>ama</td>
</tr>
<tr>
<td>TWO</td>
<td>ani</td>
<td>anei</td>
<td>ani</td>
</tr>
<tr>
<td>THREE</td>
<td>atum</td>
<td>adim</td>
<td>acam</td>
</tr>
<tr>
<td>FOUR</td>
<td>bali</td>
<td>bulai</td>
<td>balai</td>
</tr>
<tr>
<td>FIVE</td>
<td>banga</td>
<td>banga</td>
<td>bangi</td>
</tr>
<tr>
<td>SIX</td>
<td>taruk</td>
<td>turuk</td>
<td>tarok</td>
</tr>
<tr>
<td>SEVEN</td>
<td>mashi</td>
<td>mishi</td>
<td>mishi</td>
</tr>
<tr>
<td>EIGHT</td>
<td>tachat</td>
<td>tüchat</td>
<td>techat</td>
</tr>
<tr>
<td>NINE</td>
<td>takru</td>
<td>tükau</td>
<td>tak(a)u</td>
</tr>
<tr>
<td>TEN</td>
<td>rok-shi</td>
<td>rauk-shi</td>
<td>ro-shi</td>
</tr>
</tbody>
</table>

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):
5.4.4.6 WHERE THE LOWEST RUN HAS A FULLY SYLLABIC CVC- PREFIX

Nocte-Maram

<table>
<thead>
<tr>
<th></th>
<th>Nocte (GEM)</th>
<th>Nocte (Dubey)</th>
<th>Maram</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>vanthe</td>
<td>wanthe</td>
<td>hang-li-ne</td>
</tr>
<tr>
<td>TWO</td>
<td>vanyi</td>
<td>wanni</td>
<td>hangna</td>
</tr>
<tr>
<td>THREE</td>
<td>vanram</td>
<td>wanrom</td>
<td>hangtum</td>
</tr>
<tr>
<td>FOUR</td>
<td>beli</td>
<td>bali</td>
<td>madai</td>
</tr>
<tr>
<td>FIVE</td>
<td>bang a</td>
<td>bang</td>
<td>mingu</td>
</tr>
<tr>
<td>SIX</td>
<td>irok</td>
<td>iro:k</td>
<td>saruk</td>
</tr>
<tr>
<td>SEVEN</td>
<td>ingit</td>
<td>ingit</td>
<td>sina</td>
</tr>
<tr>
<td>EIGHT</td>
<td>isat</td>
<td>itset; iset</td>
<td>sachat</td>
</tr>
<tr>
<td>NINE</td>
<td>ikhu</td>
<td>ikhu</td>
<td>soki</td>
</tr>
<tr>
<td>TEN</td>
<td>ichi</td>
<td>ichi</td>
<td>kero</td>
</tr>
</tbody>
</table>

- 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

<table>
<thead>
<tr>
<th></th>
<th>Kezhama</th>
<th>Sema</th>
<th>Rengma</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>kele</td>
<td>laki; khe</td>
<td>me</td>
</tr>
<tr>
<td>TWO</td>
<td>kenhi</td>
<td>kini</td>
<td>kholhung</td>
</tr>
<tr>
<td>THREE</td>
<td>katsü</td>
<td>küthu</td>
<td>keshan</td>
</tr>
<tr>
<td>FOUR</td>
<td>pedi</td>
<td>bidhi</td>
<td>pezi</td>
</tr>
<tr>
<td>FIVE</td>
<td>pangu</td>
<td>pongu</td>
<td>pfü</td>
</tr>
<tr>
<td>SIX</td>
<td>sarü</td>
<td>tsogho</td>
<td>tsaro</td>
</tr>
<tr>
<td>SEVEN</td>
<td>sinyi</td>
<td>tsini</td>
<td>tsanü</td>
</tr>
<tr>
<td>EIGHT</td>
<td>tiche</td>
<td>thache</td>
<td>tütse</td>
</tr>
<tr>
<td>NINE</td>
<td>tepfü</td>
<td>toku</td>
<td>tükhü</td>
</tr>
<tr>
<td>TEN</td>
<td>chiro</td>
<td>chiği</td>
<td>tsarü</td>
</tr>
</tbody>
</table>

- 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

<table>
<thead>
<tr>
<th>Idu (Talukdar)</th>
<th>Mishmi (Dubey)</th>
<th>Chulikata13</th>
<th>Taraon14</th>
<th>Deng Darang (Sun et al. 1980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>khe(ng)ge</td>
<td>e:khe:</td>
<td>khing</td>
<td>kʻun55</td>
</tr>
<tr>
<td>TWO</td>
<td>kanyi</td>
<td>ka:ni</td>
<td>ka:ing</td>
<td>kə21n55</td>
</tr>
<tr>
<td>THREE</td>
<td>kasō</td>
<td>ka:sh</td>
<td>kə:səng</td>
<td>kə21səung45</td>
</tr>
<tr>
<td>FOUR</td>
<td>kapri</td>
<td>ka:ppi</td>
<td>ka:pra:i</td>
<td>kə21pəəi55</td>
</tr>
<tr>
<td>FIVE</td>
<td>manga</td>
<td>ma:nga:</td>
<td>ma:nga:</td>
<td>ma:nga:45</td>
</tr>
</tbody>
</table>

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:

<table>
<thead>
<tr>
<th>Miju Mishmi (LSI)</th>
<th>Miju (Das Gupta 1977a)</th>
<th>Deng Geman (Sun et al. 1980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>kuo:/komo:</td>
<td>kumo</td>
</tr>
<tr>
<td>TWO</td>
<td>ka:ning/kinnin</td>
<td>kinin</td>
</tr>
<tr>
<td>THREE</td>
<td>ka:-sam</td>
<td>kə:səm</td>
</tr>
<tr>
<td>FOUR</td>
<td>kambrin</td>
<td>kambran</td>
</tr>
<tr>
<td>FIVE</td>
<td>ka-li:n</td>
<td>kə:li:n</td>
</tr>
<tr>
<td>SIX</td>
<td>kə:ta:m</td>
<td>katam</td>
</tr>
</tbody>
</table>

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:-səng, 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-

<table>
<thead>
<tr>
<th></th>
<th>Abor-Miri</th>
<th>Gallong</th>
<th>Minyong</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>a-ko; a-ter/-tel</td>
<td>ako/aken</td>
<td>akon; atir/ayirr</td>
</tr>
<tr>
<td>TWO</td>
<td>a-nyi</td>
<td>anyi</td>
<td>anyi</td>
</tr>
<tr>
<td>THREE</td>
<td>a-um/a-ngum</td>
<td>aum</td>
<td>aum</td>
</tr>
<tr>
<td>FOUR</td>
<td>a-pi</td>
<td>appi</td>
<td>aki15</td>
</tr>
<tr>
<td>FIVE</td>
<td>a-ngo</td>
<td>ango</td>
<td>ango</td>
</tr>
<tr>
<td>SIX</td>
<td>a-keng/a-ke’</td>
<td>akke</td>
<td>akeng/akkeng</td>
</tr>
</tbody>
</table>

(B) With vocalic variation in the prefix (1-6 run)

<table>
<thead>
<tr>
<th>Tagin</th>
<th>Lhopa</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Das Gupta 1975)</td>
</tr>
<tr>
<td>ONE</td>
<td>akin</td>
</tr>
<tr>
<td>TWO</td>
<td>anyi</td>
</tr>
<tr>
<td>THREE</td>
<td>aum</td>
</tr>
<tr>
<td>FOUR</td>
<td>epí</td>
</tr>
<tr>
<td>FIVE</td>
<td>ango</td>
</tr>
<tr>
<td>SIX</td>
<td>aké</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Padam</th>
<th>Nishi</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Dubey 1983)</td>
</tr>
<tr>
<td></td>
<td>akem/atel</td>
</tr>
<tr>
<td></td>
<td>(Das Gupta 1977b)</td>
</tr>
<tr>
<td>ONE</td>
<td>ani</td>
</tr>
<tr>
<td>TWO</td>
<td>aum</td>
</tr>
<tr>
<td>THREE</td>
<td>appi</td>
</tr>
<tr>
<td>FOUR</td>
<td>pilingo</td>
</tr>
<tr>
<td>FIVE</td>
<td>akke</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Dafla (LSI:Robinson)</th>
<th>E. Dafla (LSI:Hamilton)</th>
<th>Yano Dafla (N.L. Bor 1938)</th>
<th>Dafla (Das Gupta 1969)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>a:-kin</td>
<td>akkin</td>
<td>aking/aku</td>
</tr>
<tr>
<td>TWO</td>
<td>a:-ni</td>
<td>anyi</td>
<td>anyi/ain(^{17})</td>
</tr>
<tr>
<td>THREE</td>
<td>a:-a:m</td>
<td>a-om</td>
<td>um</td>
</tr>
<tr>
<td>FOUR</td>
<td>a:-pli</td>
<td>apli; appi</td>
<td>api</td>
</tr>
<tr>
<td>FIVE</td>
<td>a:-ngo:</td>
<td>a:-ng</td>
<td>ango</td>
</tr>
<tr>
<td>SIX</td>
<td>akple</td>
<td>a:-kr</td>
<td>aké</td>
</tr>
</tbody>
</table>

- 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-Mila ng-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).

<table>
<thead>
<tr>
<th>Aka (= Hruso) (LSI)</th>
<th>Apatani (Simon 1972)</th>
<th>Milang (Das Gupta 1980)</th>
<th>Serdukpen (Dubey 1983)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ONE</td>
<td>a</td>
<td>kū</td>
<td>akan; atel</td>
</tr>
<tr>
<td>TWO</td>
<td>kshi</td>
<td>nī</td>
<td>ne</td>
</tr>
<tr>
<td>THREE</td>
<td>zu</td>
<td>hī</td>
<td>ham</td>
</tr>
<tr>
<td>FOUR</td>
<td>fi-ri</td>
<td>pe</td>
<td>pe</td>
</tr>
<tr>
<td>FIVE</td>
<td>phum</td>
<td>ngo</td>
<td>pangu</td>
</tr>
</tbody>
</table>

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 ≠ tsiay (§3.2.2); TEN/HUNDRED *liy ≠ *lyag (§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 lè, 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 แก่น.) Roop (1970:62-63) reports a similar Lisu general classifier lyö̂̆̃, occurring only after lyi55 ‘4’, which he characterises as a ‘suppletive allomorph’ of the ordinary general classifier ma33. In the Qiangic
language Muya, a possibly cognate form $lo^{55}$ occurs as a citation classifier, but after all the numerals from 1 to 10 (for example $so^{55}-lo^{55} \cdot '3'$, $z\gamma^{35}-lo^{55} \cdot '4'$, $ga^{55}-lo^{55} \cdot '5'$, 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).
APPENDIX 1
METASTATIC FLOWCHART OF NUMERICAL SEMANTIC ASSOCIATIONS

(chart by Jonathan P. Evans)
APPENDIX 2
INDEX OF RECONSTRUCTED ETYMA

ONE
*it ≠ *yat §3.1.1
*ʔét (OC) §3.1.1
*k-(y)at ≠ *k-(y)it ≠*k-yan ≠*k-(y)in
§3.1.2
*kat §3.1.2
*khat (Kuki-Naga) §3.1.2
*kya-n ≠*kya-t §3.1.2.1
*ka and *ko §3.1.3
*kon (PTani) §3.1.3
*g-t(y)i-k ≠*tya-k §3.1.4
*d/lay ≠*d/tan §3.1.4
*tšjāk (OC) §3.1.4
*tān ‘single, simple’ (OC) §3.1.4
*day² ‘one’ (PLB) §3.1.4
*nday ‘all’ (PLB) §3.1.4
*ʔdik ‘only’ (PLB) §3.1.4
*dek ‘nothing’ (PLB) §3.1.4
*ʔgyik ‘little bit’ (PLB) §3.1.4
*-*kla (PNN) §3.1.4
*tir ≠*tur (AMD) §3.1.5.1
*tel (PEasternTani) §3.1.5.1
*tel (AMD) §3.1.5.1
*(t)se (Kamarupan) §3.1.5.2
*-*tse (PNN) §3.1.5.2
*sa or *tsa (Kamarupan) §3.1.5.2
*han or *hay (AMD) §3.1.5.3
*a (AMD) §3.1.5.4
*(k-)JV(N) (AMD) §3.1.5.5
*d’uk ‘alone; only’ (OC) §3.1.5.5

TWO
*g-ni-s/k §4.1.1
*ni-k §4.1.1.4

THREE
*?ni-t, *?ni-? (PLB) §4.1.1
*ʔi (PTani) §4.1.1.2
*ʔ-ʔi (PNN) §4.1.1.2
*njær (OC) §1.2.6
*g-g-nis (pre-Garo) §5.2

FOUR
*b-ly = *b-lay §4.1.3
*pri (PTani) §4.1.3
*bālay (PNN) §4.1.3
*hλay² (PLB) §4.1.3.4.1
*m-lay < *b-lay §4.1.3.2
*g-lay §4.1.3.6
*g-b-lay §4.1.3.6
*g(N)-b-lay-(N) §4.1.3.6
*śəd (OC) §1.2.6
*k-b-lay §5.2

FIVE
*l-ga ≠ *b-ŋa §4.1.4
*b-l-ŋa §4.1.4.4
*m-ŋa §3.5.2.2
*r-ŋa (‘Old Kuki’) §4.1.4
*ŋa² (PLB) §4.1.4.5
*g-l-ŋa1 (AMD) §4.1.4.7
*ŋo (OC) §1.2.6
*k-m-ŋa < *b-ŋa (Proto Rgyalrong) §5.2

122
*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
*ljôk (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-nis §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(j-i-t) (PLB) §4.2.2
*si² (PLB) §4.2.2.4
*hnas (PKaren) §4.2.2.3
*kv-nut (PTani) §4.2.2.5
*ts'îët (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-îni (PTani) §4.2.3.8; §5.1.3.1
*gyat §4.233; §4.2.3.5
*pwât (OC) §1.2.6; §4.2.3
*d-ryat < *g-ryat (PKN) §4.2.3
*b-ryat (Gurung-Tamang-Thakali) §4.2.3.4
*rit ≈ *ryat (PLB) §3.5.4.4; §4.1.1.4; §4.2.3; §4.2.3.4

NINE
*d-kəw (= *d-kuw) ≈ *s-gəw ≈ *d-gaw §4.2.4
*d-gaw ≈ *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)wə-gə (AMD et al.) §4.2.4.5
*kV-(n)ə (PTani) §4.2.4.5
*kjug (OC) §1.2.6

TEN
*gip ≈ *gyap §3.2.1
*ts(y)iy ≈ *tsyay §3.2.2
*syṣye < *s-tsyiy (Proto Rgyalrong) §3.2.2
*tși’ (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
*rjuŋ (PTani) §3.2.3.1
*pal or *bal §3.2.3.2
*bon (PNN) §3.2.3.2
*s-r/lin ≈ *s-r/lyañ (AMD et al.) §3.2.3.3
*riŋ ≈ *yìn §3.2.3.3
*liŋ ≈ *lyañ ‘ten/hundred’ §3.2.3.3
*s-ryak ≈ *s-rwak §3.2.3.4
*roik (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
*_dpøj (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
HUNDRED
*b-r-γya ≠ *b-g-rya §3.5.4.5
*b-rya §3.5.4.5
*m-rya (Naga) §3.5.4.5
*hra’ (PLB) §3.5.4.4
*m-liγya (AMD) §3.5.4.6
*m-li(ŋ) (? < *m-l-ŋya) (AMD) §3.5.4.6

THOUSAND
*s-τoŋ §3.5.4.7
*s-riŋ ≠ *s-ranγ §3.5.4.7
*gheslo- (PIE) §3.5.4.7.2

*pāk (OC) §1.2.6
APPENDIX 3
INDEX OF LANGUAGES AND SOURCES

Abor-Miri-Dafla
Bogaer Luoba: TBL.
Bokar Adi: ZMYYC.
Chulikata (Taying Mishmi): LSI III.1:623.
Deng (Kaman, Geman): Sun Hongkai et al. 1980. ZMYYC. TBL.
Hrusso (Aka): LSI III.1, 622-623.
Idu: Talukdar et al. 1962:15. ZMYYC.
Idu Luoba: TBL.
Mishing: Dubey 1983.
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
Dimasa: Marrison 1967.

Himalayish
Kanawari: Joshi 1909:2-3.
Lepcha: Mainwaring/Grünwedel 1898.
Sikkim Bhutia (Dzongkha): Sandberg 1895:59.
Tibetan (Written): Jäschke 1881.

Jingpho-Nungish
Anong Nu: TBL. ZMYYC.
Dulung (Trung): Sun Hongkai 1982b:54. ZMYYC. TBL.
Jingpho (Kachin): Hanson 1906/1954. Maran (in prep.).

Karenic
Palaychi: Jones 1961.
Kelun: TBL.

Kuki-Chin-Naga
Bawm (=Laizo): Osborne 1975.
Chang: Marrison 1967.
Gangte: Bible Society of India 1972a:512.
Lakher (=Mara): Lorrain 1951.
Lotha: Marrison 1967.
Lushai: Marrison 1967.
Mao: Marrison 1967.
Maring: Marrison 1967.
Meluri: Marrison 1967.
Nocte: Marrison 1967.
Phom: Marrison 1967.
Tangsa (Muklom): Dubey 1983.
Thado: Thirumalai 1972.
Yimchungru: Marrison 1967.

**Lolo-Burmese**

Achang: ZMYYC. TBL.
Ahi: Yuan Jiahua 1953.
Akha: Lewis 1968.
Anong: ZMYYC.

Bola: TBL.
Burmese (spoken): ZMYYC. TBL.
Gazhuo: TBL.
Hani (Caiyuan = Biyue): ZMYYC.
Hani (Dazhai): ZMYYC.
Hani (Lüchun): TBL.
Hani (Mojiang): TBL.
Hani (Shuikui = Haoni) ZMYYC.
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.
Naxi (Lijiang): ZMYYC.
Naxi (Yongning = Moso): ZMYYC.
Naxi: Rock 1963. TBL.
Nusu: ZMYYC.
Sani: Ma Xueliang 1951:81. TBL.
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.
    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 (Zhábà): ZMYYC. TBL.
    ZMYYC. TBL.
Shixing: ZMYYC. TBL.
Zhábà: ZMYYC. TBL.

Other
Sulong (Sulung): ZMYYC.
Tujia: ZMYYC. TBL.
REFERENCES


Hamilton, R.C., 1900, An outline grammar of the Dafla language, as spoken by the tribes immediately south of the Apa Tanang country. Shillong.


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


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


1990a, On megalocomparison. Language 66/1:106-120.


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


In preparation, Languages of mainland Southeast Asia. Cambridge University Press.


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


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


Robinson, W., 1851, Notes on the Dophlās and the peculiarities of their language. JASB 20:126-137.


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.


1984, Kayah numerals. MS.

In preparation, A grammar of eastern Kayah. (Will include a collection of texts and a glossary.)


Taid, Tabu, 1993, Mising wordlist. MS.


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.

