PAPERS IN SOUTHEAST ASIAN LINGUISTICS NO. 14:

TIBETO-BURMAN LANGUAGES OF THE HIMALAYAS

edited by

David Bradley

Department of Linguistics
Research School of Pacific and Asian Studies
THE AUSTRALIAN NATIONAL UNIVERSITY

TABLE OF CONTENTS

INTRODUCTION

This volume contains eight papers on the Tibeto-Burman and Sino-Tibetan languages of the Himalayan region.

The first, 'Tibeto-Burman languages and classification', by David Bradley, is a general overview of the genetic classification and previous studies of these languages, which evaluates the various competing claims on the subgrouping of Tibeto-Burman (TB), and provides a new classification. This is followed by an up-to-date summary of the locations, names and populations of groups speaking these languages. Also included are references to recent and forthcoming linguistic studies.

The remaining papers are arranged geographically from west to east, an arrangement which corresponds to the genetic classification proposed. Four are concerned mainly with verbal morphosyntax, synchronic and diachronic. Of the others, one deals with phonesthetics, one with the syntax and semantics of possessives, and one with development and change within an orthography.

The West Himalayish languages of Lahul and Kinnaur in north-western India were among the earliest Tibeto-Burman languages of the Himalayas to receive attention from scholars, in the mid-nineteenth century. Little further study was undertaken until quite recently, but now Anju Saxena has done so in her paper 'Towards a reconstruction of the Proto West Himalayish agreement system'. This paper presents a summary of her own and other recent descriptive studies on verb morphology and a proposed reconstruction. The data and results are extremely useful to compare with the reconstructions of Kiranti verbal morphology by van Driem and others, and to provide further data for the current debate by various scholars including DeLancey (1989), LaPolla (1992) and van Driem (1993) concerning the status of verbal morphology in Proto TB.

Ross Caughley, whose previous studies such as Caughley (1982) have also contributed very substantially to this debate, takes up another fascinating topic in his contribution, 'Semantically related vowel gradation in Sunwar and Chepang'. Sound symbolism is quite widespread in the South and Southeast Asian linguistic areas, but most published descriptions relate to Austroasiatic languages. Caughley shows that similar phenomena exist in TB languages, as indeed they also do in Indic and Tai languages. His study shows the presence of semantically based vowel alternations in two sections of TB, both in Chepang and in Sunwar, which forms part of the Kiranti or Rai group.

The following four studies all describe other members of the Kiranti or Rai group. Gerard Tolsma's article 'The verbal morphology of Kulung' provides a thorough summary on this topic for one of the western Kiranti languages.

The paper by Novel Kishore Rai and Werner Winter, 'Triplicated verbal adjuncts in Bantawa', like Caughley's contribution, discusses the continuum of forms extending towards sound symbolism. The paper provides extensive data on the morphosyntactic and phonological characteristics and sources of triplicated verbal adjuncts, including some nominal stems, more verbal stems, and a much larger group of forms not attested except when triplicated. Bantawa is a central Kiranti language with a substantial number of speakers and internal dialect diversity, and has replaced or is influencing some of the smaller adjacent languages; it was formerly used as a Rai lingua franca in various areas including even the Darjeeling tea plantations. This paper has previously appeared elsewhere but with very restricted distribution and a number of misprints; thus it has been republished here.

Balthasar Bickel presents a convincing summary of the syntax and semantics of possession in his paper 'The possessive of experience in Belhare'. This language, as noted elsewhere, was newly discovered by the Linguistic Survey of Nepal; due to Bickel's work it is becoming one of the better described Kiranti languages.

In his paper 'A new analysis of the Limbu verb', George van Driem continues to extend his descriptive and comparative studies of Kiranti morphosyntax. Like Sunwar in the west, Limbu and a few other languages at the eastern end of the Kiranti/Rai area are often not included within the Rai group, though linguistically they form part of it. Limbu is unusual among Kiranti/Rai languages in several ways: the large number and wide distribution of speakers; the existence of a traditional Indic-derived orthography used by a literary elite and to some degree in education; the degree of diversity among its dialects; and the number of recent descriptive studies of these dialects, most notably by van Driem himself (see the reference in the van Driem paper in this volume).

Another traditional orthography of the area is that of Lepcha, the autochthonous group of the Darjeeling/Sikkim/south-western Bhutan area. R.K. Sprigg's contribution 'Lepcha orthography: an earlier and a later stage' presents a thorough summary of the changes within this orthography from the earliest records in 1828 to the present. This continues a long series of studies by the same author of the development of and connections between orthographies of South and Southeast Asia. Since Lepcha is regarded by many scholars as a key language in the reconstruction of Proto TB, Sprigg's contribution is particularly timely and will enable scholars to use the recent dictionary by Tamsang (1994) and earlier studies in a fully informed way.

Of course a great deal of descriptive and comparative work remains to be done on the TB languages of this area. The scholars whose papers are included in this volume, among many others, are carrying our understanding of TB forward rapidly. I hope that this volume will assist in the process.

REFERENCES

Caughley, Ross C., 1982, *The syntax and morphology of the verb in Chepang. PL*, B-84.
DeLancey, Scott, 1989, Verb agreement in Proto-Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* LII/2:315-333.

van Driem, George, 1993, The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* LVI/2:292-334.

La Polla, Randy, 1992, On the dating and nature of verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* LV/2:298-315.

Tamsang, Kharpoo, 1994, *Lepcha-English encyclopedic dictionary*. Kalimpong: Mani Press.

DAVID BRADLEY
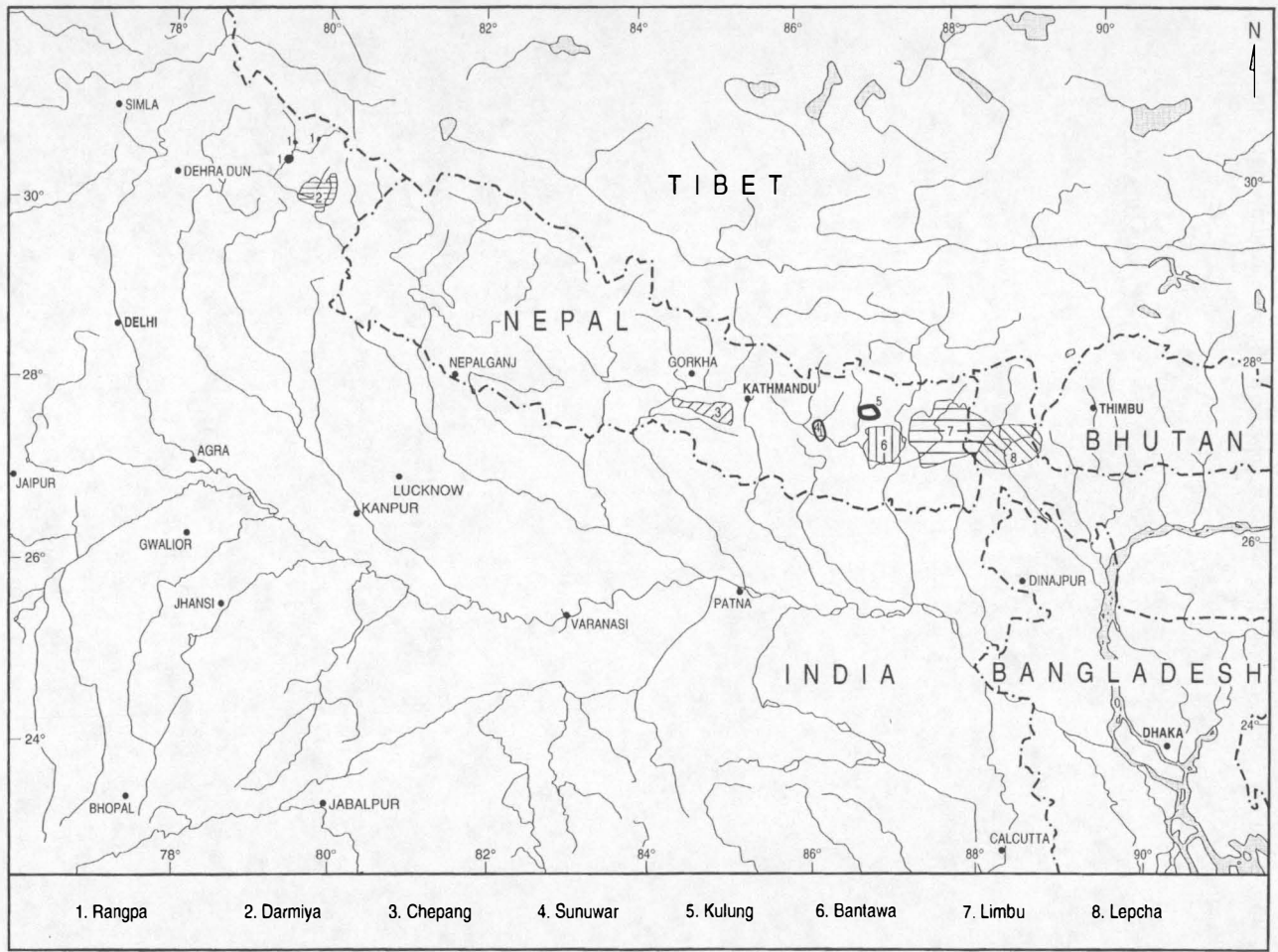
MAP 1: LANGUAGES IN THIS VOLUME

1. Rangpa   2. Darmiya   3. Chepang   4. Sunuwar   5. Kulung   6. Bantawa   7. Limbu   8. Lepcha

TIBETO-BURMAN LANGUAGES AND CLASSIFICATION

DAVID BRADLEY

1. INTRODUCTION

The Tibeto-Burman (TB) languages are the principal languages of the Himalayan region, spoken from Kashmir in the west, across the Himalayan and sub-Himalayan regions of India, Nepal, Bhutan, Bangladesh, Tibet and China, and into Southeast Asia across Burma, Thailand, Laos and Vietnam. There are several hundred languages known, and doubtless some others yet to be identified.

Various maps of the TB languages have been produced recently. For those in Southeast Asia, see Wurm and Hattori (1981/1983); concerning China, see Wurm et al. (1987/1991). A relatively comprehensive picture of TB and other languages in Burma is given in Moseley (1994). The TB languages which are used as languages of wider communication are discussed and mapped in Wurm, Mühlhäusler and Tryon (forthcoming), and those which are dying are found in Wurm (1996). Van Driem (forthcoming) deals with the languages of Nepal, and van Driem (1992) those of Bhutan. However, detailed maps of the TB languages of the rest of South Asia are provided here for the first time, as well as an appendix showing the names and approximate 1995 populations of all TB groups.

Hale (1982) is a recent and extremely useful bibliographical summary for all TB languages. A fuller bibliography of linguistic studies of all the TB and other languages of Nepal by Toba (1991) provides an excellent resource for languages of this area, including some which extend outside Nepal. In the discussion and references below, the major recent studies of TB languages of the Himalayan region, especially in South Asia, are cited. Two excellent recent compilation volumes for the TB languages of China, Sun (1991) and Dai (1992), have provided further data on some of these as well as many other TB languages.

There are two main classifications of TB languages in use; one is that of Shafer (1974), which splits it into four main parts: Bodic, Baric, Burmic and Karenic; the other is that of Benedict (1972), with eight subgroups plus an 'other' category. These nine subgroups have been reclassified by Benedict (1976) into three groups. For a comparison of these and earlier classifications, see Hale (1982).

With additional data on languages of China and north-eastern India, it has become clear that some revisions are needed; specifically, some of Benedict's 'other' languages, classified tentatively as Burmic by Shafer, form an additional group called the Qiang group by Sun (1983b), some of the Rong languages by Thurgood (1985), and here the North-eastern TB

1

languages. Shafer's classification of a number of specific languages and language groups has proven doubtful; for example, most languages of his Burmic group are here included in the Burmese-Lolo or Kuki-Chin subgroups of the South-eastern TB group, but some others such as Jinghpaw are in the North-eastern India TB group and others such as Qiang are in the North-eastern TB group.

The classification of some languages is uncertain, in most cases due to very extensive contact: Bai and Tujia with Chinese, the Nungish languages with Burmese-Lolo languages, Lepcha with a variety of TB languages and so on. It appears that the overall pattern can be summarised as below:

```
                        Tibeto-Burman

        North-eastern India

                Western

                                 South-eastern    North-eastern
```
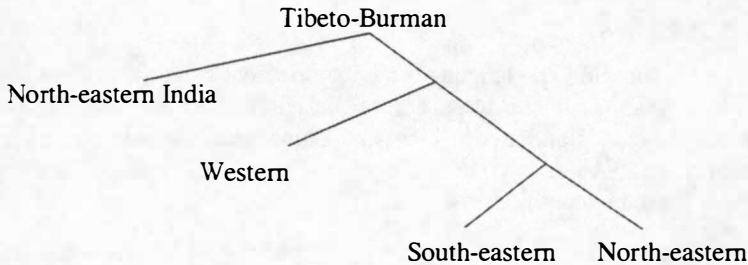
The North-eastern India group includes Shafer's Baric group, also known as Benedict's Bodo-Garo-Konyak group, now usually known as the Sal group from a name suggested by Burling (1983b); plus the Jinghpaw (Kachin) and Sak or Luish group; it has some lexical peculiarities not shared with other TB languages; most languages have some morphology including parts of that reconstructed for Proto TB by Wolfenden (1929), but relatively simple tonal systems. This group and others have been linked geographically by Matisoff, in his Kamarupan group.

The Western group corresponds mainly to Shafer's Bodic group and Benedict's Tibetan/Kanauri plus Himalayan, with a few additions. In this classification the Western group is divided into Bodic (including Tibetan), and Himalayan. Bodic has four subgroups: Tibetan proper; Western Bodish (Gurung or Tamang) subgroup, Eastern Bodish or Monpa, and eastern Bodic or Tshangla subgroup, and the Kanauri subgroup, also known as West Himalayish. Himalayan falls into the relatively homogeneous Kiranti or Rai subgroup and the more disparate western subgroup, which includes various languages not classified as Bodic by Shafer, including Newari and Chepang. Most of these languages have extensive secondary morphology, especially on verbs; many have word rather than syllable tone systems, often involving phonation as well, while some are non-tonal.

The North-eastern India group includes the Central subgroup (languages of the border between north-eastern India and Tibet, also northern Burma and adjacent areas of China); some scholars such as Thurgood have linked this with the core North-eastern or Qiang subgroup. The latter languages have substantial shared verb morphology; most are tonal. Some of the southern North-eastern group languages are lexically transitional to South-eastern, but are phonologically and morphologically more typical of North-eastern. These include the Nungish and Naxi languages.

The South-eastern group includes Shafer's Kukish/Benedict's Kuki-Chin-(Southern) Naga; the Burmese-Lolo subgroup; and the Karen subgroup. Apart from Kuki-Chin, which

is typologically similar to the adjacent North-eastern India languages in having some preserved and some innovative morphology and relatively simple tonal systems, these languages tend to have very little morphology and complex tonal systems typical of the northern South-east and southern East Asian linguistic area. Of all the TB languages, the Karen subgroup is the only one to have SVO syntax; but even Karen retains various verb-final syntactic characteristics. Bai also shows SVO as an alternative possibility; it has been heavily influenced for millennia by SVO Chinese. All other TB languages are SOV.

2. WESTERN TB OR BODIC

This group comprises two main branches: Tibetan and other closely related languages on the one hand, and the TB languages south of the main Himalayan range, from north-western India across Nepal and Sikkim. The relationship between these languages can be shown as follows.

WESTERN TB

HIMALAYAN

WEST BODISH (Gurung)   CENTRAL BODISH (Tibetan)   EAST BODISH (Bumthang   TSHANGLA   WEST HIMALAYAN   CENTRAL   KIRANTI

2.1 BODISH

This subgroup includes Tibetan proper. Apart from literary Tibetan with its long history and continuing use as the liturgical language of a variety of Buddhism, there are many regional and local varieties which serve as the spoken Low corresponding to the literary High in a diglossic pattern. Some of these regional Lows have more widespread use; this includes the Low of Lhasa, the traditional capital of Tibet, which is used as a spoken lingua franca among Tibetans in South Asia and elsewhere. In much of central and eastern Tibet a variety of Kham spoken Tibetan is used as a lingua franca by herdsmen, and is thus known as Brokpa, literally 'herdsman'.

Many TB groups on the margins of Tibet have adopted Tibetan Buddhism, and some of these speak languages which are also Bodish, while others speak North-eastern TB languages. Those in more intimate contact with Tibetans may tend to become Tibetanised, with their languages showing this process. One such example is Baima, the northernmost North-eastern Tibetan language, which some scholars prefer to regard as a variety of Tibetan; see for example Nishida and Sun (1990) who support this view.

There are two non-Tibetan subgroups of Bodish. One is West Bodish, also sometimes known as the Gurung group or the Tamang group from the two languages with the largest number of speakers. This includes Gurung, Tamang, Thakali, Manang, Kaike and Ghale; the last two are close to Tibetan linguistically, and thus perhaps transitional between West Bodish and Tibetan. The other is East Bodish, which includes the large Tshangla group.
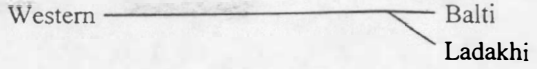
2.1.1 CENTRAL BODISH (TIBETAN)

Some sources suggest as many as 6.5 million speakers of Tibetan, but this is an overestimate; the actual total is probably less than five million. Also very close to Tibetan but linguistically distinct are the Monpa group of eastern Bhutan and adjacent areas of India and Tibet to the east. Within 'Tibetan' itself there is a vast range of varieties, nearly all of which are linked by sharing Tibetan Buddhism and thus literary Tibetan as a koine. Scholars tend to divide this range into Western, Central, Southern, Amdo (mostly north-eastern) and Khams (mostly eastern) subgroups. In India and Nepal most Tibetans are pejoratively called Bhotia, and in China they are called Zang [tsaŋ^{51}]. Apart from its role as the language of Tibet, varieties of Tibetan are or were the official language of various kingdoms, from Ladakh in the west to Mustang in north central Nepal, Sikkim, Bhutan and so on.

The Tibetan-speaking area has been divided among various provinces in China, including the 'Tibetan Autonomous Region', north-western Yunnan, western Sichuan, much of Qinghai and parts of Gansu. Tibetan Buddhist influence formerly spread even further, into the Mongol areas to the north-east. The events of 1959 added substantial numbers of 'standard' Lhasa speakers to India and Nepal as refugees. Closure of the Indian-Tibetan border in 1965 further disrupted traditional trading relationships and left small 'Bhotia' populations speaking divergent Central and Southern Tibetan varieties cut off. In Nepal, the northern quarter of the country is inhabited mainly by Central and Southern Tibetan speakers, with large post-1959 refugee groups around Kathmandu and elsewhere. Apart from the speakers of Nepali, nearly all the population of Bhutan and Sikkim speak some variety of Tibetan or Monpa. Tibetan is also spoken by a few in northernmost Burma. A conservative total is nearly 4.9 million speakers. This total does not include the West Bodish (Gurung, Tamang) Group with nearly 800,000 and the East Bodish, Bumthang or Monpa Group with nearly 150,000 speakers. More distant are the Tshangla Group, three languages with nearly 150,000 speakers in Bhutan. Not included are the rGyarung and several other groups of western Sichuan who are within the cultural orbit of Tibetan Buddhism but speak distinct languages of the North-eastern TB Group. These latter groups, however, are included within the Tibetan nationality in China; hence the total population of the Tibetan nationality in China, 4.6 million in 1990, is higher than the number of Tibetan speakers there, approximately 4.3 million.

I am grateful to George van Driem and Warren W. Glover for some of the information in this section. It should be noted that the names of many Tibetan varieties include the suffix *-pa/-po* 'people' or the suffix *-sKad/-kad/-kat* or *-ka/-kha* 'speech'. The following indicates the names and relationships within Central Bodish.

Central Bodish/Tibetan   Western ——————————— Balti

Ladakhi

Central   Western

Tod
Ranglo
Spiti
mNyam
Jad

mNgahris (Ali)

Ruthog
sGar
rTsamdah
sPuhreng
mTshochen
dGergyas
Humli Tamang
Khan
Karmarong

gTsang

Dolpo
Reng Pungmo
Tichurong
bKag
gLo
Nar
Gyasumdo
Nubri
Tsum
'Kachad'
Langtang
Sherpa
Kagate
Jirel
Lhomi
Halung
sKyidgrong
Dingri
gZhiskartse
rGyalrtse
sNadkarrtse

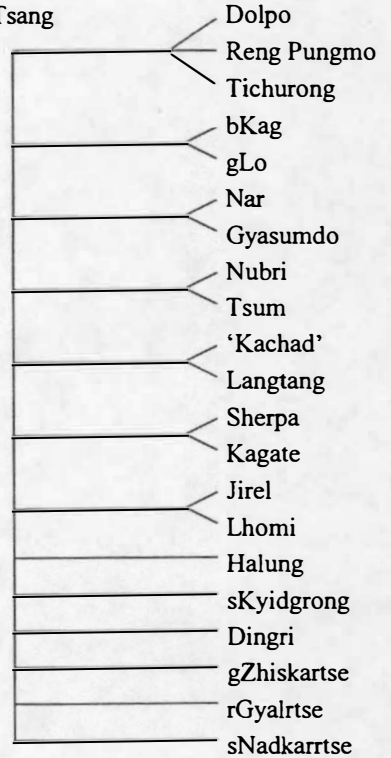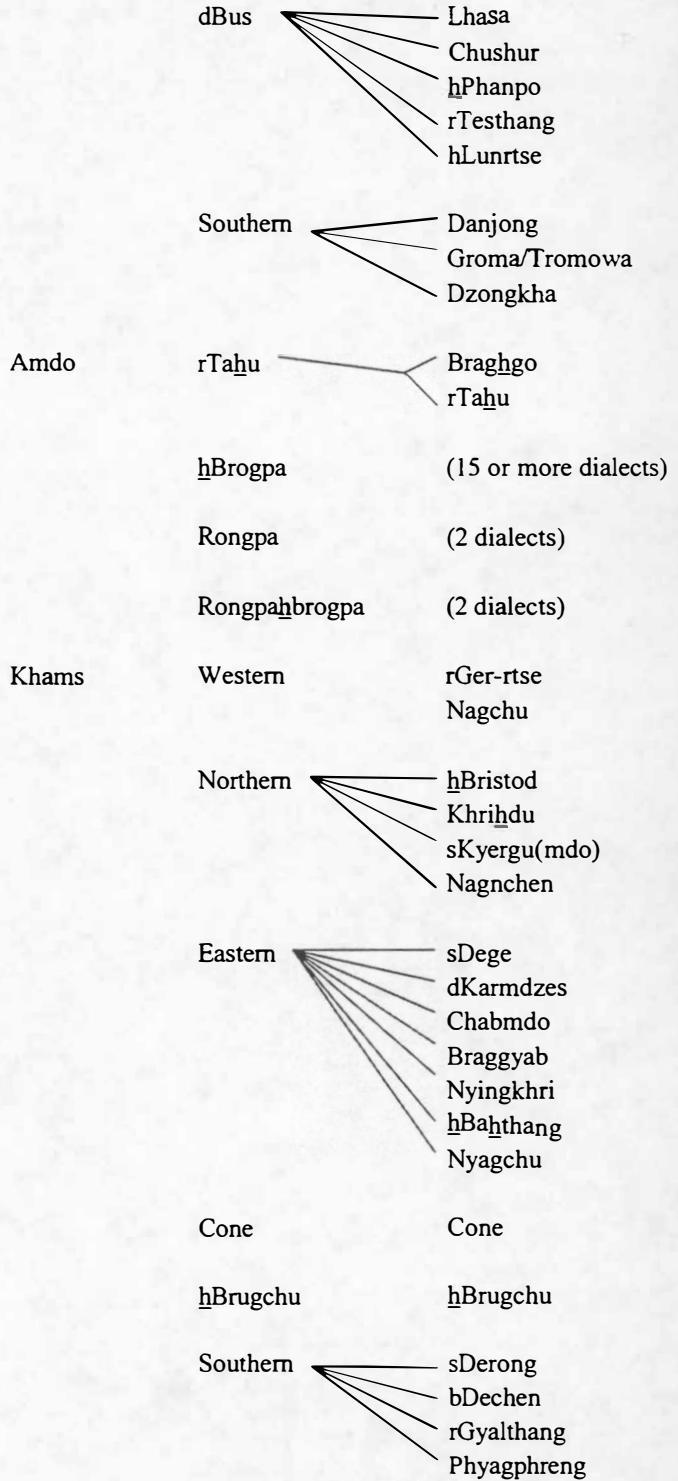|		
	dBus	Lhasa
		Chushur
		hPhanpo
		rTesthang
		hLunrtse
	Southern	Danjong
		Groma/Tromowa
		Dzongkha
Amdo	rTahu	Braghgo
		rTahu
	hBrogpa	(15 or more dialects)
	Rongpa	(2 dialects)
	Rongpahbrogpa	(2 dialects)
Khams	Western	rGer-rtse
		Nagchu
	Northern	hBristod
		Khrihdu
		sKyergu(mdo)
		Nagnchen
	Eastern	sDege
		dKarmdzes
		Chabmdo
		Braggyab
		Nyingkhri
		hBahthang
		Nyagchu
	Cone	Cone
	hBrugchu	hBrugchu
	Southern	sDerong
		bDechen
		rGyalthang
		Phyagphreng

2.1.1.1 WESTERN TIBETAN

These languages are spoken in Kashmir, and are non-tonal. They are in various other ways phonologically conservative, for example reflecting the written Tibetan b-, d-, g-, r- and l- prefixes directly.

Balti is spoken by a Moslem population of over 300,000 in the Baltistan area of northern Kashmir around Skardu and Khappalu, and written in an Arabic script. In the Indian-administered area around Kargil is the Purik dialect of Balti. There are about 45,000 speakers in the Indian-controlled area, and over 250,000 in the Pakistan-controlled area.

Ladakhi is the local Tibetan variety of the Ladakh area around the town of Leh in Indian-administered north-eastern Kashmir; it is also known as Ladwags from its local name. The most recent description is Koshal (1979). As the lingua franca of the area, it is also spoken by some Sh(r)ina (Dardic) speakers and others. As a first language, it has about 75,000 speakers. In the Zangskar valley to the south-west there is a distinct dialect with about 5,000 speakers.

2.1.1.2 CENTRAL TIBETAN

2.1.1.2.1 WESTERN SUBGROUP

Some of these varieties, spoken by about 25,000 people in India, are included in the Western Group by other scholars; this includes the varieties of Tibetan spoken in Lahul, sometimes incorrectly lumped into a category 'Lahuli' along with several West Himalayish languages of the lower valleys in the same area. Nishi (1986) calls these 'Western Innovative' and hence suggests a closer grouping with Ladakhi, Purik and Balti to the north-west.

Several closely related varieties of Tibetan locally known as Tod and Ranglo are spoken in Lahul, and are sometimes also called 'Lahuli'; these should be distinguished from the other TB languages of Lahul, Pattani, Tinan and Bunan. Tod (Tibetan for 'upper') is spoken in the Tod valley by about 1,700; and Ranglo (also sometimes known as Khoksar from the name of the largest village) to the north of the Rohtang Pass by some 1,000 in four villages. These and other varieties of Tibetan indigenous to northern India are often classified as Bhotia in census and other Indian sources.

Known to its over 12,000 speakers as Piti from the local placename Spiti, mNyam is spoken in the Spiti valley. To its south in Kanaur is mNyamskad or Nyamkat, with fewer speakers. Spiti can be divided into four dialects, Tod ('upper'), Bhar, Pin and Sham.

The Jad variety of Tibetan is underenumerated; census data suggest only a few hundred 'Bhotia' speakers in the area, but this is incorrect. They are locally known as Garhwal Bhotia, but should not be confused with the Rangpa or Marchha people to the east, who are also sometimes classified as 'Bhotia' but speak the West Himalayish language Rangkhas.

The mNgahris subgroup of Central Tibetan is mostly found in Ngari (Chinese Ali) Prefecture in western Tibet. It also extends into north-western Nepal, where it includes the so-called 'Humli Tamang' (who of course are not Tamang), Khan, and Karmarong varieties. There are seven varieties in Tibet, named from the principal towns of each: Ruthog, sGar (Gartok), rTsamdah, sPuhreng, mTshochen and dGergyas are six varieties, with a seventh

in the western part of adjacent Shigatse Prefecture. However, the north-eastern part of Ngari is inhabited by speakers of a Western Khams variety, sGer-rtse. For more details on these varieties, see Qu and Tan (1983); the total number of speakers is about 40,000.

2.1.1.2.2 GTSANG

This subgroup of Central Tibetan includes most of the Tibetan varieties in northern Nepal as well as those of the large towns of Shigatse and Gyantse. There is more detail available on the subvarieties of Nepal. The total 'Bhote' population of Nepal in 1981 was enumerated at about 74,000; this figure is too low, as the Sherpa total alone for Nepal is nearly 50,000. Most other varieties have a few thousand speakers each, for a total of about another 50,000, as well as some 460,000 in China or over 560,000 in total.

In some valleys of northern central Nepal, there are various gTsang varieties spoken. The Dolpo, Reng Pungmo and Tichurong varieties are sometimes collectively known as Dolpo, and are spoken in the Dolpo region of north-western central Nepal. The bKag and gLo varieties are spoken in the Mustang area; gLo or Lo in Mustang itself, and bKag or Kag to the south. The four varieties known (from west to east) as Nar, Gyasumdo, Nubri and Tsum are spoken to the north of the Gurung area. Webster (1992) surveyed the eastern part of this area and reports a population of about 3,200 Nubri and about the same number of Tsum speakers. Kachad and Langtang 'Bhotia' are spoken to the north of Kathmandu.

The Sherpa group of over 70,000 with some 50,000 speakers is well known for its mountaineering exploits. It is found mostly in north-eastern Nepal but also in China (about 800) and in the Darjeeling area of India (about 20,000, but many do not speak Sherpa). Its name means 'eastern people'. Closely related is Kagate, spoken to the south of the eastern Sherpa area by a small group.

The Jirel, Lhomi and Halung varieties are spoken in north-eastern Nepal. Jirel lies between the two main Sherpa areas and is spoken by about 3,000; Lhomi (hLomi, Kath Bhotia) is spoken by over 4,000 to the east of Sherpa; and Halung is spoken further east, north of the Limbu. See Vesalainen and Vesalainen (1980). Jirel and Lhomi are particularly similar.

2.1.1.2.3 GTSANG VARIETIES OF TIBET

In south-western Tibet various local varieties are spoken, including sKyidgrong, Dingri, gZhiskartse (Shigatse), rGyalrtse (Gyantse) and sNadkarrtse. In China, the total number of speakers is nearly 460,000.

2.1.1.2.4 DBUS

This is 'central' Central Tibetan, with about 900,000 speakers including most Tibetan refugees. It includes Lhasa (hLasa), Chushur, hPhanpo, rTsethang and hLunrtse varieties, among others, and extends into north-western and north-eastern Bhutan with about 50,000 speakers there. There are over 150,000 speakers among the various Tibetan refugee communities around the world, mainly in India and Nepal.

The local speech of Lhasa is the 'standard' variety of Tibetan spoken in and around the traditional capital; it is more generally known as 'central' dBus (transliterated from the written name) or Ü ([y], from the spoken version of the same word). It has diglossia between a literary variety used mostly for religious purposes but also for other reading-related activities, and a spoken variety. The literary diglossic 'High', which is phonologically conservative, is used throughout the Tibetan Buddhist area, but the spoken Lhasa variety was mainly limited to the central area until it was spread by education. Educated people from other areas of Tibet traditionally retained their local variety and learned the literary variety, which also served as a koine. The Lhasa variety has about 450,000 first-variety speakers, including most of the nearly 150,000 Tibetan refugees in Nepal, India and most Western countries. A classic description of this is found in Yu and Chao (1930); for more recent spoken materials, see Chang and Shefts (1978-1981).

2.1.1.2.4.1 SOUTHERN TIBETAN

These groups speak a slightly divergent type of Central Tibetan; most of the approximately 500,000 speakers live in Sikkim and Bhutan.

Tromowa or Gromo is the speech of the Chumbi valley between Sikkim and Bhutan. Danjong(ka), the language of the Sikkimese court, is spoken by over 70,000 people in Sikkim and adjacent areas and used in education in Sikkim; a slightly modified version of the Tibetan script is in use.

Dzongkha, the language of the fort (dzong), also sometimes known as Drukpa, is the Tibetan variety of the western third of Bhutan, with about 225,000 first-language speakers. For details of Tibetan varieties in Bhutan, see van Driem (1992). Recent language policy activities have developed a slightly modified Tibetan script for Dzongkha, and have spread Dzongkha as the national language throughout Bhutan, to speakers of Bumthang and Tshangla among other languages.

2.1.1.3 AMDO (NORTH-EASTERN) TIBETAN

These varieties are archaic in some ways, for example in the retention of nasal prefixes including h- (known in the Tibetan orthography as h-chung). Included is the frequently cited Golok (mGolog) variety among many others. Chinese sources divide Amdo into four subgroups: hBrogpa (nearly 540,000), Rongba (nearly 100,000), Rongba-hBrogpa (113,000) and the isolated southern rTahu (60,000); they suggest a total of about 810,000 speakers, which may be underenumerated. Within rTahu, Nishi (1986) identifies Braghgo and rTahu varieties. Various sources give fifteen or more named varieties within hBrogpa. Rongba includes two or more varieties, and Rongba-hBrogpa is transitional between them, with two or more varieties. One variety is described in Sun (1986).

2.1.1.4 KHAMS (SOUTH-EASTERN) TIBETAN

Not all scholars agree with the Chinese in attributing rGer-rtse (spoken in north-eastern Ngari Prefecture in western Tibet) to the Khams group, but if this is accepted then Khams is the most widespread subtype of Tibetan, extending from fairly far west to the furthest east

and south of all Tibetan traditional territory. Chinese sources suggest nearly 1.5 million speakers; there are also eight villages in northernmost Burma. Khams is also the lingua franca of Tibetan herdsmen over an even wider area.

A Western Khams variety known as rGer-rtse is spoken in Ngari Prefecture, with very sparse population in a band to the north-east and extending to the north of almost the entire Central Tibetan area, with Nagchu among other varieties included. These varieties account for about 160,000 speakers.

The largest group of Khams varieties is Eastern Khams, with nearly 960,000 speakers; this includes at least ten varieties including sDege, dKarmdzes, Chabmdo, Braggyab, Nyingkhri, hBahthang, Nyagchu and others. Northern Khams has about 91,000 speakers in a sparsely settled region; this includes hBristod, Khrihdu, sKyergu(mdo) and Nangchen varieties. All of the Tibetans in Yunnan Province and some of those in south-westernmost Sichuan Province speak Southern Khams varieties; speakers number about 135,000, with varieties being sDerong, bDechen, rGyalthang and Phyaphreng.

The Cone variety of Khams, with about 77,000 speakers, along with the hBrugchu to their east, are separated from the rest of Khams by much of the Amdo-speaking area. The hBrugchu type of Khams, spoken by about 30,000 people, is the easternmost of all Tibetan groups.

2.1.2 WEST BODISH (GURUNG, TAMANG) GROUP

Many members of these groups have moved away from traditional areas, and now speak only Nepali. Within some groups such as the Gurung there is a cultural and religious continuum from north (Buddhism) to south (Hinduisation). Transitional between the core Gurung Group and Tibetan are Ghale and Kaike. Some scholars prefer to call this the Tamang Group, after the language with the most speakers. The following shows the relationships within West Bodish. As noted below, there is very substantial dialect diversification within Gurung and Tamang.
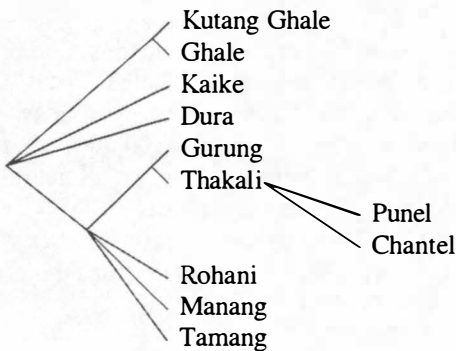
Kutang Ghale was first reported in Webster (1992), and is the north-easternmost West Bodish language, spoken between the Tibetan varieties of Nubri and Tsum. Kutang is a local name for the area; the 1,300 speakers call themselves Bhotte and are culturally Tibetan and in close contact with the adjacent Nubri and Tsum; however, what they speak is linguistically closer to Ghale. Hence Webster (1992) coins the name Kutang Ghale for the language.

The Ghale language, with about 15,000 speakers in 33 villages, and Kaike are culturally closer to Tibetan than to the rest of this group; they also show the linguistic effects of contact with Tibetan. Some 'Gurung' actually speak Ghale, and some 'Ghale' speak Gurung; the map shows the linguistic situation.

Kaike is spoken by the 'Magar' of three villages, a total of about 2,000 people; it is also sometimes known as Tarali Kham though it is quite different from Kham, a Himalayan language of western Nepal.

In the eastern part of the Gurung range, there are small numbers of Dura speakers; the language is being replaced by Gurung and Nepali, but is still spoken by some older people. No linguistic description is available, but it is said to be quite distinct from Gurung.

Gurung, with over 200,000 members of the ethnic group and perhaps 150,000 speakers, has two major dialects, west and east, which are so different as to lead to lack of mutual intelligibility. There are subdialects within each variety. There is a recent tendency to call this group by its autonym, Tamu.

The Thakali are a numerically small but economically important group of about 5,000 speakers. They are also known as Thaksya, or by the autonym Tapaang. Usually included in the Thakali are the closely related Panchgaon ('five village') people to the north, including Punel (Marpha village), Syangtani (Syang village) and Chhimtani. The Thakali are vigorous and successful traders throughout Nepal, and traditionally into Tibet. The Chantel to the south-west of the Thakali are concentrated in Myagdi and Baglung villages. Linguistic data is limited, but de Sales (1993) suggests that Chantel is a variety of Thakali spoken by part of a composite group of former copper miners in Myagdi. The Chantel of Baglung, she claims, have always spoken Nepali.

Rohani is a small TB-speaking group to the south of the Thakali and west of the Gurung. Due to lack of linguistic data it is not clear whether or where it fits into the Gurung Group.

Manang is spoken by a strongly Buddhist group of about 3,000 also known as Nyishang or Nyeshang. Their language shows Tibetan influence and is in close contact with nearby varieties of Tibetan.

Tamang was formerly known as Murmi, Ishang or Sain, and is widely spoken in the hills around the Kathmandu valley; there are very large dialect differences among the approximately 600,000 speakers. Tamangs are scattered throughout Nepal and into India, but many of these no longer speak Tamang. Webster (1992) reports a small group of 3,000 to 4,000 Tamangs, who refer to themselves (in Nepali) as Gurung, at the north-western end of the Tamang area, adjacent to the Ghale in north-eastern Gorkha District.

2.1.3 EAST BODISH GROUP

This subgroup is considerably divergent from other varieties of Tibetan, and appears to form a further subgroup of Bodish. It includes some 80,000 speakers of Bumthang, Khengkha and Kurtöpkha in central Bhutan, and also the smaller 'Nyenkha or Henkha (10,000), 'Olekha Monpa, Chalikha and possibly Dakpakha languages (1,000 each) of eastern Bhutan, as well as the Dzalakha/Central Monpa/Cuona Menba and related languages or dialects. The following are the relationships within East Bodish.

Bumthang
Khengkha
Kurtöpkha

'Nyenkha/Henkha
Chalikha

'Olekha Monpa

Dakpakha

Dzalakha
East Monpa
Sherdukpen
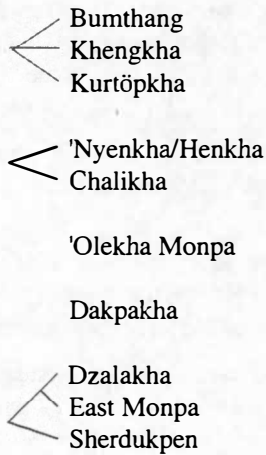
This subgroup is sometimes known as the Bumthang group from the westernmost members of the cluster. The best described language in this group is known to its speakers as Dzalakha and is spoken in extreme north-eastern Bhutan by some 15,000, in Tawang District of north-western Arunachal Pradesh, India by over 30,000 where it is known as Central Monpa, and in adjacent areas of Tibet around Cuona by a further 7,000 where it is included in the Menba nationality and known as the Cuona Menba dialect. For descriptions, see Lu (1986), Das Gupta (1968) and Sun et al. (1980). Further east in north central Arunachal Pradesh and in adjacent areas of Tibet there is a smaller group of about 5,000 'East Monpa' (Indian term) or 'Motuo Menba' (Chinese term) speaking a variety similar to but distinct from Dzalakha; see Zhang (1986) for a description. In addition, there is a small group in the south-eastern part of West Kameng District; this is usually collectively known as Sherdukpen from two large villages of this group, Shergaon and Rupa (Tukpen). This group of about 4,000 is sometimes subdivided according to other village names, such as the Bot subgroup centred on Bot village, and the Lish subgroup. Their autonym is Mei.

2.1.4 TSHANGLA AND EAST BODIC

The Tshangla group lives in south-eastern Bhutan and western West Kameng District of Arunachal Pradesh; in this area, Tshangla was the traditional lingua franca among the various 'Monpa' groups. Another name for the Tshangla is Sharchop 'eastern people', which refers to their geographical position in Bhutan. They are sometimes also called Southern Monpa in the literature, especially in India, and they are thus sometimes confused with the Dzalakha and other groups.

Two other small groups of Bhutan speak non-Tibetan varieties of East Bodic. These are the Lhokpu (in Dzongkha, Lhobikha) of south-western Bhutan and the Gongduk (in Dzongkha, Gongdubikha) of south central Bhutan. Van Driem (1992) suggests that these may have been the autochthonous groups of the area, prior to the migration southwards of the various Tibetan groups. The exact classification of these two groups is not yet certain.

2.1.5 WEST HIMALAYISH

Influenced by contact with Tibetan to its north but still quite distinct, various languages of Lahul, Kinnaur, Almora and western Nepal are grouped by Benedict (1972) under the term Kanauri (the former spelling of Kinnaur); earlier classifications link them instead with the TB languages of central and eastern Nepal.

West Himalayish/Kanauri

NNW (Lahul)
Pattani (Manchati)
Tinan (Gondhla), Ranglo
Bunan (Gahar)

NW (Kinnaur)

Kanauri/Kinnauri
Upper (Thebar)
Lower
Chitkhuli

Kanashi/Malana

Almora
Rangkas (Rangpa) (1)
+Rangkhas (2)
Darmiya
Chaudangsi/Byangsi

Eastern (Nepal)
+Bhramu
Thami
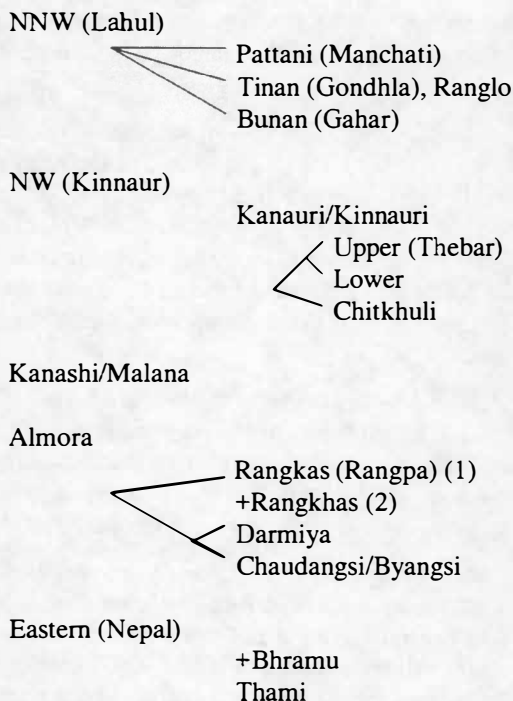
These groups are Hinduised residents of the valleys of northern Himachal Pradesh and northern Uttar Pradesh. At the northern edges of this region they are in contact with the local Tibetans, and on the southern edge there is also a much larger Indic-speaking population in the lower valleys. The names for them which appear in the early literature are based mainly on Indic placenames. Some of these groups have winter villages lower down in the valleys and summer pastures and houses higher up; others are more sedentary.

The term 'Lahuli' is sometimes used to refer generally to the non-Tibetan languages of Lahul, including Pattani, Tinan and Bunan, but is not used in this way locally. What is now known in the literature as Pattani or PaTani [paṭṭani] was formerly called Manchati. It is spoken along the Chandra (upper Chinab) River by about 20,000 people who are Hindu, and shows extensive lexical influence from Indic. What is sometimes known in the literature as Chamba Lahuli is a variety of Pattani spoken by about 6,000 people in what was formerly north-eastern Chamba, along the Chinab (Chandra) River at the eastern edge of the district; this area has now been transferred administratively to Lahul. Pattani has recently been described by Devidatta Sharma (1989a), by Suhnu K. Sharma, and most recently and thoroughly by Anju Saxena, some of whose results are reported in this volume.

Tinan, Tinani or Gond(h)la is spoken by about 2,500 people south of the Chandra (upper Chinab) River just above (to the south-east of) its junction with the Bhaga River, east of Pattani. Ranglo ('near the pass') is spoken by about 1,000 people further to the east; as noted above, this is a variety of Tibetan. Tinan is closely related to Pattani. Again, much improved recent descriptions for Tinan and the adjacent varieties of Tibetan have been provided by Devidatta Sharma (1989a).

The Bunan or Ga(h)ri language is spoken along the Bhaga River in western Lahul by about 5,000 people; it shows more Tibetan influence than Pattani. Devidatta Sharma (1989a) has also described this language. To its immediate north is a variety of Tibetan locally known as Tod.

Kanauri, now also known as Kinnauri (from the modern Indian spelling of the area's name) is the non-Tibetan language of the Kinnaur area. It is spoken by about 60,000 people, and includes closely related lower and upper Kanauri (with the latter also known as Thebor or Thebar) and a divergent variety, Chitkhuli, spoken in two south-eastern villages. Again, recent descriptive work by Devidatta Sharma (1988, 1992) and Anju Saxena (1992), and further comparative work by Saxena, have greatly improved our knowledge of this language. Varieties of Tibetan are spoken further up the same valleys; in this case, Nyam and Spiti.

The Kanashi language, spoken in Malana village near Kulu by about 1,100 people, is a separate subgroup within West Himalayish. There is no recent description of this language; most scholars continue to use the hundred-year-old *Linguistic survey of India* materials; see Grierson (1903-1909). Devidatta Sharma (1992) summarises and updates these, with some additional lexical material gleaned from recent non-linguistic sources.

The Rangpa people, formerly known as March(h)a, whose language Rangkhas has recently been well described by Zoller (1983) and also by Devidatta Sharma (1990), live in north central Garhwal in Uttar Pradesh. Zoller estimated about 5,000 speakers in 1983; a current estimate would be some 7,500 speakers. In addition to being used to refer to this group, the term Rangpa has come to be used as a collective term for all of the Hinduised non-Tibetan TB groups of north-eastern Uttar Pradesh, in place of the Indic term Bhotia which is both somewhat pejorative and also does not distinguish them from the local Buddhist Tibetans further up the same valleys. The adjacent Indic-speaking Tolcha group intermarry with the Rangpa; Zoller suggests that they were also formerly speakers of a TB language.

The Darmiya are a small group in northern Almora; to their east are the Chaudangsi, and to the north-east are the very similar Byangsi. The Byangsi also live in north-eastern Nepal; sources such as Devidatta Sharma suggest that Chaudangsi and Byangsi are varieties of one language. Another group, known by the placename Johar or by the term Rangkhas (but not to be confused with the Rangpa who speak Rangkhas further to the west), live to the west of the Darmiya, but their TB language has recently been completely replaced by the local Indic language; Devidatta Sharma (1989b) summarises the available data. The total number of speakers of the three remaining languages is about 12,000, mostly in India. Recent descriptions in Devidatta Sharma (1989b) have greatly improved our knowledge of these languages.

The Bhramu or Baram language was spoken to the north-west of Kathmandu; but existing descriptions date from the 1850s and the language is probably now dead. It is poorly described. The Thami language in eastern Nepal is spoken by about 14,000; it is the easternmost of what Shafer classifies among the West Himalayish (Benedict's Kanauri) languages. Toba (1990) is a dictionary of this language, which is unfortunately not yet published.

2.2 HIMALAYAN

These languages appear to form a group within TB, but some of the relationships are remote and obscured by contact. They comprise nearly all the non-Tibetan TB languages of Nepal. In general Tibetan-derived groups inhabit the northern quarter of Nepal; and Indic groups inhabit the southern quarter, the plains of the Terai. In addition there has been extensive migration which has spread Nepali, the Indic national language, northward and eastward throughout the country and into north-eastern India; and to a lesser but still substantial extent eastward migration of TB groups who retain their ethnic identity outside their traditional areas but mostly speak only Nepali. I am most grateful to George van Driem for the detailed information in a preliminary version of his forthcoming volume, on which much of this map is based. Another important source is Hansson (1991).

Grierson's (1903-1909) division into pronominalised and non-pronominalised Himalayan languages has been shown by Bauman (1975) and Caughley (1982) to be based on secondary and independent morphological developments. Shafer (1974) divides these languages into West Himalayish, West Central Himalayish and East Himalayish sections of Bodic, with Newari less closely and clearly linked. Benedict (1972) prefers to connect the West Himalayish or Kanauri languages more directly to Tibetan in his Tibetan-Kanauri, with Magar intermediate between this and his Bahing-Vayu which comprises the rest of the languages here. Glover (1974) does not consider the West Himalayish/Kanauri languages, but the rest he divides on lexical grounds into East Himalayish (Limbu plus the 'Rai' languages) and Central Himalayish, which includes the rest as well as Tibetan and its outliers. This last classification accords with the traditional classification in Nepal, which groups the 'Rai' or Kiranti TB languages of eastern Nepal as opposed to the others.

What are not included here, contra Glover, are the Bodish languages of western central Nepal: Gurung, Tamang, Thakali, Manang, Ghale and Kaike; nor the numerous Tibetan dialects along the northern borders; the best known of these is Sherpa in the north-east, but virtually every northern valley in Nepal has one. Also not included are the numerous post-1959 Tibetan refugees.

Central

Magar
Raute
Raji

Kham

Chepang

Newari

'Rai'Kiranti — Hayu

Sunwar
Surel
Bahing (Rumdali)
Chaurasia (Umbale)
Jerung

Thulung

Khaling

Dumi
Kohi
Kulung
Sotang
Nachering
Parali
Kudak 'd.' of Sangpang
Chukwa

Sangpang
Bantawa
Dungmali
Khesang
Chamling
Puma

Athpare
Belhare
Chintang
Chulung
Yakkha
Lumba

Mewahang (Newahang)
Lohorong
Yamphu
Yamphe

Tilung

Chathare Limbu
Limbu

2.2.1 CENTRAL HIMALAYAN LANGUAGES

The Magar are a numerous and widely scattered group, found throughout the lower hills in western Nepal; this was formerly the language of a major kingdom in the area. The language was underenumerated at 212,681 by the 1981 census; van Driem (forthcoming) estimates 290,000 speakers. Many more Magar no longer speak the language, and some people classified as Magar speak other TB languages such as Kham.

Raute and Raji are two names for small nomadic groups of western Nepal and adjacent areas of India. Devidatta Sharma (1990) provides some data on Raji as spoken in the Kumaon region of north-eastern Uttar Pradesh in India, derived from recent work by Shobha Ram Sharma and from Grierson. His conclusion is that Raji is a Munda language with extensive contact lexicon from TB and from Indic sources. Some sources have suggested that the Raute and Raji speak Magar; more fieldwork is needed to determine whether Raute and Raji are the same or not, and if not what the linguistic position of Raute may be.

The Kham group is usually included with Magar, but speak the quite distinct language Kham (with various dialects) and call themselves Buhda. There are about 40,000 speakers, with dialects Mhai and Takale. Extensive descriptive work has been carried out by Watters.

The Chepang group calls itself [tʃjoʔbaŋ], hence its Nepali name Chepang. There are about 17,000 speakers, including some 2,000 Bujheli (autonym Gharti) in the north-west; a few hundred nomadic Bankariya in the east may also be included. Excellent descriptive work has been carried out by Caughley, with a grammar in print and a dictionary in press. This volume also contains a paper on Chepang by Caughley.

Newari is the traditional language of the Kathmandu valley, where it was the vehicle of a high civilisation using an Indic script; the earliest dated manuscript is from 1113 AD. The status of Newari has gone down since the Gurkha (Nepali-speaking) conquest of Nepal over two centuries ago, but the language is still very widely used in the Kathmandu valley and in the low hills to the east. Of its approximately 600,000 speakers, nearly all are bilingual in Nepali. Dialect differences are major; the Dolakha dialect described by Genetti is very distinct, and even between the three traditional centres of the Kathmandu valley, Kathmandu, Patan/Lalitpur and Bhaktapur, there are some differences. There has been a recent flowering of Newari literature, using the modern Devanagari script instead of the traditional Newar script, and monolingual dictionaries and grammars as well as several Newari-English dictionaries have appeared. Malla (1985) is a useful English-language grammar. A major long-term effort to produce a dictionary of classical Newari under the direction of Kamal P. Malla is also about to bear fruit. Another centre of Newari language studies is the Newari Department of the Lalitpur campus of Tribhuvan University, headed by Professor Sunder Krishna Joshi.

2.2.2 KIRANTI OR 'RAI' LANGUAGES

Apart from Limbu and Yakkha in the east and Thami (and sometimes Sunwar) in the west, all the TB languages of eastern Nepal are grouped in the category 'Rai', also known as Kiranti (or Kirat) from the former kingdom of this area. In 1981 the 'Rai' mother tongue total thus defined was 221,353; provisional 1991 results indicate about 400,000 Rai, but many do not speak their traditional languages. Some Rai languages are nearly extinct, being replaced by Nepali or by other Rai languages. For example, Bantawa is replacing some of

the smaller adjacent languages to its east. Virtually all of the Rai languages are endangered; they are being replaced by Nepali. Few of the Rai outside eastern Nepal can speak anything but Nepali. Linguistically Sunwar, Yakkha and Limbu form part of the Rai group; the total of speakers of all these languages is over 400,000. The Linguistic Survey of Nepal, a detailed survey of eastern Nepal, was undertaken with German support and directed by Werner Winter some years ago, and preliminary results (Hansson 1991) have recently become available. Some Rai languages have been well described; many still need further research, which is very urgent as most of these languages are dying. Several other Rai languages were studied by members of the Summer Institute of Linguistics, notably Toba's work on Khaling and that of Schulze and Bieri on Sunwar. More recently, Ebert has studied Chamling, Athpare and other languages, and a variety of studies by van Driem and his students have greatly enhanced our knowledge and understanding of Rai languages, especially verb morphology. In the following discussion, these languages are listed starting from the west.

The Hayu or Vayu language is virtually extinct, with only a small number of older speakers; it has recently been excellently described by Michailovsky (1988).

Sunwar (Nepali names Sunuwar or Bahrathar 'twelve clans') was enumerated at 10,650 in the 1981 census, but is estimated by van Driem (forthcoming) to have about 20,000 speakers, while Hansson (1991) estimates up to 25,000. Schulze and Bieri have pubished a number of descriptive studies.

Bahing is also known as Rumdali; the language is closely related to Sunwar and Chaurasia. Chaurasia is also known as Umbale; the language is extremely closely related to Jerung. The Jerung language is known to its speakers as [ɟero mala].

With about 8,000 speakers, Thulung forms a separate subgroup within the Rai languages. It has been described by Allen (1975).

With about 12,000 speakers, Khaling also constitutes an independent subgroup within Rai. Some scholars suggest that it is mutually intelligible with Dumi. This language has been extensively described by Toba.

The Dumi language is spoken fluently mainly by older people; this group has about 8,000 members by some estimates, but very few speakers, all of them old. A grammar (van Driem 1993) has recently appeared.

The Kohi or Koi group speaks a language very similar to Dumi; the two are closest to Kulung and Nachering.

Kulung, also known as Kulunge, is spoken by about 9,000 people. Sotang, also known as Sotange and centred around the village of Sotang, is said to be a dialect of Kulung. The combined total for the two is 15,000 speakers. Ongoing studies by Tolsma include the article which appears in this volume.

Nachering is the language of a fairly small group; another group, Dimali, is sometimes separated from it. The Parali dialect of Nachering is almost undescribed; it is spoken by only a very small group.

Chukwa is another independent subgroup of Rai, close to Kulung-Nachering but not part of it.

One dialect of Sangpang (or Sangpahang), spoken at Kudak is actually closer to Kulung-Nachering, but other dialects form a subgroup of Rai close to but distinct from Bantawa; this may be due to recent contact with Bantawa.

The large and dialectally diverse Bantawa Rai group is widely scattered outside its traditional area, and was also used as a lingua franca at an earlier period. Various named subgroups exist, including Panguwali, Amchoke, Arthare, Dilpali, Wahitpang and probably others. Lambichong or Mugali is a small but distinct group whose language is being replaced by Bantawa and Nepali. Novel Kishore Rai (1978) is a pedagogically oriented grammar.

The Dungmali and Khesang language forms a subgroup of Rai fairly close to Bantawa; on the map both are shown together.

Chamling, also known as Rodong, is fairly numerous and widely scattered. Ebert has worked and published extensively on this language.

The Puma are a small Rai group whose language is similar to Chamling. Of the other Rai subgroups, Chamling and Puma are closest to Bantawa.

The Athpare language is in the Athpare-Yakkha subgroup of Rai along with Lumba, Lambichong and Chulung. Not suprisingly given its location, it shows some affinities with Chattare Limbu as well. Elbert has also worked with this language.

Belhare is the language of Belhariya village in the Athpare area, the most recently recognised of the Rai languages; it was formerly thought to be a variety of Athpare. There is substantial recent descriptive work on this by Bickel, including a paper in this volume.

Chintang is a small language in the process of being replaced by Bantawa; it is closest to Chulung. Both form part of the Athpare-Yakkha subgroup of Rai.

The small Chulung Rai group speaks a language close to Chintang.

The Yakkha are a fairly large group not usually included within the 'Rai' category, but the language is closest to Rai languages such as Athpare.

Lumba is another of the Athpare-Yakkha subgroup of Rai.

Mewahang, also referred to as Newahang, is a large language with eastern and western dialects. Several other extinct or nearly extinct Rai languages are very closely related to it; these include Saam, Sambya, Bungla and Pongyong.

With two dialects, northern and southern, the Lohorong or Lorung language forms part of the Lohorong-Yamphe group.

The Yamphu language occupies an intermediate position in its subgroup between Lohorong, Yamphe and southern Lohorong.

Yamphe is a small Rai group whose language is also known as Yakkhaba and sometimes included in Newahang/Mewahang.

The Tilung Rai group speaks a quite distinct language; it probably includes the so-called Dorungkecha Rai and Choskule Rai.

The Limbu group is divided into four subgroups, of which the south-western Chathare ('six clans') Limbu speak a distinct language. The south-eastern Panchthare ('five clans') dialect extends into India (Darjeeling and Sikkim); central Phedappe and northern

Tamarkhole varieties are spoken only in eastern Nepal. Overall, there are nearly 200,000 speakers, with over 20,000 Limbu (but fewer speakers) in India. The traditional 'Kiranti' (Limbu) script is being brought back into use, both by Limbu literati in Nepal and in the education system of Sikkim State in India. There are now published grammars of the Panchthare dialect by Weidert and Subba (1985) and of the Phedappe dialect by van Driem (1987), with an unpublished study by Michailovsky of the Tamarkhole dialect still to appear.

The Dhimal in the south-eastern corner of Nepal speak a language which is also spoken by a group known as Toto in Jalpaigiri District of West Bengal in India, just south of the south-western corner of Bhutan. Ongoing studies by Toba provide recent data on this language. Though the Dhimal themselves believe that they are closely related to the Limbu and are thus Kiranti, there appear not to be historical linguistic grounds for this belief. The language also does not fit with the reconstructions of Sal in Burling (1983), nor with those of Sun (1993) for the Adi-Mising-Nishi, Misingish or Tani portion of Central TB. Dhimal/Toto may thus form a separate subgroup within TB, or it may fit in some as yet undetermined way within Himalayan or Central TB.

3. NORTH-EAST INDIA/SAL

This subgroup was named by Burling (1983) from the distinctive etymon for 'sun' *sal found in most such languages; it is characterised by numerous other innovative etyma, such as *war 'fire', *s-raŋ 'sky' *wa 'father' and *nu 'mother'. It was first proposed as a subgroup in Benedict (1976), and comprises:

(1) Shafer's Baric group (Benedict's Bodo-Garo-Northern Naga), which covers the plains of north-eastern India as well as a large area of the hills to the east of this;

(2) Jinghpaw (Kachin), the central nucleus of TB according to Benedict (1972), a part of Burmic according to Shafer;

(3) Luish (Shafer, another part of Burmic), also known as the Sak group;

(4) Pyu, an extinct language of central Burma, also known as Tircul.

SAL

BODO-GARO   NORTHERN NAGA   LUISH   JINGHPAW   PYU   KUKI/CHIN/NAGA
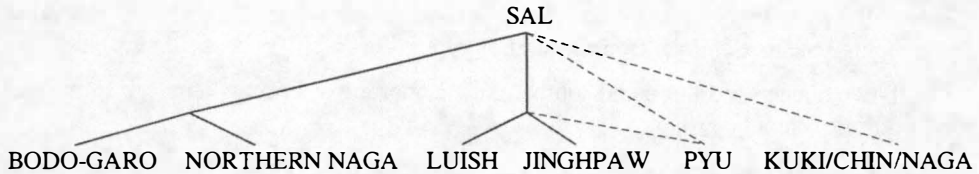
All these languages are SOV, with substantial prefix and suffix morphology. The Luish or Sak group is scattered and moribund but formerly covered a much wider area; for one view see Luce (1985). Jinghpaw is the core group in the Kachin cultural system, which also includes several Burmish and a few other groups which fit elsewhere linguistically. Baric includes Boro (Bodo, Bara, Bårå [bɔrɔ], or 'plains Kachari'), formerly the main language of the upper Brahmaputra valley in north-eastern India, with very closely related languages such as Dimas(h)a ('hills Kachari'), Kokborok (Tripuri), Lalung and so on covering the plains and low hill areas to the south, and fairly closely related Garo in the hills to the south-west; also the 'Koch' languages such as Atong, Rabha, Wanang and so on generally in the plains

to the west, with the still rather closely related Northern Naga languages of northern Nagaland, Tirap District of Arunachal Pradesh, and adjacent areas of Burma to the east. Jinghpaw is spoken in a large area immediately to the south-east of the latter, with the Sak group scattered (in an area mainly inhabited by speakers of other TB languages) to the south. Löffler (1966a) has suggested that Chakma (spoken in Bangladesh and India) and its dialect Tongcengnya (Daingna, Daingnet, Doingnak, Dengnak, Tunzunnya, spoken in Bangladesh and Burma) should be included with the Sak group, but whatever the historical facts, linguistically these are now divergent dialects of Bengali spoken by Buddhists.

BODO-GARO-NORTHERN NAGA

Bodo-Garo
- Boro
- Lalung
- Dimasa
- Hojai
- Kokborok
- Garo
  - Achik dialect
  - Abeng dialect

'Koch'
- Atong
- Wanang
- Rabha
- Hajong
- Deori

Northern Naga
- Khienmungan
- Chang
- Phom
- Konyak
- Htangan
- Wancho
- Haimi
- Nocte
- Tangsa/Rangpan

Jinghpaw (Kachin)

Jinghpaw
  Gauri dialect

Sak Group/Luish
- Kadu
- Ganan
- Taman
- +Malin
- +Chakpa
- +Phayeng
- +Sekmai
- Sak

+Pyu
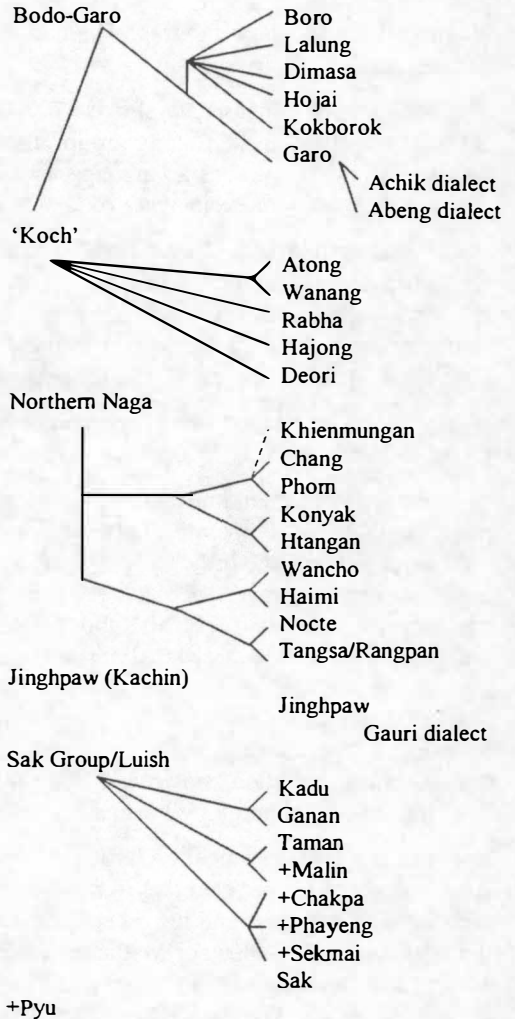
3.1 BODO-GARO LANGUAGES

It has recently been suggested that the large and widely-distributed Boro group has over four million members, though many of these now speak mainly or exclusively Assamese. According to All Bodo Students' Union (ABSU) (1987) the total is 4,104,000, of whom most live north of the Brahmaputra in Assam, with smaller numbers south of the river, in adjacent states and into Bhutan, Nepal and Bangladesh. In some areas, such as Nepal, they are known as Mech or Meche; this is now viewed as a pejorative name in India. Another general term is Kachari; the 'plains Kachari' are Boro. The number of speakers is much smaller. 1971 census figures give over 600,000 speakers, though this is an underenumeration; a more plausible current estimate is about a million speakers. The language has a roman orthography and a Devanagari one, and is used as a medium in some primary and secondary schools.

Though 1971 census data shows only 10,650 Lalung speakers, ABSU (1987) claims 200,000 members of the Lalung group. Intermingled with Boro to its north and Garo to its south-west, this group has a language very close to Boro; though again most of the group now speaks mainly or exclusively Assamese.

Centred in the North Cachar Hills, the Dimasa or Dimasha group is also known as 'hills Kachari'; it is also scattered in adjacent areas. Census data from 1971 give nearly 38,000 speakers, ABSU (1987) claims 150,000 group members. There is an Assamese-based orthography in which a recent dictionary, Baruah (1992), has appeared; also an older but little-used roman orthography. The language has some dialect diversification.

The Hojai language of the Nowgong area in central Assam is claimed to be used by a group of 20,000 (ABSU 1987); only very limited linguistic data are available.

Kokborok is better known as Tripuri and is the indigenous language of the former princely state 'Hill Tipperah' (Tripura) in north-eastern India. It has a roman orthography as well as an earlier Indic orthography; Pai (1976) is a recent grammar. The current name Kokborok means 'language – people'; the word for 'people' is of course related to the name of the Boro. Officially there are about 350,000 speakers, with substantial numbers in Bangladesh but concentrated in Tripura; ABSU (1987) estimates 700,000. The main dialects include the standard Debbarma (western), also Riang (Reang, south-eastern), Noatia or Tipra (eastern), plus various smaller dialects: Jamatia and Darlong (north-eastern), Aiang, Dahula, Karpong, Koloi, Laitong, Muslung and Rupini. In the Chittagong Hill Tracts of Bangladesh, the Kokborok are often known as Mrung from the Arakanese name for this group. Recent descriptions of Kokborok by Pushpa Pai have been published in India.

Garo, the language of the Garo Hills in western Meghalaya, has about half a million speakers; about ten per cent live in Bangladesh. The standard dialect, Achik, is to the east in two-thirds of the area, and the other main dialect, Abeng, is in the west. Between the two in the south is the Matabeng or Matjanchi dialect. Within Achik there are Gara and Ganching; Matchi and Dual; Chisak; and Awe or Akawe subdialects from south to north. Some 'Koch' languages (Atong, Wanang, Hajong) are also officially (but incorrectly) regarded as dialects of Garo. There is a well-established roman orthography which has replaced a former Indic orthography.

Spoken to the south-east of Garo and included in Garo census figures, Atong is a 'Koch' or 'Konch' (Western Bodo-Garo) language; available linguistic data are very limited.

Another such language is Wanang, spoken by small groups of 'Garo' at the western edge of the Garo area and by a smaller group in Jalpaiguri District of West Bengal; Burling (1961) prefers to classify it as a dialect of Rabha. The major 'Koch' group is the Rabha, immediately to the north of the Garo in the Brahmaputra valley with a number of dialects including Tintikia and possibly Wanang. ABSU (1987) claims 400,000 members, but even if there are this many Rabha, most of them now speak only Assamese. A more conservative estimate would be about 50,000 speakers. A roman orthography exists.

Hajong is another small 'Koch' language spoken to the north-west of Garo and scattered elsewhere. Burling (1961) reports another 'Koch' language, Ruga, to the south of the Garo, and other names appear elsewhere in the literature. All are poorly known; one estimate of the total number of speakers of 'Koch' kanguages (other than Rabha and Atong) is 35,000, centred around Cooch Behar.

Also known as Chutiya or Deori Chutiya, the Deori group is scattered along the Brahmaputra, and acording to Saikia (1976) is still spoken only by the Dibongiya subgroup. Figures from the 1971 census show 14,937 speakers, though ABSU (1987) suggests 150,000 members of the group. This is the most divergent of the core Bodo-Garo languages according to Benedict.

3.1.1 NORTHERN NAGA

This subgroup is Shafer's Nagish portion of Baric, where Benedict and more recently Burling also place it. Its subgrouping has been the subject of French (1983); Marrison (1967) independently separates it from the other Naga languages as his Naga A. I am very glad to acknowledge personal communications from G.E. Marrison and J. Morse in preparing this portion of the map and text. The names of 'Naga' groups are notoriously confusing and confused; Assamese or other names of villages, rivers or towns where contact occurred, clan names for the very numerous subgroups of each group, names used by other 'Naga' groups to refer to a group, autonyms and descriptive names are all used.

The Khienmungan are known in the anthropological literature as 'Kalyokengnyu' from their stone-roofed houses, this large group is the southernmost Baric 'Naga' group. About one-sixth of the group is in India but most are in Burma where parts of it are known as Nok-aw or Nauk-o (a clan name), Ponyo (a village name), Para or Paya (a Burmese name of uncertain origin) and so on. In India the Sema call them Tukhemmi and the Chang call them Aoshedd; many alternative representations of the autonym also are seen: Khiamngan, Khemungan and so on. The language is virtually undescribed, but is probably closer to the adjacent groups, that is the southern Nagish Chang, Konyak and so on, than to the northern Nagish groups.

The Chang, a small group of northern Nagaland, has nearly 20,000 speakers; a roman orthography exists, but little recent linguistic data. The Ao and Konyak name for the group is Majung, Mojung or Manjung, and the Sangtam name is Machongrr.

Phom is another relatively small group with over 20,000 speakers and a roman orthography but little linguistic data. Their former autonym was Chingmengnu, and they appear as Assiringia (a village name) and Tamlu in the literature.

Konyak is the largest 'Naga' group in India, with over 90,000 speakers and a roman orthography. It is at the northern tip of Nagaland, with a small number of speakers in adjacent areas of Burma as well. The 'standard' dialect is spoken at Wakching; in older sources this is also referred to as Tableng, Mulung, Kongon or Angwangku. Various sources list over thirty current clan or village names or varieties.

In Burma to the east of the Konyak and south of the Wancho are the Htangan; Marrison (1967) suggests that it is closely related to Konyak, if not a dialect of it; but no current linguistic or population information is available.

The Wancho are a substantial group of about 40,000, mainly at the southern tip of Tirap District of Arunachal Pradesh and extending into Burma. A roman orthography exists but is not in use. This group was formerly known as Banpara, Mutonia, Joboka or Jokoba, with subgroups Khulung-Muthun, Bor-Muthun and Horu-Muthun. It has two main subgroups: Changjan and Tangjan. Its genetic position is not agreed: French tentatively links it with Chang and attributes lexical similarities with northern Nagish to contact, while others regard it as part of northern Nagish.

Haimi is a large group in Burma with nearly twenty named clan-dialects. No linguistic data are available, but it appears to be a northern Nagish language as its speakers are developing a shared roman orthography based on the Moshang clan dialect of Rangpan.

Nocte is a large group of about 40,000, about half in India and half in Burma. There is a roman orthography in India as well as the new shared orthography in Burma. Formerly known as Namsangia (a village), Borduria (another village) or Jaipuria (a town which many Nocte visit), it has six main dialect groups: Hawajap, Japejap, Kapajap, Lazujap, Photungjap and Tangjap.

In one case, a group has two names: Tangsa in India and Rangpan in Burma. They number about 15,000 in India and somewhat more in Burma, for a total around 40,000. Again a multitude of clan and other names appear in the literature; Moshang (Mawshang) and Shangge are two such. There is a roman orthography for Tangsa and a new orthography based on the Moshang clan dialect in Burma, to be used by Rangpan/Tangsa, Nocte, Haimi, and Wancho there. This orthography is a considerable improvement on most such, as it indicates the tones.

3.2 JINGHPAW

Known in India as Singpho, in China as part of the Jingpo nationality, and in Burma as Kachin, this group (whose own name is [tʃiŋ³¹ phɔʔ³¹]) forms the core of the Kachin culture complex in northern Burma, with minor extensions into China and India. The official 1983 population of 'Kachin' in Burma was 465,484. This includes a large number of Burmish language speakers: Maru, Atsi, Lashi and Ngochang, but is certainly an underenumeration for the wider 'Kachin' group. In China there were nearly 93,000 members of the Jingpo nationality, but only a bit over 20,000 speak Jinghpaw as their first language. In India the total is about 2,000, plus some former speakers of Thai languages who now speak Jinghpaw. A conservative current estimate of the 'Kachin' population is 0.9 million, of whom over 600,000 speak Jinghpaw as a first language, with the rest speaking it very fluently or bilingually as a second language and using it as the medium of literacy; less-fluent second-language speakers include many Rawang, Shan, Lisu and Khamti in Burma. The

Jinghpaw roman orthography is very widely used, but unfortunately does not indicate tones. Separate roman orthographies are now being developed for some Burmish 'Kachin' languages.

Within Jinghpaw there are some divergent dialects, most notably the Gauri (Kauri, Hkauri) dialect of the area just north of Bhamo in Burma. Sometimes other clan names are cited as if they were also dialect names.

3.3 SAK OR LUISH GROUP

The Kadu appear in the history of Burma as the former dominant group of the Tagaung kingdom in Upper Burma, under the name Kantu; they should of course not be confused with the eponymous Mon-Khmer group in Vietnam. Their language is moribund, surviving with only about 20,000 speakers in the hills north of Mandalay; the group's autonym is [ɑsa?].

Closely related to Kadu and now spoken to its immediate west, Ganan has some 7,000 speakers in about 20 villages. Both Kadu and Ganan are poorly described.

Taman is reported in one village north of Homalin in western Burma. It has fewer than a thousand known speakers, though others may be scattered nearby. The probably extinct language Malin was very close to Taman, and very small groups in adjacent areas of Manipur in India which formerly spoke closely related languages, known in the literature as Andro or Undro; Sengmai; and 'Chairel' (the last is a village name), and more recently reported under the names Phayeng, Sekmai and Chakpa. None of these languages is well described.

The Sak group is often known from the Burmese form of its name, Thet, or Arakanese That. Their autonym is [atsa?]; they should not be confused with the Bengali-speaking but Buddhist Chakma. The best source is Bernot (1967), but see also Luce (1985). They total about 7,000 speakers, with several thousand each in Burma and Bangladesh.

3.4 PYU

The Pyu kingdom of the central plains of Burma used a TB language probably related to the Luish group, though available data is limited. Stargardt (1990) suggests that the Pyu kingdoms were based on irrigated agriculture in side valleys of the central plain of Burma from the second century BC, with major centres at Halin (near modern Shwebo), Beikthano and later a capital at Sriksetra (near modern Prome) which developed from the fifth century AD and was probably formally established as capital in 638 AD; this date is the basis for the present Burmese era. The Pyu came into direct contact with the Chinese court in 800 AD and were conquered by the Nanzhao kingdom of western Yunnan in 832 AD and again by Burmans in the tenth and eleventh centuries AD, and fully incorporated into the Burmese Pagan kingdom from about 1050 AD. Presumably the Pyu language survived for several centuries thereafter; Luce (1985) notes a Burmese inscription as late as 1369 AD mentioning Pyu villages.

Luce (1985) suggests that the Pyu name for themselves was Tircul; Chinese sources call them Piao, and suggest an autonym Tuluozhu (probably pronounced [thuət la tɕju] in late Tang times). The Burmese name Pyu is clearly related to the Chinese and Nanzhao terms.

The language is extinct, but survives in some inscriptions; the best-known is from 1112 AD and is quadrilingual with Burmese as well as Pali and Mon. Other inscriptions from the seventh to eleventh century AD also survive; some are bilingual with Pali. Earlier inscriptions (from the fourth century AD) use the same script for Sanskrit/Pali Buddhist texts. Thus, the Pyu script is the earliest script developed for a TB language, slightly antedating Old Tibetan.

4. KUKI-CHIN

This subgroup is relatively cohesive, both geographically and linguistically, and has been extensively investigated by Shafer, who classifies it as part of Burmic. Benedict likewise links it to Burmese-Lolo. Both include here all of the 'Naga' groups apart from those in the Sal group (Northern Naga or Nagish languages). However it shows substantial lexical and morphosyntactic similarities with the adjacent Sal group, and may be more appropriately linked with it.

Names for these groups are much more numerous than distinct languages. Firstly, there are overall names: in India those who live in Nagaland and northern Manipur are often called Naga, those who live in southern Manipur and points west are usually still called Kuki, while the largest group in Mizoram has lately renamed itself Mizo (formerly Lushai). In Burma all are linked under the Burmese category Chin. For example, the Thado (or Thadou) are usually called Kuki or sometimes Thadou Kuki in India, but form part of the northern Chin group and are thus called Chin or sometimes (more specifically) Thado in Burma. One group of eastern Manipur, the Anal, decided a few years back to reclassify themselves as Naga rather than Kuki. Secondly, there are more specific names for subgroups; in many cases former names and names used by outsiders as opposed to autonyms. For example, the Arleng were formerly known as Mikir, and are now officially known by an alternative autonym, Karbi. Thirdly, some of these groups have recently been amalgamating and new names have been coined to refer to these larger groups; for example the Zeme/Nzeme (also formerly called Empeo), Liangmai (formerly called Kwoireng) and Nruanghmei (formerly called Kabui, including the Puiron dialect often referred to as a separate language) now refer to themselves as Zeliangrong, although some Zeme prefer to remain Zeme; previously, before the addition of the Nruanghmei, the term Zeliang was coined to refer to Zeme plus Liangmai, and this is still used as well. An older cover term for these three groups, Kachha Naga ('bad Naga'), for obvious reasons is no longer used. Fourthly, geographical names are sometimes used instead of the more specific subgroup names; for example, Tiddim Chin instead of Kamhau Chin.

Some of the languages included are more divergent; the foremost example is Arleng (Mikir, Karbi) which has long been in contact with Sal Group TB languages as well as non-TB languages; it has even been suggested that there may be a connection between Kuki-Chin and Lepcha, with Arleng as the link (Bauman 1976). Also somewhat different is Meithei (Manipuri), which has long been the language of a Hindu civilisation in the Manipur valley and thus shows more Indic influence. Most of the other languages are spoken by hill groups, some extremely small.

Linguistically the Kuki-Chin languages are characterised by tones (mostly unrecorded and not indicated in orthographies), extensive verb morphology involving tonal alternations and extensive suffixing with some prefixes. There is a widespread *ni ergative suffix on NPs;

the basic word order is SOV. I am grateful to various colleagues for information on Kuki-Chin languages and populations: F.K. Lehman, L. Löffler, G.E. Marrison and others.

4.1 SOUTHERN NAGA

The Ao language, with roman orthography based on the Chungli dialect, has about 110,000 speakers, including some eastern dialects, such as Yacham-Tengsa, which show contact effects with Phom and Chang.

The Sangtam language is found in three main locations, with some dialect differences; the 'standard' basis for the roman orthography of the northern Lophomi dialect. There are about 30,000 speakers.

The Rengma group, perhaps formerly more widespread, lives in various locations. Though regarded as one group, it probably includes three languages: Western, for which Tseminyu is the basis for the orthography; Northern or Ntenyi; and Eastern or Meluri (with the autonym Anyo); the three total about 15,000 speakers. Ntenyi and Meluri are genetically closer to Lhota and Yimchungrü than to Western Rengma, which fits with Maram and Zeliangrong according to Marrison (1967).

The Lhota total about 60,000 speakers; Wokha is the 'standard' dialect, basis for the roman orthography.

The Yimchunger language has some 30,000 speakers in two areas; the roman orthography represents the Yachumi dialect.

Tangkhul are also known as Luhupa or 'savages' to the Manipuri. This large group is almost as numerous in Burma as in India; there are probably about 100,000 speakers. The 'standard' written dialect is that of Ukhrul.

The small Maring group of about 15,000 lives to the south of the Tangkhul; until the Anal declared themselves 'Naga' it was the southernmost 'Naga' language.

The Sema are a large group of about 100,000, with a roman orthography. Ntenyi and Meluri 'Rengma' are linguistically close to Sema rather than Lhota and Yimchungrü according to some Indian sources.

Formerly included in 'Angami' were most of the Chakhesang and the Mao; what is now known as Angami is only the western part of this larger group. Altogether it numbers about 60,000, of whom about 35,000 speak the western dialect and are still called·Angami. Their writing is based on the Tengima dialect.

Chakhesang is the new name for a composite group of eastern 'Angami': the Chokri (about 15,000) and the Kezhama (about 10,000). It also includes a few Sangtam, who live to the north and south of the eastern edge of this group.

The town of Mao (known to its inhabitants as Sopvoma) is the largest village of the Mao group; the language is close to Angami and Chakheseng, and has about 60,000 speakers. There is a roman orthography.

Maram again is named from its main village; the language has over 5,000 speakers. Linguistically it is similar to the Zeliangrong languages.

Ao
Sangtam (N)

Lhota
Yimchungrü
Ntenyi/Meluri

Tangkhul
Maring

Sema
Angami (W)
Chakhesang
　　Chokri
　　Khezhama

Mao

Rengma (W)
Maram
Zeliangrong
　　Mzieme
　　Zeme
　　Liangmai
　　　　Puiron
　　Nruanghmei

Meithei

'Old Kuki'

Rangkhol
　　Bete
Hallam
Langrong
Hmar
Anal
Kom
Chawte
Mayol
Lamgang
other Old Kuki

North Chin   (various)

Central Chin   Lai(zo) (various)
                      Mizo

South Chin   Ashö

                      Khami/Khumi

                      Mara

                      Arleng (Karbi, Mikir)

The Zeliangrong group represents the amalgamation of three distinct 'Naga' groups: in the west and north-west, the Zeme, Mzieme or Nzeme (about 35,000; old exonym Empeo); Liangmai (about 10,000; old exonym Kwoireng, also referred to as Liyang, Lyeng, etc.) in the north-eastern area; and Nruanghmei (about 40,000; older name Kabui, including the Puiron dialect) to the east and south. A former pejorative term for these three groups, Kachha Naga ('bad Naga'), is no longer used. The small Khoirao group (several thousand) has also been included here, so the overall total is about 90,000 speakers. Roman orthographies exist for Zeme, Mzieme, Liangmai and Nrunghmei.

Also known as Manipuri, the Meithei language has a Bengali-based Indic orthography which does not indicate the tones. It is spoken by about a million people, mostly in the Manipur valley but also by small numbers in Burma and other parts of north-eastern India. So-called Bishnupriya 'Manipuri' is not a TB language; it is an Indic language spoken by former Bengali slaves, with some Manipuri lexicon and reduced morphology; most speakers of Bishnupriya 'Manipuri' now live elsewhere. Many non-Meithei Kuki-Chin and other people in the valley also speak Meithei; some of their languages (especially Sak group and some Old Kuki languages) are being replaced by it. In August 1992 Manipuri became the first TB language to receive recognition as a Schedule VIII official language of India.

4.2 KUKI

The term 'Old Kuki' was coined for the *Linguistic survey of India* to refer to various small Kuki groups around the Manipur valley and to the west. It includes some of the Kuki groups of Tripura and nearby parts of Bangladesh and Cachar, for example Rangkhol and Bete, Hallam, and Langrong. These have a total of well over 20,000 speakers. It also includes the strongly Mizo-influenced group Hmar (about 30,000 speakers); the Anal group of south-eastern Manipur and into Burma (about 12,000 speakers); the Chawte (Chote, Chaw, Kyaw) group of eastern Manipur and into Burma (about 1,000); the nearby Mayol (Moyon, 'Mon'; about 1,000) and Lamgang (about 2,000); the Kom of south central Manipur (about 7,000); and various other groups whose languages are nearly extinct, having been replaced by Meithei: Aimol, Kolhreng, Purum, Tarao and so on. Roman orthographies have been created for some 'Old Kuki' languages, but are not widely used. Some groups classified as 'Old Kuki' in Indian sources fit better in Northern Chin, and so are discussed there: Chiru, Gangte, Pawi, Simte, Thado(u) and so on.

4.3 CHIN

4.3.1 NORTH CHIN

This group includes what is perhaps the largest 'Kuki' group in India, the Thado(u), with over 50,000 speakers, as well as such other 'Kuki' groups as Chiru (5,000), Gangte (7,000), Pawi (10,000) and Simte (5,000). Its speakers are very widely scattered throughout Manipur and adjacent areas to the north-west, but concentrated in southern Manipur; some also live in Burma. Along with the North Chin languages of Burma, spoken by about 135,000 people concentrated in the Tiddim area of northern Chin State and along the western edge of Sagaing Division, there are well over 250,000 speakers in all. The local variety of Tiddim has been described in Henderson (1965). In Burma some of the named subgroups

included here are Thado; Siyin, Paite and Vuite; Sokte; Kamhau and so on. There are roman orthographies in use for most of these.

4.3.2 CENTRAL CHIN

The Central Chin are sometimes collectively known as Laizo ('central people'); this includes a very large number of dialects, some of which are so different as not to be mutually intelligible. In fact Mizo (formerly Lushai) is another variety of Central Chin, but with a separate literary tradition. The group includes the western varieties Bawm (formerly Banjogi, 6,000 speakers) as described by Reuchle (1981) and Paangkhua (formerly Pankho, 3,000 speakers) in Bangladesh; northernmost Zahao or Laizo in Burma and extending into India; also Tashon, Ngawn, Zanniat, Zophei, Lawtu, Lailen, Senthang, Tawr and many other groups. They inhabit the central area of the Chin State, including Falam, Haka and parts to the south. For Burma, the population is about 400,000, with a further 10,000 in Bangladesh; in Mizoram they are included as dialects within Mizo. A couple of roman orthographies exist and are in use.

The Mizo, formerly known as Lushai, is the largest Chin group with various dialects; it is also known as Hualngo where it is spoken in Burma. There are about 400,000 speakers in all, including nearly 2,000 in Bangladesh and substantially more in Burma. Some other Kuki-Chin groups are gradually becoming assimilated to the Mizo, culturally and linguistically; these include the Hmar ('Old Kuki') in the north and the Mara (Lakher or Maram) in the south. In India, Mizo is widely used in its roman orthography.

Some Chin nationalists have taken to using the new term Zomi from the North and Central Chin term Zo 'Chin' plus *mi* 'person'. This is intended to refer to all the Chin, including the Mizo. It would certainly be more politically correct to call the Kuki-Chin by some such autonym, rather than the somewhat pejorative exonyms (Naga, Kuki, Chin) now in use.

4.3.3 SOUTH CHIN

Again, this group has many different names. In Bangladesh there are more than 1,000 speakers under the name Khyang. In Burma there are roughly 250,000 group members, but some of those living in the plains no longer speak Chin. Northern subgroup names include Zolamnai ('on the trail'); Welaung and Matu (placenames; the latter call themselves Ngala). In the central part are M'kang, Ng'men, Nitu (or its new name Daai, about 30,000 speakers), Hngizung, Utpu (a large subgroup, called Chinbon by the Burmese); and to the south the Chinbok or Saingbaung Chin (Burmese names), also known as Ashö Chin from the South Chin form of the word for 'person'.

4.3.4 OTHER CHIN GROUPS

The Khami/Khumi group includes several diverse dialects which fall into two subgroups: Khami versus Khumi. In Burma there are about 50,000, with several thousand more in Bangladesh; a roman orthography exists.

Mara is better known as Lakher from the Lushai name for the group. It is also known as Maram from their Central Chin name. They total about 20,000 speakers, mostly in India but

some in Burma. There is a roman orthography, but the language is being replaced by Lushai. The best description is Lorrain (1951).

4.4 ARLENG/KARBI/MIKIR

Arleng is a very widespread language with various dialects: Amri (western), Bhoi (south-western, with Khasi influence), Rengkhang (southern) and so on. All Bodo Students' Union (1987) suggests 600,000 members of the group, while underenumerated census data suggest about 200,000 speakers. The reality is between the two figures; many Arleng have become Assamese speakers. Formerly known as Mikir and now known officially as Karbi (which is one autonym), this group's main autonym is Arleng. The language is rather divergent from the other Kuki-Chin languages, and is well-described in Grüssner (1978).

5. CENTRAL TIBETO-BURMAN

The classification of these languages has not been finally determined, mainly because good data have until very recently been lacking. The subgroup includes most of the languages spoken along India's north-eastern border, the northern tip of Burma, and the adjacent border area of north-western Yunnan and south-eastern Tibet. More or less recent intrusions by Tibetan from the north and Assamese from the south-west have influenced some of these languages more than others. My thanks are due to G.E. Marrison and to members of the Morse family for detailed comments on these languages.

There are five main clusters of languages, which may form a subgroup within TB: the West Arunachal languages; the Adi-Mising-Nishi, Misingish or Tani languages of much of northern Arunachal Pradesh, extending slightly into adjacent parts of Assam and Tibet; the rather diverse 'Mishmi' languages of India's extreme north-east and into Tibet, forming two subgroups, Digarish and Miju/Keman; and the Rawang or Nungish group of northern Burma and adjacent border regions of China. Benedict (1972) classifies the first as the Abor-Miri-Dafla subgroup of TB, does not mention the second, and classifies Nungish as a link between Burmese-Lolo and the rest of TB. Shafer (1974) calls the first Misingish, breaks the second into two groups, Digaru and Miju (but omits Idu), and includes Nungish in Burmic. He is uncertain of the position of these groups within TB, tending to group them with Bodic or Burmic. It is clear from new data that they do not form part of the Sal group, Burling's (1983) name for Benedict's (1976) extension of Shafer's Baric to include part of Burmic. Marrison (1989) provides comparative data on some eastern Adi-Mising-Nishi languages, and J. Sun (1993) is a detailed comparative study of the same subgroup. All these languages are verb-final; many are tonal, but some sources do not indicate these tones.

Another language which may fit here is Lepcha. Its genetic position is debated; while clearly TB, it has been put with the TB languages of Nepal in the *Linguistic survey of India* (Grierson 1903-1909), with Ao and thus in Kukish by Shafer, and recognised as aberrant for its geographical location by Benedict. Bauman (1976) notes some similarities with the aberrant (and geographically nearest) Kukish language, Mikir (Karbi, Arleng). Most recently, Bodman (1987) suggested a fairly close connection with Adi-Mising-Nishi and Nungish, and a more indirect one with Jinghpaw; he also casts doubt on the Austroasiatic connection suggested in Forrest (1962).

Lepcha

West Arunachal ⟨ Sulung
Bugun (Khowa)

Hruso (Hrusso, Aka, Angka)
Dhammai (Miji)
Bangru

Adi-Mising-Nishi/Misingish/Tani Group

Nishi/Bangni/Apa Tani (Dafla)
Hill Miri
Adi (Abor)
Mising (Miri)

Digarsh 'Mishmi'   ⟨ Idu ('Chulikata, Mishmi')
Taraon (Digaru, Taruang, Darang Deng)

Keman 'Mishmi'   Keman (Miju, Geman Deng)

Nungish   Gvnøng ⟨ Trung (Dulong)
Zørwang (Jerwang)
Dvru (Daru)

Anung
Nusu

Rawang ⟨ Dvngsar (Tangsarr)
Mvtwang
Lungmi (Longmi)

5.1 LEPCHA

Lepcha, known to its speakers as Rong, was the traditional language of much of Sikkim and parts of Darjeeling District of West Bengal in India and of south-western Bhutan, but not many speakers remain; probably as few as 4,000. In terms of ethnic group population, this is approximately 26,000 in India (of whom over 22,000 are in Sikkim) and 2,000 in Bhutan. Lepcha has a traditional script; see Sprigg (1983) and Sprigg in this volume for details. In 1977 it was made one of the official languages of Sikkim; since then a substantial effort has been made to teach the script in schools. See Sprigg in this volume for a discussion of changes in this script over the last two centuries. Mainwaring (1876) provided an early description of this language.

5.2 WEST ARUNACHAL LANGUAGES

The Sulung, a small group who were formerly slaves of the dominant groups (especially Bangni) around them, speak a distinct and poorly described language. Their autonym is Puroik, and they have recently been resettled in separate villages. The 1961 Indian census gives 1,516 as the population, but this is too low; there are also some speakers in Tibet. According to the Kho(w)a, autonym Bugun, who are a Monpa group, the Sulung are closely related to them.

The Bugun group of about 1,700, better known as Khoa or Khowa, are under the cultural influence of the Monpa, but they claim close relationship with the Sulung. Limited linguistic information is available in Simon (1976).

The Hruso (autonym; also spelled Hrusso in the literature) are better known by the exonym Aka or Angka; their language, spoken by about 4,000, is poorly described.

The Dhammai [ðɯmmai], better known by the exonym Miji and not to be confused with the Keman (Geman Deng or Miju) far to the east, is a group of over 5,000 culturally similar to the Hrusso, but only limited linguistic data are available. Some early descriptions have also called them Aka, which led to some confusion.

The smallest group of this area is the Bangru (the Bengni name for this group); their autonym is [lə³¹væ⁵⁵]; they number only about 1,000 in Tibet and in Arunachal.

Shafer (1974) and J. Sun (1993) point out that these languages are quite different from the Adi-Mising-Nishi languages to their east.

5.3 ADI-MISING-NISHI (ABOR-MIRI-DAFLA), MISINGISH OR TANI GROUP

The Nishi, Bangni and Apa Tani group was formerly known collectively as Dafla, a name no longer used as it is pejorative. They are now more commonly referred to by various subgroup names such as Bangni (Bengni in Chinese sources) in the west, Apa Tani (Apa Tanang) around Ziro, Tagen or Tagin in the north-east, and elsewhere as Nishi, Nyishang or Nishang. They total about 170,000, of whom only a few hundred (Bangni) live in Tibet (Ouyang 1985). Outside India, officially they (along with all other non-Tibetan non-Buddhist groups of south-eastern Tibet) are included in the Luoba nationality whose name is derived from the Tibetan 'Hlopa', a pejorative word for 'savages' in this area. J. Sun (1993) prefers to use the word for 'person' in these languages, Tani, to refer to the group. Other terms used include Shafer's Misingish. Das Gupta (1969) describes Nishi. J. Sun (1993) has undertaken a detailed reconstruction of this subgroup.

Hill Miri is a small and scattered group in central Arunachal Pradesh; their language is closely related to Nishi and less closely to Miri or Mising. There are several thousand speakers, but census information includes them with the Miri or Mising.

Adi, formerly known as Abor, is now divided into a large number of named subgroups all speaking very similar dialects. It includes Gallong (autonym Galo) in the south-west as described by Das Gupta (1963), Bokar (Bogar in Chinese sources) in the north-west, Pailibo, Bori and Ashing in the north central area, Tangam and Shimong (Simong) in the north-east, Minyong at the centre (including the Karko subgroup), Milang to their east, Pasi and Panggi at the south centre, and Padam in the south-east. These groups total over 125,000 speakers. Bokar extends into Tibet, where there are a few hundred speakers.

Formerly known as Miri, the Mising group is scattered over the upper Brahmaputra valley and into the hills to the north. It may have as many as 500,000 members, but not all speak the language. A romanisation has been developed but not yet approved for use in education. The language, with dialects, is particularly closely related to Adi.

5.4 DIGARISH 'MISHMI'

Idu is one of three 'Mishmi' groups; the Idu were formerly known as Chulikata (crop-haired) Mishmi. The language is spoken by about 20,000, of whom two-thirds live in Arunachal Pradesh and the rest in Tibet. In Tibet they are included with the Luoba nationality.

The Taraon 'Mishmi' group, autonym [ta^{31}ɹuɑŋ^{55}], is also known as Tain, Taying or in China Darang Deng; they were formerly called Digaru Mishmi. There are some 12,000 speakers, most in India, less than half in Tibet and a few in northernmost Burma.

5.5 KEMAN 'MISHMI'

The Keman 'Mishmi' group, autonym [kɯ^{31}mɑn^{35}], is the third 'Mishmi' group; other names include Kaman, Geman Deng and formerly Miju Mishmi. They total about 20,000, two-thirds in Arunachal Pradesh and the rest in Tibet.

5.6 RAWANG/NUNGISH

I am glad to acknowledge personal communications from many members of the Morse family, especially Stephen, Joni and Nangsar, in preparing this map and discussion. Morse and Morse (1966) provide some historical background, Morse (1965) describes the standard dialect in Burma, and Morse (1989) is a brief survey of the dialects in Burma. The Jinghpaw name for this group is Nung, which is the source for the terms used by Benedict and Shafer, and should not be confused with the Central Thai group of north-eastern Vietnam and adjacent areas of Guangxi, China. Former Chinese names include Nuzi, Luzi and Jiuzi. It includes a very large number of subgroups with rather different languages and many subdialects. In Burma, where there are nearly 150,000 speakers, the term Rawang (which fomerly referred only to the largest supergroup in Burma) is now used for these groups as a whole. In China speakers are included in two nationalities, Dulong and Nu; there they total about 25,000 speakers. Sun (1982) describes Dulong; Sun and Liu (1986) describes Nusu, one of three varieties within Nu. In Burma the Mvtwang clan dialect of the Mvtwang clan cluster in the Rawang supergroup has been chosen as the standard, and a romanisation implemented among Christians; in China a Dulong pinyin orthography exists, and a roman script was devised for Anung and used in a New Testament; but it is unclear how widely either is known. Among Nungish groups there are various levels of self-classification: by clan, of which there are probably nearly two hundred; by clan cluster (as shown on the map); by supergroups of several clan clusters; or recently by the overall exonymic collective terms Rawang, Dulong or Nu.

The Dulong nationality in China, plus about 6,500 of the northernmost members of the Nu nationality, form the Trung [tɯ^{31}ɹuŋ^{53}] clan cluster, the northernmost Nungish group. The total number of speakers is thus about 11,000. This should not be confused with the

Taraon/Digaru Mishmi/Darang Deng group. Even within Dulong there are substantial differences between eastern and western dialects.

In Burma, the Zørwang clan cluster is also known as Jerwang or Tvluq; it forms part of the Gvnøng supergroup including Dulong, which is quite distinct from the Rvwang (Rawang) supergroup. A rough estimate is 15,000 speakers. The clan cluster known as Dvru or Daru is also part of the Gvnøng supergroup, and has approximately 35,000 speakers in Burma. The Anung clan cluster, living mainly in China, numbers about 6,000 there, plus a few in Burma. Its autonym is [ɑ^{31}nuŋ^{53}]. They form part of the Nu nationality, which includes several different types of Rawang/Nungish groups.The Nusu group of over 8,000, found in China, has northern [nu^{35}su^{35}], central [nu^{55}su^{55}] and southern [nu^{31}ʃu^{31}] dialects with substantial tonal and other differences. They are also of Nu nationality.

The Dvngsar clan cluster, also known as Tangsarr, forms part of the Rvwang (Rawang) supergroup. It has roughly 15,000 speakers. The Mvtwang clan cluster is the largest, with about 50,000 speakers; it is the main part of the Rvwang supergroup, and its Mvtwang clan dialect forms the basis for standard Rawang orthography. The Dvmang clan cluster is included within Mvtwang as well. Many Rawang in Burma who have a different first dialect can also speak Mvtwang as this has been used as the literary dialect.

Lungmi, also known as Longmi, is the southernmost clan cluster, and has undergone considerable Jinghpaw influence. With nearly 30,000 speakers it is rather different from both Rvwang and Gvnøng supergroups.

6. NORTH-EASTERN TIBETO-BURMAN

While virtually all scholars agree that all the languages here are TB, the exact grouping is not generally agreed. One proposal by Sun (1983b) is that most of them (except Bai, Naxi, Baima and Tujia) form a subgroup which he calls Qiangic. Another by Nagano and by Thurgood is that at least some of these languages (Jiarong, Qiang) can be grouped with the Central TB group. Indeed there are some similarities between languages of this group and the adjacent Central TB Rawang/Nungish languages; these similarities are transitional and may be contact-induced.

The standard classification in China is to group Bai, Naxi and Tujia with the Loloish languages, with Pumi and Qiang more distantly related, and not to recognise the others in western Sichuan as separate languages from Tibetan, since the speakers are classified as members of the Tibetan nationality; all this is clearly incorrect.

Some scholars, such as Nishida and Sun (1990), prefer to classify Baima as a variety of Tibetan, which is arguably correct; or Baima may be transitional between North-eastern TB and Tibetan.

Another language, the furthest north of any TB language, was Xixia; this is now extinct, but survives in numerous manuscripts of the eleventh to thirteenth century AD. Mention should also be made of two TB languages recorded in early Chinese sources: Bailang, for which some songs appear in Later Han history (second century AD); and Nam. Both are rather inaccurately represented using Chinese characters. From their geographical location it is most likely that these were NE TB languages. It is not clear which if any of the modern languages are the modern descendants of Bailang and Nam.

I am glad to acknowledge the assistance of Liu Huiqiang, who has done extensive research on the core Qiangic languages.

```
                          +Xixia

                          Baima

core Qiangic              Jiarong (rGyarung)
                          Qiang          Northern language
                                         Southern language; 5 dialects

                          Ergong
                          Zaba
                          Guichong
                          Muya
                          Ersu (Tosu)     3 dialects
                          Choyo
                          Narmuyi
                          Shixing
                          Pumi (Primi)

other                     Naxi
Northeastern              Moso            ('East Naxi')
TB                        Tujia           2 dialects, north and south
                          Bai             3 dialects, south, central, north(west)
```
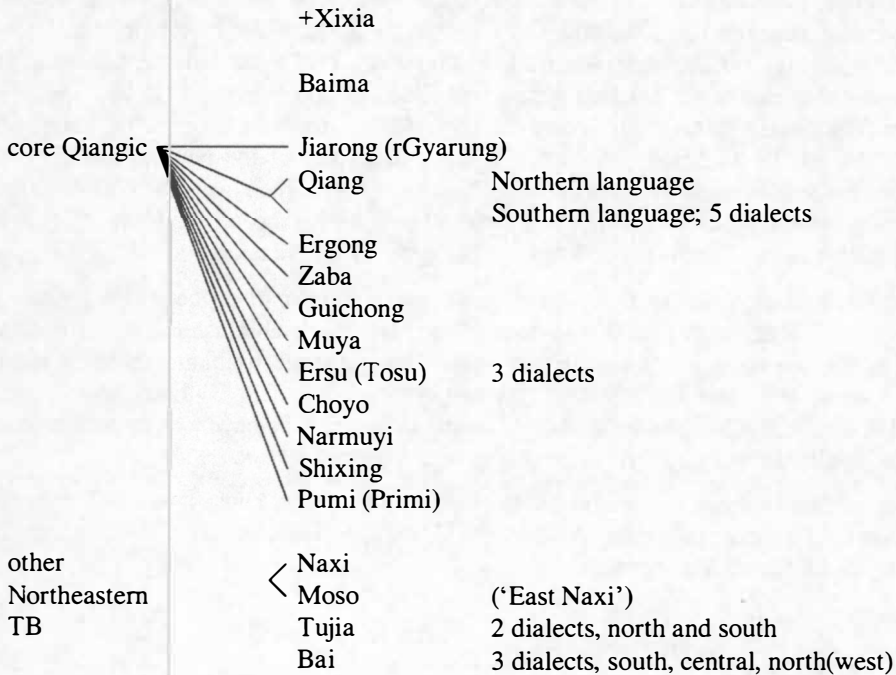
Like nearly all TB languages, the Qiangic languages are SOV. Unlike most other subgroups (and all adjacent subgroups), they have extensive verb morphology which can be reconstructed for this subgroup. Nearly all are tonal.

6.1 CORE QIANGIC

An older Chinese name for this group was Xifan 'western barbarians'. All these languages fall into the Tibetan cultural orbit, and most speak some Tibetan and use Tibetan as the medium of literacy. Apart from the Qiang and the Pumi, which are recognised as separate nationalities in China, all are simply classified as Tibetan, despite their languages. Qiang is another old Chinese name for an ethnic group of this area, but it is unlikely that it has always had only its current referent. These languages are quite diverse, but probably more closely related to each other than to anything else in TB. A pinyin (roman) orthography has recently been developed for Qiang, but no other core NE TB language has an orthography. Qiang includes two 'dialects' which are clearly distinct languages; the northern 'dialect' is non-tonal, unlike the southern one, which has substantial internal diversity.

Population figures for these languages are quite uncertain, as most of them are not recognised. There are probably 20,000 Baima, 150,000 Jiarong, 45,000 Ergong (with three dialects), 8,000 Zaba, 7,000 Guichong, 20,000 Ersu, 7,000 Choyo, 5,000 Namuyi, 15,000 Muya and 2,000 Shixing. For both Qiang and Pumi, the census figures are substantially too low: the census gives just over 24,000 Pumi, but scholars estimate over 55,000 speakers,

most of whom are members of Tibetan or other nationalities; for Qiang, scholars estimate over 195,000 speakers, which is substantially greater than census figures for the Qiang nationality. As is frequently the case in China, the Qiang nationality includes quite distinct languages; in this case, northern is non-tonal and southern is tonal; other differences are also very substantial. Some recent descriptions of these languages have appeared in China, notably Sun (1981) for Qiang, Nagano (1983) for Jiarong, and Lu (1983) for Pumi.

6.2 OTHER NORTH-EASTERN TB LANGUAGES

Naxi, often written as Nakhi in the western literature, appears to be transitional between Qiangic and Burmic; it shares lexical material with both subgroups, but like Bai and Tujia lacks the extensive morphology of core Qiangic. Its two main western dialects are mutually intelligible. The 'eastern dialect', Moso, is a distinct language spoken around Luhu Lake in Yunnan and Sichuan. The traditional Naxi pictographic writing system, extensively studied by Rock, is not widely used now; it is a mnemonic for religious texts known by heart; there is also a related syllabic system which can be read without knowing the text. About 210,000 of the Naxi nationality speak Naxi, and about 40,000 speak Moso. A romanisation for Naxi has recently been developed and put into use; but the Moso do not use this. He and Jiang (1985) provide a recent description of both Naxi and Moso; Bradley (1975) shows that it is not Burmese-Lolo, as it is sometimes classified in China.

Tujia is a very large nationality, with several million recognised members; but only about 170,000 of them speak Tujia. This nationality and language were 'discovered' after the Chinese Revolution, before which they were regarded as Han Chinese. The language has very numerous Chinese loanwords, and is mostly spoken by older people even in its remaining core area on the borders of Sichuan, Hunan and Hubei. It is not written, but has recently been described in Tian (1986).

By contrast Bai has a long and distinguished history, as the main language of the Nanzhao kingdom of western Yunnan. Due to more than a millennium of Chinese contact, with extensive borrowings from various Han dialects at various stages in their development, it is extremely difficult to determine the exact position of Bai within TB. Some linguists have even suggested that Bai has by now become a Chinese dialect. Its syntax is sinicised and shows SVO order. The widespread Chinese idea that it is a Loloish language is incorrect. A roman orthography based on the central dialect has recently been developed, but the speakers have long been accustomed to use Chinese for writing. The three dialects have substantial differences, but not such as to lead to mutual unintelligibility; they are described and compared in Xu and Zhao (1984). Over a million members of the Bai nationality speak Bai; mainly those who still live in the traditional area of north-western Yunnan.

6.3 XIXIA

The Xixia also appear in the literature (according to the Wade-Giles romanisation of Chinese) as Hsi-hsia. Another term for the same group is Tangut, the Turkic and Mongol term for this group. The Xixia kingdom with its capital at what is now Ningxia was founded about 990 AD; the Chinese histories date it from 1032 AD. Though the Xixia state was conquered in 1227 AD, with Genghis Khan dying during the siege and his forces later destroying the city and massacring its inhabitants, Marco Polo (1938:1/150) reported that in

the 1280s the city (which he calls Suçio, a good approximation of a Yuan period Chinese pronunciation of Xixia) existed, and its people (whom he calls Tangut) had their own language.

The writing system is said to have been devised in 1037 AD by Li Yuanhao, a Xixia scholar literate in Chinese. It thus antedates the Burmese script and the old Newar script which are both from the early eleventh century. The meanings of the numerous surviving Xixia texts, mainly translations from Chinese, are usually clear, due to an 1190 Xixia-Chinese dictionary. However the exact phonetic value of Xixia is uncertain. There are various competing hypotheses: a Russian theory; a Chinese theory; and the work of Nishida forming a third alternative. The claims by Kwanten that Xixia was not TB and his rather different ideas about the phonetic value of the script appear to be wrong.

Some aspects of Xixia phonology (such as *a > /i/) and lexicon suggest that it was part of the North-eastern TB group.

7. SOUTH-EASTERN TB

7.1 BURMESE-LOLO

Burmese-Lolo (BL; also more recently known as Lolo-Burmese, or as Burmese-Yipho with Yi (a modern Chinese term) plus the Northern Loloish form of the TB male human suffix) is a large and diversified part of TB; it can be subdivided into two main subgroups, Burmish (including Burmese) and Loloish. A third subgroup, Ugong, appears to be intermediate between the two. For details see Bradley (1979a). The Mru language of western Burma may also be a remotely related part of the Burmese-Lolo group, though some scholars suggest otherwise All BL languages are verb-final, with complex tonal and initial consonant systems but little or no morphology

BURMESE-LOLO

MRU   GONG   BURMISH   LOLOISH

Burmese Burmish   Northern Central Southern

Mru

Ugong

Burmish

Burmese
Northern
Arakanese
Tavoyan
Beik
Intha
Danu
Taungyo
Yaw

Maru
Atsi
Lashi
Achang
Bola
Chintau

Hpun

Loloish

Northern

Nosu

Yinuo
Shengza
Sondi/Adur

Nasu
Laka
Kepo

Sami
Samatau
Sangwie/Sanie

Nisu
Gazhuo
Phula
Kathu
Moang

Central
Sani
Axi
Azha
Azhe

Lisu
Lipho
Lolopho
Lalo
Lahu
Kucong
Jinuo
Zaozou

Southern
Akoid
Akha
Hani
Haoni
Baihong
Akeu
Phana
Sila

Bi-Ka
Khatu
Piyo
Mpi

Bisoid
Bisu
Phunoi
Sangkong

7.1.1 Mru

This language is spoken by about 40,000 people, most of them in Arakan State of Burma and some in Bangladesh. Its exact position in TB is not certain, but it shows various layers of contact vocabulary from Kuki-Chin and from Burmese. According to Löffler (1966b) it is not Kuki-Chin as sometimes suggested, but may be remotely related to the Burmese-Lolo group; this is also Shafer's view.

7.1.2 UGONG

The Gong language is spoken by a small and diminishing population of some 300 Ugong (*u* 'person') in western central Thailand. All speakers are to some degree bilingual in some variety of Thai; younger members of the group tend to be semi-speakers of Gong. The language is no longer spoken in most of its former locations.

7.1.3 BURMISH

7.1.3.1 BURMESE

Burmese is the national language of Burma, with about 32 million speakers of the standard dialect as a first language, usually known as Burmans, and over ten million speakers as a second dialect or language; most of the latter speak another Burmese dialect or a TB language as their first language. Some members of other nationalities, especially many Mons and some Shans, are also monolingual in Burmese. It has a long literary history, with the earliest dated inscription from 1112 A.D.; there are diglossic high and low varieties, with the former used mainly in written or other formal contexts. The preferred variety is as spoken in Mandalay; the most divergent, and in some ways conservative, variety is spoken in the north. Distinct dialects include Arakanese in the west, with about 1.8 million speakers; this is also spoken in south-eastern Bangladesh and adjacent areas of India, where it is usually known as Mogh or Magh. Arakan was reconquered by the Burmans about two centuries ago, and at that time much of the Arakanese court fled to what is now Bangladesh, where they now call themselves [mərəma] but are better known as Marma. Other dialects are south-eastern Tavoyan (400,000) and Beik (250,000); east central Intha (90,000, around Inle Lake), Danu (100,000) and Taungyo (40,000); and west central Yaw (20,000). All the regional dialects are in some ways more conservative phonologically than standard Burmese; for example, Arakanese retains the distinction between [r] and [j], Tavoyan keeps medial [l], and so on.

Since 1990 the government of Burma has changed the English name of the country to Myanmar; this is the name of the country in literary Burmese. The English name of Burma is derived from the spoken Burmese name for the country. The government distinguishes those who are Burman by race, now calling them Bamar, as opposed to members of other indigenous races of Burma, who are now called Myanmar; the Burmese language is also called Myanmar. This new terminology has not been accepted outside Burma, except by those who view the SLORC military government as politically acceptable.

7.1.3.2 OTHER BURMISH LANGUAGES

To the north-east of the Burman area is a hill area with some inhabitants speaking closely related languages. There are four main groups, all to some degree integrated into the Kachin or Shan cultures of the surrounding majorities. These groups go by different names as shown below:

Autonym	Jinghpaw name	Burmese name	Chinese name
Lawngwaw	Maru	Maru	Langsu
Tsaiwa	Atsi	Zi	Tsaiwa
Lachik	Lashi	Lashi	Lachi
Ngochang		Maingtha	Achang

Of the main four, the most scattered is the Achang, with about 30,000. In China the bulk of this group formed a separate Shan-like valley state and are recognised as a separate nationality, but in Burma most are mingled with the Lashi; a few, known as Tai Sa, are blacksmiths scattered in Shan villages over a wide area of the southern Kachin State in Burma (Jerry Edmondson, personal communication). The Lashi number some 30,000, mostly in Burma. The 100,000 Maru are widely dispersed, as are the 150,000 Atsi. For most purposes the Atsi, Maru, Lashi and some Achang operate as clans within the Kachin culture complex, bilingually speaking Jinghpaw as their literary language and intermarrying. Even smaller clan-based groups, Bola and Chintau, are found only in China among the Atsi and Achang respectively; there may be others. Roman orthographies for Atsi, Maru and Achang exist but have just started to be used. Within each group the regional differences are substantial; perhaps even as great as those between different Burmish groups living together and intermarrying. Maru and Lashi are characterised by the addition of final stops to some syllables; these are absent in Atsi and Achang. In China, most members of the 'Jingpo' nationality are Atsi, with smaller numbers of Maru, Lashi, Bola, and Chintau as well as some speakers of Jinghpaw; in Burma the proportion of the 'Kachin' who speak Burmish home languages is smaller, but still substantial, especially in the Northern Shan State and the south-east of the Kachin State. There are also very small communities in Thailand. Xu (1984) describes Atsi, Dai (1985) describes Achang.

The final Burmish language is Phun (Hpun, Hpon, Hpön), spoken in the gorges of the upper Irrawaddy north of Bhamo; it has two dialects, north and south. Both are extremely moribund, and may be nearly extinct; at the most there are a few hundred speakers left. It is phonologically more conservative than the 'Kachinised' Burmish languages.

7.1.4 THE LOLOISH LANGUAGES

For details of the phonological and lexical subgrouping of these languages, see Bradley (1979a). Basically, all share an innovative two-way tonal contrast in original stop-final syllables; the Northern Loloish languages have mostly reversed phonetic values for these two tones compared to other Loloish languages. Central Loloish is characterised by extensive tonal splits leading to complex tonal systems including contour tones.

In China, the 'Yi' nationality includes six languages, three of which (Nosu or Northern Yi, Nasu or Eastern Yi and Nisu or Southern Yi) are closely related to each other and form Northern Loloish. The remaining three, South-eastern Yi (including the Sani, Axi, Azhe and Azha), Central Yi (Lipho and Lolopho) and Western Yi (including Lalo and others), form part of Central Loloish. None of these six languages is spoken outside China (apart from a few Southern and South-eastern Yi in northernmost Vietnam), but most other Loloish languages are.

7.1.4.1 NORTHERN LOLOISH

The three languages in the Northern Loloish group have a traditional autonym derived from *ni*, with various forms according to sound changes; they also have a more recent group name derived from *nak su* 'black people'. All three, as well as the South-western Yi, had a character-based traditional script known to traditional religious practitioners. Revised versions of these scripts are now being disseminated; for details see Bradley (forthcoming).

Nosu is the largest group speaking a Loloish language; there are about two million speakers in the Liangshan (cool mountains) area of southern Sichuan province, and about 250,000 more, mainly in the Xiaoliangshan (small cool mountains) area of north-western Yunnan province, but also a few in north-eastern Yunnan. It has three main dialects: northern, with subdialects Tianba (north-western) and Yinuo (north-eastern); central Shengza; and south-eastern Sondi with subdialects Sondi and Adur; the Sondi subgroup is fairly distinct from the other two, which are more closely related. The largest group, over half of the total, is the Shengza, and the local variety of Shengza spoken at Xide county (a Chinese placename) has been selected as the standard, with a new syllabic orthography based on the traditional Nosu characters in extensive use since 1978; a romanisation is used occasionally.

Very closely related to Nosu but not mutually intelligible, the Nasu group is spread throughout western Guizhou and north-eastern Yunnan, with about 6,000 in north-western Guangxi. Of approximately 850,000 speakers, more than half are in Yunnan; there are several hundred thousand more non-speakers, mostly young people in all three provinces, who are members of the group. Dialect differences are very substantial and complex; there are three main subgroups, south-eastern (known also as Panxian from the county in south-western Guizhou where they are concentrated; about 150,000 speakers); north-eastern (most of the other Nasu in Guizhou, and some in extreme north-eastern Yunnan and south-eastern Sichuan; with subdialects named after the four traditional Nasu kingdoms of the area, Shuxi, Wusa, Mangbu and Wumeng, about 300,000 speakers); and western (all in north central Yunnan, about 250,000 speakers with two subdialects: Black and much less numerous Red). The traditional characters are being brought back into use in Guizhou; separate Pollard (missionary) scripts existed for Black Nasu, Laka and Gepo, and these still enjoy some very limited use. In Guizhou the traditional script is being promoted; in Yunnan a new combined Yi script has been created and is starting to be used by the Nasu, Nisu and South-eastern Yi groups. In addition to the Nasu there are several small groups speaking closely related but distinct languages including about 30,000 Laka (sometimes known to the Chinese as Gan Yi), about 90,000 Kepo (found in the western missionary literature as Köpu), Sami (known to the Chinese as Samei, and spoken just south-east of Kunming by about 10,000, mainly older speakers). The moribund language Sangwie or Sanie is spoken by older members of the Bai Yi (White Yi) to the west of Kunming; there are about 20,000 members of the group, with perhaps half this number of speakers. There are probably other moribund related languages in this area.

The Nisu or Southern Yi are also rather diverse; in addition to about 600,000 speakers (including 3,200 in Vietnam where they are still known as Lôlô), there are several hundred thousand non-speaker members of the group. Only very limited descriptions are available.

Smaller related groups speaking distinct languages include the Phula, with over 100,000 speakers scattered over south-eastern Yunnan and 6,500 in north-eastern Vietnam; the

Gazhuo, spoken by about 4,000 members of the Mongol nationality just north of Tonghai county in Yuxi, south of Kunming; the Kathu of Guangnan county in extreme eastern Yunnan; and the 5,000 Mo'ang [muɑŋ^{51}] of Funing county in south-eastern Yunnan and a few in adjacent areas of Guangxi.

7.1.4.2 CENTRAL LOLOISH

The first subgroup of Central Yi comprises four named groups, Sani, Axi, Azha and Azhe, speaking very similar dialects; the Chinese offical classification groups these as 'South-eastern Yi'. They number about 400,000 in south-eastern central Yunnan. The best described is Sani with about 200,000 speakers south-east of Kunming; see Ma (1951). A new Sani syllabary based on traditional characters was recently introduced in Lunan County. This is in competition with the newer Yunnan-wide Yi script. There are some 150,000 Axi, mainly in Mile county south-east of Lunan. The 100,000 Azhe are further south, with about 50,000 Azha to the south-east.

The Lisu are a large and widely dispersed group totalling about 850,000, with 575,000 in China (most in Yunnan, especially the north-western part, but about 13,000 in southern Sichuan as well); nearly 250,000 in Burma, mainly in the north; about 25,000 in Thailand, and a couple of thousand in north-eastern India, where they are known as Yobin from the former Burmese name Yawyin. Dialect differences are substantial; the Thailand dialect as described in Hope (1974) has extensive Chinese loans. Various orthographies exist; those now used include the 'Fraser' script, devised by a missionary of that name, which uses upper-case roman letters, upright and inverted, and punctuation marks for tones; and the new Chinese romanisation, which uses numerous digraphs, as well as consonants after the vowel to indicate tones. Use of the latter is on the decline.

The names Lipho and Lolopho are used to refer to groups classified by the Chinese as 'Central Yi' who speak a language which is linguistically very close to Lisu; some scholars have even regarded them as an eastern dialect of Lisu. There is a total of about 450,000 speakers and some non-speaker group members, mostly young people, in north-western central Yunnan, surrounded by Nosu or Northern Yi on the north, Nasu or Eastern Yi to the east, Nisu or Southern Yi to the south, and Lalo or Western Yi to the west; hence their designation as 'Central Yi'. Some of the Lipho use a Pollard (missionary) script; there was no traditional Lipho/Lolopho script.

The Western Yi subgroup includes Lalo and various other named groups, with a total of about 300,000 speakers. No full description of any variety is available, and in many areas the language is being replaced by Chinese.

The Lahu total about 650,000, with 360,000 in China, over 200,000 in Burma, 60,000 in Thailand, and about 10,000 in Laos; there is also a small group of about 1,000 refugees from Laos in the United States. The main dialect difference is between Black Lahu (Lahu Na) and Yellow Lahu (Lahu Shi), but there are extensive smaller differences within each; for details see Bradley (1979b). This is one of the best described TB languages, mainly due to the work of Matisoff, for example Matisoff (1982).

There are about 55,000 Kucong (Chinese name), mainly scattered south of the Red River in southern central Yunnan, who call themselves Lahu but speak a rather distinct language. In Vietnam they are sometimes known as Cosung and sometimes as Lahu. The Kucong were

officially amalgamated into the Lahu nationality in China in mid-1989. There are nearly 50,000 in China and 5,400 in Vietnam.

The Jinuo group of about 20,000 speakers, who live in one area of south-western Yunnan, was recognised as a separate nationality in China only in 1979, the most recent 'new' nationality in China. It is described in Gai (1986). On its position in Central Loloish, see Bradley (1983). There is no orthography.

The Zauzou group of about 2,500 is classified as part of the Nu nationality in north-western Yunnan, China; but their language is Loloish. Based on limited available data it appears to be Central Loloish.

7.1.4.3 SOUTHERN LOLOISH

As for the Northern Loloish 'Yi' nationality, there is a Southern Loloish 'nationality' in China that includes a variety of distinct languages: Hani. Chinese linguists break this into three main subgroups: Hani-Akha, Haoni-Baihong, and Biyue-Kaduo. The first two are fairly closely related within the Akoid subgroup of Southern Loloish, but the third forms a separate cluster of languages.

7.1.4.3.1 AKOID

In the Akoid group are the 550,000 Akha – about 35,000 each in northern Thailand and northern Laos, some 250,000 in southernmost China, and about 220,000 in Burma; it is well described in Lewis (1968). Dialect differences within Akha are fairly minor, apart from the few thousand Akeu whose language is quite different from Akha. According to Akha traditions, in which the ritual group name is [za^{21}ni^{21}], they migrated from the north-east; to the north-east, the roughly 520,000 Hani [xa^{21}ni^{21}] in southern central Yunnan and another 12,500 in Vietnam speak a cluster of very similar varieties, more or less mutually intelligible with Akha. To the north-west of the Hani (upriver) are about 120,000 Haoni [xɒ^{21}ni^{21}] who speak a variety quite similar to Hani, and further north some 60,000 Baihong, whose language is somewhat less similar. Some of the most northerly 'Hani' in China no longer speak anything other than Chinese. Within this cluster there are also several smaller groups: the Phana of Laos (about 500) and the Sila or Sida of Laos and Vietnam (about 2,000). For Akha there are three competing romanisations: Catholic, Baptist and a third based on the Hani romanisation. The Hani script now used in China is a romanisation using the principles of pinyin. For Hani, this was first proposed in 1957 but introduced (in revised form) only in 1984.

7.1.4.3.2 BI/KA

This subgroup is named from the first syllables of the Chinese terms for its two main components: Biyue (autonym Piyo) and Kaduo (autonym Khatu). Nearly 300,000 of these are included in the 'Hani' nationality of southern Yunnan; a small number of Khatu live in northernmost Laos as well. There is a third related group of 1,500, the Mpi, who live in a village just east of Phrae in northern Thailand; their tradition reports that they were brought as war captives from the north about 200 years ago – presumably from the Piyo area. Mpi is

described in Duanghom (1976). All Mpi are now bilingual in Northern Thai and culturally assimilated, but the village maintains the language.

7.1.4.3.3 BISOID

The two well-known Bisoid languages, Bisu of northern Thailand and Phunoi of northern Laos (also known as Công in north-western Vietnam), share the development of voiced stops corresponding to some initial nasals in other TB languages; by contrast, they are the most conservative Loloish languages for finals including nasals. Dialect differences within Bisu and Phunoi are substantial; among the 500 or so Bisu in Thailand there are distinct dialects for each of the four villages, and five dialects within Phunoi are reported. About 6,000 Bisu live in south-western Lancang county in south-western Yunnan, whence the Bisu of Thailand were probably brought in about the 1850s. A language called Pyen was moribund when last reported in the 1931 census in the Shan State of Burma, and may now be extinct; from the limited data available, it appears that it was another dialect of Bisu. Công is one of the smaller nationalities of Vietnam, with only about 1,300 people; linguistic data is virtually unavailable, but it appears to be quite similar to Phunoi. There are over 28,000 Phunoi in Laos, so the total Phunoi/Côông population is about 30,000. A distinct Bisoid language, Sangkong, was recently described by Li (1992); it is spoken by about 2,000 people of Hani nationality in Jinghong county.

7.2 KAREN

The Karen languages, mostly in Burma but extending into western Thailand, are clearly TB but have a number of distinct characteristics. The main one is that all are SVO, the only such languages in TB. Extensive data and a reconstruction are provided in Jones (1961). There is also a substantial non-TB element in the lexicon, as demonstrated in Luce (1985). Benedict (1972) treated Karen as a coordinate subgroup with TB and Sinitic within Sino-Tibetan, but has more recently (Benedict 1976) stated that Karen appears to be within TB, and indeed close to the Burmic portion of TB. Shafer (1974) places Karenic within TB as a separate group coordinate with Burmic, Bodic and Baric.

I am most grateful to R.B. Jones, F.K. Lehman and D.B. Solnit for extensive personal communications which have been essential to the preparation of the information on Karen languages. There is considerable disagreement on the subgrouping of Karen, which includes a number of languages, many with several alternative names.

Karen population was reported at 2,122,825 in Burma in 1983 and 292,814 in Thailand in 1992; Kayah are separately counted at 141,028 in Burma. All of these figures are substantially underenumerated; many Karens try to pass for members of the majority group, Burman or Thai, and many in Burma live in areas not directly controlled by the central government. Moreover many of the smaller central and northern Karen groups are not included in these totals. The Karen who are scattered in the Lower Burma delta region, enumerated at about 1.5 million, are mostly Sgaw with a smaller proportion of Pho; many of these are in the process of becoming Burmanised. The 'official' balance in Burma, about 0.8 million, may safely be doubled, with several hundred thousand Pa-O, Padaung and others giving a total well over three million for Burma, or close to four million including Thailand. There is also a small Karen community in the Andaman Islands.

```
Karenic        Northern ─────────────── Pa-O (Taungthu)
                                         Padaung/Kayan
                                              Yingbaw/Ka-ngan
                                              Zayein/Latha
                                              Gekhu

               Central/Bwe
                         Eastern
                                 Kayah
                                 Manu
                                 Yintale
                         Western
                                 Blimaw
                                 Bre
                                 Geba

               Southern              Sgaw
                                         Mopwa/Palaychi
                                         Paku
                                         Wewaw
                                         Monnepwa
                                     Pho/Pwo/Phlong
                                     Lekhe
```
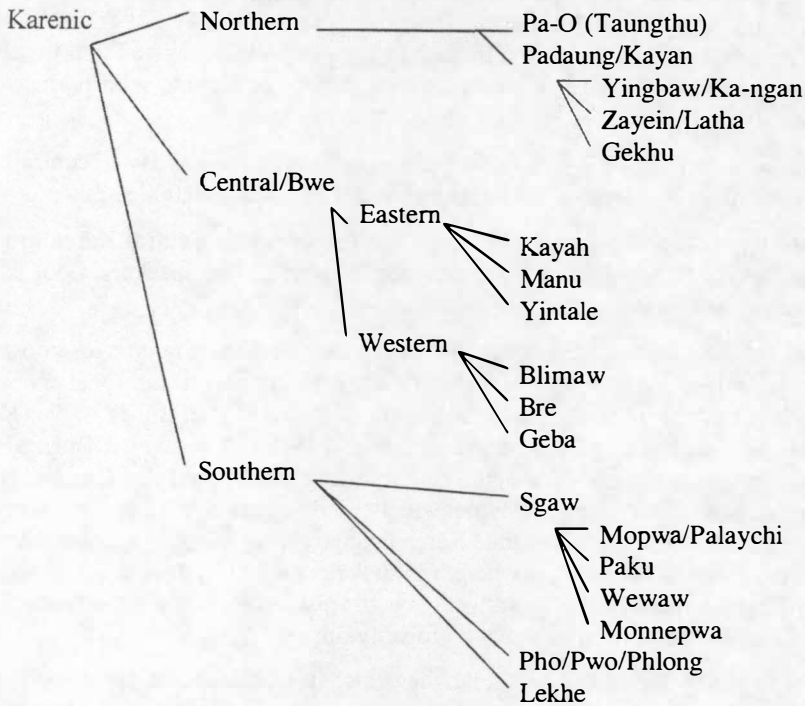
Jones (1961) suggests a subgroup which includes Pa-O, Pho and Lekeh versus the rest, which he divides into Sgaw and its dialects versus a central group with three subgroups: Padaung, Eastern Bwe and Western Bwe. Solnit and Lehman share the classification as shown above, linking Pa-O and Padaung in Northern and Sgaw with Pho in Southern. Lehman differs from Solnit and Jones in grouping Gekhu with Bwe rather than Padaung.

The Pa-O were formerly called Taungthu (Burmese for 'hill people') and Shan Tonghsu; these names are now regarded as pejorative. This group probably totals over half a million speakers.

The Padaung are the group whose women traditionally wear neck and knee rings. The new autonym (replacing somewhat pejorative Padaung) is Kayan; also included here linguistically are Yinbaw (autonym Ka-ngan), Zayein (Latha), Gekhu ('upper', also seen as Gheko, Geko, Gekho, etc.) and probably Sawntung, with a total of about 85,000 speakers.

Sometimes still known as Karenni or Red Karen, the Kayah group has substantial dialect differences between east and west. It is a separate nationality with a separate state in Burma; there are some speakers in north-western Thailand as well. Officially it has over 140,000 speakers in Burma, but this is underenumerated (and probably includes Manu, Yintale and perhaps some other Karen langages within the Kayah State). Including Thailand, there are probably a quarter of a million speakers.

For the small western subgroup Manu, the Burmese name is Manumanaw, the Kayah name is [puɯnuɯ]. This means 'western' (dialect of Kayah); it may have 10,000 or more speakers. The name of the subgroup Yintale has been folk-etymologised into the Burmese Yin-Talaing or Mon Karen. It is a south-western variety of Kayah, with perhaps 10,000 speakers.

The Blimaw subgroup of Karen is sometimes simply known as Bwe 'central' Karen. There is a dictionary, Henderson (forthcoming), and it has about 20,000 speakers.

Known to its speakers as [brɛʔ] and hence Bre or Brek, this Central Karen group now prefers the autonym Kayaw [kəjɔ]; it probably totals about 25,000 speakers. Geba is another Central Karen group, with about 10,000 speakers.

Sgaw, in the Southern Karen subgroup, is the largest Karen language, known to its speakers as [sɣɔʔ] from the word for 'person'; it has about 1.6 million speakers, with just under half in the delta region of Lower Burma and nearly two-thirds of the Karen in Thailand. It is fairly well described in Gilmore (1898). There is a Burmese-based orthography devised in the mid-nineteenth century using extra vowel and tone symbols, and a traditional 'chicken track' orthography not widely used anymore. Various divergent dialects have appeared in the literature as separate Karen languages; for example Mopwa (also known as Palaychi from one of the villages where it is spoken; studied by Jones) in the north-west, Paku [pakuɯ], which is the Kayah word for Sgaw and also refers to a north-eastern dialect of Sgaw, as well as Wewaw, Monnepwa and probably others.

Briefly reported by Jones, the Lekeh language also has a traditional orthography derived from Burmese but not widely known or used. The number of speakers is unknown; some at least live in the delta region around Rangoon.

To its speakers known as Sho, to the Sgaw as Pwo or Pho, and in Thailand (from the word for 'person' there) as Phlong, Pho Karen has about 1.4 million speakers, with nearly one-third of these in the delta region, and about one-third of the Karen in Thailand. This language extends quite far to the south, virtually to Burma's southern extremity, and onto some adjacent islands. There are very substantial dialect differences; the standard dialect has an orthography derived from Burmese, devised shortly after the Sgaw orthography and using different conventions for the vowels and tones of Pho. Duffin (1913) provides a good description.

8. CONCLUSION

Altogether there are nearly 250 known TB languages, with about 65 million speakers. Doubtless some others remain to be 'discovered', especially in China and Burma. Many others are already dead; in some cases there are written records of these dead languages, but others have completely disappeared. Quite a large number are endangered; so salvage work is urgent.

Of all TB languages, the one with the largest number of speakers is Burmese, which is about half the total; another ten million people use it as a second language. The following table shows the top ten TB languages in terms of number of first language speakers.

	Language	Speakers	Official status	Location
1.	Burmese	32 million	Burma	Burma
2.	Tibetan	4.6 million	Tibet, Bhutan	Tibet, India, Nepal, Bhutan, etc.
3.	Karen	3.9 million	Karen/Kayah states	Burma, Thailand
4.	Nosu	2.25 million	Liangshan	China
5.	Lisu/Lipo	1.3 million	Nujiang	China, Burma, Thailand, India
6.	Bai	1.1 million	Dali	China
7./8.	Meithei	1 million	Manipur	India
7./8.	Bodo	1 million		India
9.	Nasu	800,000		China
10.	Lahu	650,000	Lancang	China, Burma, Thailand, Laos

The totals for Tibetan and Karen include speakers of a very wide variety of languages and dialects who feel a cultural unity. In the case of Tibetan this is founded on Tibetan Budhism and the written Tibetan language. For the Karen, this unity is more tenuous; 3 million Karen speak either Sgaw or Pho Karen, each of which has a long-standing Burmese-based orthography, and the rest speak a variety of distinct languages in the Kayah State and to the north.

The development of orthographies for TB languages has been going on for well over a millenium. On the one hand, some scripts are based on an Indic model. These include the Pyu script (seventh century AD), the Tibetan script (eighth century AD), the Burmese script (1112 AD) and the old Newari script (1113 AD). Those based on a Chinese model, and thus using the character principle, include Xixia (1037 AD), Nasu/Nosu (undated but probably at least 600 years old) and Naxi. During the last century, roman scripts have been devised for many TB languages; initially by missionaries, and since the 1950s in China following the principles of pinyin, the official romanisation for Chinese. Indic models continue to be used for new scripts in the South Asian region, alongside roman scripts.

Another fascinating phenomenon is language contact and language shift among related TB languages. Many non-Tibetan TB languages of the northern and eastern Himalayas are in close contact with and are influenced by Tibetan; all TB languages of Burma are influenced by Burmese; and so on. Some languages are being or have been absorbed into others; notably in the Rai area of eastern Nepal and the Kuki-Chin-Naga area of north-eastern India.

Wurm, Mühlhäusler and Tryon (forthcoming) summarise the use of TB languages as lingua francas; in many areas speakers of various TB languages use another TB language for interethnic communication. Such languages include Tibetan (literary High, Lhasa spoken, Khams, Dzongkha), Burmese, Lahu, Mizo and others. Speakers of TB languages in many countries use non-TB languages for this purpose: Nepali in Nepal, Darjeeling and Sikkim, Nagamese in Nagaland, English in Meghalaya and elsewhere in the hills of north-eastern India, and the national languages in Thailand, Laos, Vietnam, China, India and Pakistan.

Another interesting sociolinguistic phenomenon is the development of diglossia in several TB languages. The two principal examples are Burmese and Tibetan. In each case the literary High is more archaic, and is used in some formal spoken contexts as well as nearly all written contexts. The spoken Low shows a great deal of dialect diversity in Tibetan, and somewhat less in Burmese.

Official policy or indigenous unification movements have resulted in the amalgamation of a variety of more or less closely related TB languages into new ethnicities, sometimes with a standard written language. This has happened in several cases in south-western China (the Yi, the Hani and the Lahu among others), in north-eastern India (various Naga groups, the Mizo and so on), and may be starting among the Rang of northern Uttar Pradesh in north-western India. Conversely recent political considerations have divided some TB groups, such as the Balti and Purik in Kashmir, or various groups which live across the borders of China and adjacent countries, such as the Kucong, the Bisu/Phunoi/Công and numerous others.

The politics of language and language policy is another fascinating issue for many TB groups. In Burma and Bhutan, TB languages have national-level official status. Tibetan is the official language of the Tibetan government in exile and is widely used in the portions of Tibet under Chinese control. Manipuri has long been a state language in Manipur, and has recently achieved Schedule VIII status in the constitution of India, conferring on it official status at the national level. Newari was the court language of the Malla kingdoms of the Kathmandu area until they were conquered in the late 18th century by the Nepali speakers of Gorkha. At a much earlier stage, some now-extinct TB languages such as Xixia in what is now western central China, and Pyu in what is now central Burma, had a similar status.

Recent changes have led to a broadening in the use of TB languages in education. In China, most minority languages are used to some extent in education and administration, especially in designated 'autonomous' regions, prefectures and counties of that minority. In Sikkim state, the indigenous Lepcha and Limbu as well as the Danjong variety of Tibetan are used in education up to matriculation. Most other states of India with a large TB-speaking population have chosen to use English for official purposes; but education and public life using TB languages is widespread in Meghalaya, Mizoram, Arunachal Pradesh and Nagaland. In Bhutan, official and educational status have been accorded to Dzongkha, another variety of Tibetan. Since political liberalisation a few years ago, Nepal has again begun to use some TB languages such as Newari on radio, reversing the previous policy of Nepali for all purposes. Burma has accorded some status to the languages of its seven states including Arakanese (a dialect of Burmese) as well as TB Chin, Kachin (Jinghpaw), and the Kayah and (Sgaw/Pho) Karen along with non-TB Shan and Mon; however education and public life continues to be in Burmese.

Political movements and rebellions seeking independence or autonomy for certain TB groups have flared into violence continually since the 1950s. These movements include the Naga and Mizo in India, and the Kachin, Karen, Arakanese and many others in Burma. In north-eastern India over the last ten years there has been agitation, sometimes violent, for official status for Bodo. Similar tension, mainly caused by competition for increasingly scarce land and other resources, has arisen between TB groups such as the Naga and Kuki in Manipur and Nagaland. Tension between indigenous Kokborok and migrant Bengalis also simmers in Tripura, parallel to the more extreme levels of conflict between the Bengalis and indigenous TB and non-TB groups of Meghalaya and Assam. Recent movements of Rohingya (Bengali Moslem) refugees out of Burma back to Bangladesh and of Nepalis back from Bhutan and north-eastern India to Nepal, Darjeeling and Sikkim are a direct result of indigenous hostility to encroachments by Indic language speakers into traditional TB territory. Sometimes these conflicts also have a religious aspect; many hill dwellers in north-

eastern India and Burma are now Christian, while the Nepali migrants are Hindus and the Bengalis or Rohingyas are Moslems in Hindu Assam and Buddhist Burma.

In China the process of achieving recognition for one's ethnic group is also political, but usually non-violent; groups may apply for recognition as a separate national minority, and one TB group, the Jinuo, successfully did so in 1979. Other groups in China have been less successful; the Kucong also applied, but were instead amalgamated with the Lahu in 1989. Other applications are pending. It is also possible to change one's nationality in China. Because of advantages for minorities, many people did so between the 1982 and 1990 census, and so the population for many nationalities increased substantially due to these additions as well as natural increase. Sometimes the changes are because of previous misclassifications; thus, a substantial number of Lipo changed from Yi to Lisu nationality in the mid-1980s.

There are still some unresolved problems in the classification of TB languages. At the more macro level, issues such as the position of Kuki-Chin-Naga are not fully resolved; and the exact status of Central TB is uncertain. Sun (1993) casts doubt on the position of Keman in this group, and Lepcha has long been a bone of contention. The other West Arunachal languages also need more classificatory research. Some of the outliers of North-eastern TB are also of somewhat uncertain status: is Naxi closer to Burmese-Lolo, as some of its lexicon indicates, or to NE TB? Where do Bai and Tujia fit? Is Baima a Tibetanised NE TB language or a variety of Tibetan?

In general, there remains a great deal of linguistic research to be done on TB languages: basic descriptive work; sociolinguistic studies of the process of language contact, convergence, shift and death; and comparative work.

APPENDIX

NAMES, SPEAKERS AND LOCATIONS OF TIBETO-BURMAN LANGUAGES

1. INTRODUCTION

Names shown in the left column are the usual names found in the recent literature and referred to in the classification above. Where the group has an autonym which is different, this is given next. Other names which have been used for the group are given immediately prior to the population; some of these are names used by other groups or geographical names. These different names are discussed in the text above.

Population totals are 1995 estimates based on census and other information; totals given in bold are for speakers of languages. In many cases this is fewer than totals for members of the corresponding ethnic group due to language shift. For various groups in China, it is fewer than census figures for nationalities, due to misclassification of part or all of a group as members of another nationality, mainly as Tibetan. Where data on the number of speakers of dialects is available, totals for speakers of dialects are given in non-bold.

Countries where the languages are spoken are listed in order of population for each group.

B	Burma
Ba	Bangladesh
Bh	Bhutan
C	China
L	Laos
N	Nepal
P	Pakistan
r	refugees in Western countries
T	Thailand
Ti	Tibet (traditional area, including Qinghai and parts of Gansu, Sichuan and Yunnan)
V	Vietnam

The numbers given with each heading below correspond with the sections in the paper in which the languages/dialects are discussed.

2. BODIC

2.1 BODISH

2.1.1 CENTRAL

Name	Autonym	Other names	Population	Location
Balti		sBalti	300,000+	Pl
Purik d		Purki	45,000	I
Ladakhi		Ladwags	75,000	I
Zangskar d.			5,000	I
Tod		'Lahuli'	1,700	I
Ranglo		Khoksar	700	I
Spiti	Piti		12,000	I
Nyam		mNyam	3,000	I
Jad			1,500	I
mNgahris	(numerous varieties)	Ngari	50,000	Ti N
gTsang	(numerous varieties)		600,000	Ti N I
dBus	(numerous varieties)	Ü	900,000	Ti I r Bh
Southern	(several varieties)		550,000	Bh I Ti
Amdo	(numerous varieties)		900,000	Ti
Khams	(numerous varieties)		1,600,000	Ti B

2.1.2 WESTERN (GURUNG, TAMANG)

Name	Autonym	Other names	Population	Location
Kutang Ghale			**1,300**	N
Ghale			**15,000**	N
Kaike			**2,000**	N
Dura			few	N
Gurung	Tamu		**150,000**	N I
Thakali	Tapaang	Thaksya	**5,000**	N
Chantel			**2,500**	N
Rohani			few	N
Manang		Nyishang	**3,000**	N
Tamang			**600,000**	N I

2.1.3 EASTERN (BUMTHANG, MONPA)

Name	Autonym	Other names	Population	Location
Bumthang			**30,000**	Bh
Khengkha			**40,000**	Bh
Kurtöpkha			**10,000**	Bh
'Nyenkha/Henkha		Mangdebikha	**10,000**	Bh
'Olekha Monpa			**1,000**	Bh
Chalikha			**1,000**	Bh
Dakpakha			**1,000**	Bh
Dzalakha		Cuona Monpa/Central Monpa	**52,000**	I Bh Ti
Eastern Monpa		Motuo Monpa	**5,000**	I Ti
Sherdukpen	Mei		**4,000**	

2.1.4 TSHANGLA

Name	Autonym	Other names	Population	Location
Tshangla		Southern Monpa	**140,000**	Bh
Lhokpu		Lhobikha	**2,500**	Bh
Gongduk		Gongdubikha	**2,000**	Bh

2.1.5 WEST HIMALAYISH

Name	Autonym	Other names	Population	Location
Pattani		Lahuli, Manchati	**14,000**	I
Tinan		Gondhla	**2,500**	I
Bunan		Gahar, Gahri	**5,000**	I
Kanauri		Kinnauri	**60,000**	I
Chitkhuli d.			**1,000**	I
Kanashi		Malana	**1,100**	I
Rangkhas		Marchha	**7,500**	I
Darmiya			**4,000**	I
Chaudangsi/Byangsi			**8,000**	I N
Bhramu			**extinct**	N
Thami			**14,000**	N

2.2 HIMALAYAN

2.2.1 CENTRAL

Name	Autonym	Other names	Population	Location
Magar			300,000	N
Raute/Raji			500	N I
Kham			40,000	N
Chepang			17,000	N
Newari			600,000	N

2.2.2 KIRANTI

Name	Autonym	Other names	Population	Location
Hayu			100	N
Sunwar (incl. Surel d.)			25,000	N
Bahing		Rumdali	10,000	N
Chaurasia		Umbale	5,000	N
Jerung			2,000	N
Thulung			25,000	N
Khaling			12,000	N
Dumi			2,000	N
Kohi		Koi	300	N
Kulung		Kulunge	9,000	N
(Sotang d.)			6,000	N
Nachering (incl. Parali d.)			2,000	N
Chukwa			100	N
Sangpang		Sangpahang	7,000	N
Bantawa			70,000	N
Lambichong		Mugali	1,000	N
Dungmali/Khesang			5,000	N
Chamling		Rodong	30,000	N
Puma			3,000	N
Athpare			2,000	N
Belhare			1,000	N
Chintang			100	N
Chulung			1,000	N
Yakkha			10,000	N
Lumba			1,000	N
Mewahang		Newahang	4,000	N
Lohorong (N vs. S dialects)			15,000	N
Yamphu			5,000	N
Yamphe	Yakkhaba		5,000	N
Tilung			1,000	N
Chatare Limbu			30,000	N
Limbu			200,000	N I
Dhimal/Toto			10,000	N I

3. NORTH-EASTERN INDIA/SAL

3.1 BOBO-GARO

Name	Autonym	Other names	Population	Location
Boro		plains Kachari	1,000,000	I N Ba Bh
Lalung			20,000	I
Dimasa		hills Kachari	50,000	I
Hojai			10,000	I Ba
Kokborok		Tripuri	400,000	I Ba
Garo			500,000	I Ba
Atong			10,000	I Ba
Wanang			15,000	I Ba
Rabha			50,000	I
Hajong			10,000	I
Deori			15,000	I

3.1.1 NORTHERN NAGA

Name	Autonym	Other names	Population	Location
Khienmungan	Kalyokengnyu		50,000	B I
Chang			25,000	I
Phom			25,000	I
Konyak			90,000	I B
Htangan			20,000?	B
Wancho			40,000	I B
Haimi			60,000	B
Nocte			40,000	I B
Tangsa/Rangpan			40,000	B I

3.2 Jinghpaw

Name	Autonym	Other names	Population	Location
Jinghpaw		Kachin, Singpho	600,000	B C I

3.3 LUISH/SAK

Name	Autonym	Other names	Population	Location
Kadu			20,000	B
Ganan			7,000	B
Taman			1,000	B
Sak		Thet	7,000	B Ba
Chakpa		Chairel	extinct	I
Phayeng		Andro	extinct	I
Sekmai		Sengmai	extinct	I

3.4 PYU

			Population	Location
			extinct	(B)

4. KUKI-CHIN

4.1 SOUTHERN NAGA

Name	Autonym	Other names	Population	Location
Ao			110,000	I
Sangtam (N)			30,000	I
Lhota			60,000	I
Yimchungrü			30,000	I
Rengma			15,000	I
Tangkhul			100,000	I B
Maring			15,000	I B
Sema			100,000	I
Angami (W)			60,000	I
Chakhesang			30,000	I
Mao			60,000	I
Maram			5,000	I
Zeliangrong			90,000	I
Meithei	Manipuri		1,000,000	I B

4.2 KUKI

Name	Autonym	Other names	Population	Location
Rangkhol			8,500	I Ba
Hallam			12,000	I Ba
Langrong			1,000	I
Hmar			30,000	I
Anal			12,000	I B
Kom			7,000	I
Chawte		Chote, Chaw, Kyaw	1,000	I B
Mayol		Moyon, Moyol	1,000	I
Lamgang			2,000	I

4.3 CHIN

Name	Autonym	Other names	Population	Location
North Chin			250,000	B I
Central Chin			400,000	B I Ba
Mizo			400,000	I B Ba
South Chin			225,000	B Ba
Khami/Khumi			55,000	B Ba
Mara		Lakher, Miram	20,000	I B

4.4 ARLENG

Name	Autonym	Other names	Population	Location
Arleng	Karbi	Mikir	350,000	I

5. CENTRAL

5.1 Lepcha	Rong		4,000	I

5.2 WESTERN ARUNACHAL

Name	Autonym	Other names	Population	Location
Sulung	Puroik		**3,000**	I Ti
Bugun	Khowa, Khoa		**1,700**	I
Dhammai		Miji	**5,000**	I
Hrusso		Angka, Aka	**4,000**	I
Bangru			**1,000**	I Ti

5.3 ADI-MISING-NISHI/ABOR-MIRI-DAFLA/TANI GROUP

Name	Autonym	Other names	Population	Location
Nishi/Bangni/Apa Tani	Dafla		**170,000**	I Ti
Hill Miri			**25,000**	I
Adi		Abor	**125,000**	I Ti
Mising	Mishing	plains Miri	**500,000**	I

5.4 DIGARISH 'MISHMI'

Name	Autonym	Other names	Population	Location
Idu		Chulikata Mishmi	**20,000**	I Ti
Taraon	Digaru	Taruang	**12,000**	Ti I B

5.5 KEMAN 'MISHMI'

Name	Autonym	Other names	Population	Location
Keman 'Mishmi'		Miju	**20,000**	I Ti B

5.6 RAWANG/NUNGISH

Name	Autonym	Other names	Population	Location
Trung		Dulong	**11,000**	C Ti
Zørwang	Jerwang	Tvluq	**15,000**	B
Dvru	Daru		**35,000**	B
Anung			**8,000**	C B
Nusu			**8,000**	C
Dvngsar			**15,000**	B
Mvtwang			**50,000**	B
Lungmi		Longmi	**30,000**	B

6. NORTH-EASTERN

Name	Autonym	Other names	Population	Location
Xixia		Tangut	**extinct**	(C)
Baima			**20,000**	Ti
Jiarong	Karu	rGyarung	**150,000**	Ti
Qiang (N)			**75,000**	Ti
Qiang (S)			**120,000**	Ti
Ergong			**45,000**	Ti

Name	Autonym	Other names	Population	Location
Zaba		Zhaba	**8,000**	Ti
Guichong	Yutong, Guquo		**7,000**	Ti
Muya			**15,000**	Ti
Ersu		Tosu, Lusu	**20,000**	Ti
Choyo		Queyu	**7,000**	Ti
Namuyi			**5,000**	Ti
Shixing	Xumi, Shihin		**2,000**	Ti
Pumi	Primi		**55,000**	Ti/C
Naxi		Nakhi	**210,000**	C
Moso		'East Naxi'	**40,000**	C
Tujia			**170,000**	C
Bai		Minchia	**1,100,000**	C

7. SOUTH-EASTERN

7.1 BURMESE-LOLO/BURMIC

7.1.1 Mru			**40,000**	B Ba
7.1.2 Gong	Ugong		**1,000**	T

7.1.3 BURMISH

Name	Autonym	Other names	Population	Location
Burmese			**32,000,000**	B
Arakanese		Mogh/Magh	**1,800,000**	B Ba I
Tavoyan			**400,000**	B
Beik		Merguese	**250,000**	B
Danu			**100,000**	B
Intha		Angsa, Inle	**90,000**	B
Taungyo			**40,000**	B
Yaw			**20,000**	B
Hpun		Hpon, Phun, Phon	**50 ?**	B
Achang	Ngochang	Maingtha	**30,000**	C B
Lashi	Lachik		**30,000**	B C
Atsi	Tsaiwa	Zi	**150,000**	B C
Maru	Lawngwaw	Langsu	**100,000**	B C
Bola			**1,000**	C
Chintau			**1,000**	C

7.1.4 LOLOISH

7.1.4.1 NORTHERN

Name	Autonym	Other names	Population	Location
Nosu (northern Yi)			**2,250,000**	C
Nasu (eastern Yi)			**800,000**	C
Shuxi, Wusa, Mangbu, Wumeng			300,000	C
Hei Yi, Hong Yi (Black, Red)			250,000	C

Name	Autonym	Other names	Population	Location
Panxian			150,000	C
Laka			30,000	C
Gepo			90,000	C
Sami			10,000	C
Samatau			50	C
Sangwie/Sanie		Bai Yi	10,000	C
Gazhuo			4,000	C
Nisu (southern Yi)			600,000	C V
Phula			100,000	C V
Kathu			1,000	C
Mo'ang			5,000	C

7.1.4.2 CENTRAL

Name	Autonym	Other names	Population	Location
Sani	Ni	'Gni Lolo'	200,000	C
Axi		'Ahi Lolo'	150,000	C
Azhe			100,000	C
Azha			50,000	C
Lalo ('western Yi')			300,000	C
Zaozou			2,500	C
'Central Yi'/ Eastern Lisu	Lipho, Lolopho		450,000	C
Lisu		Yawyin, Yobin	850,000	C B T I
Lahu		Lohei, Muser	650,000	C B T L
Kucung		Cosung	55,000	C V
Jinuo			20,000	C

7.1.4.3 SOUTHERN

Name	Autonym	Other names	Population	Location
Baihong		Mahei	60,000	C
Haoni		Putu/Budu	120,000	C
Hani			520,000	C V
Akha		Ikaw, Kaw	550,000	B C T L
Akeu			5,000	C B T
Phana		Bana	500	L
Sila		Sida	2,000	V L
Khatu		Kaduo	180,000	C L
Piyo		Biyue	120,000	C
Mpi			1,500	T
Bisu		mBisu, Misu	6,500	C T (B)
Phunoi		Công	30,000	L V
Sangkong			2,000	C

7.2 KAREN

Name	Autonym	Other names	Population	Location
Pa-O		Taungthu	500,000	B T
Padaung	Kayan		85,000	B
Kayah		Karenni	250,000	B T
Manu	Punu	Manumanaw	10,000	B
Yintale		Yangtalai	10,000	B
Blimaw		Western Bwe	20,000	B
Bre	Kayaw	Brek	25,000	B
Geba			10,000	B
Sgaw			1,600,000	B T
Pho		Pwo	1,400,000	B T
Lekeh			?	B

In preparing the Map 1 to 8, I have benefited from the assistance and advice over many years of a very large number of colleagues in India, Nepal and elsewhere. Both the language data and language map of Bhutan have been taken from George van Driem's writings on Bhutan (e.g. van Driem 1992), and parts of the Nepal language map and the Nepal language data have been reproduced here from George van Driem's yet-to-be published handbook on the languages of the Himalayas (van Driem forthcoming), a preliminary version of which he kindly sent to me at my request. Many other sources have been useful, including various volumes of the Census of India (1981), Hansson (1991) concerning eastern Nepal, Webster (1992) concerning north central Nepal, Brauns and Löffler (1990) concerning south-eastern Bangladesh, Zoller (1983) and Sharma (1982, 1988, 1989a, 1989b, 1990, and 1992) concerning northern Uttar Pradesh and Himachal Pradesh in India, Moseley (1994) concerning areas adjacent to Burma, the various other sources cited above, as well as personal communications from colleagues too numerous to mention. Naturally any misinterpretations and inaccuracies are my own responsibility.
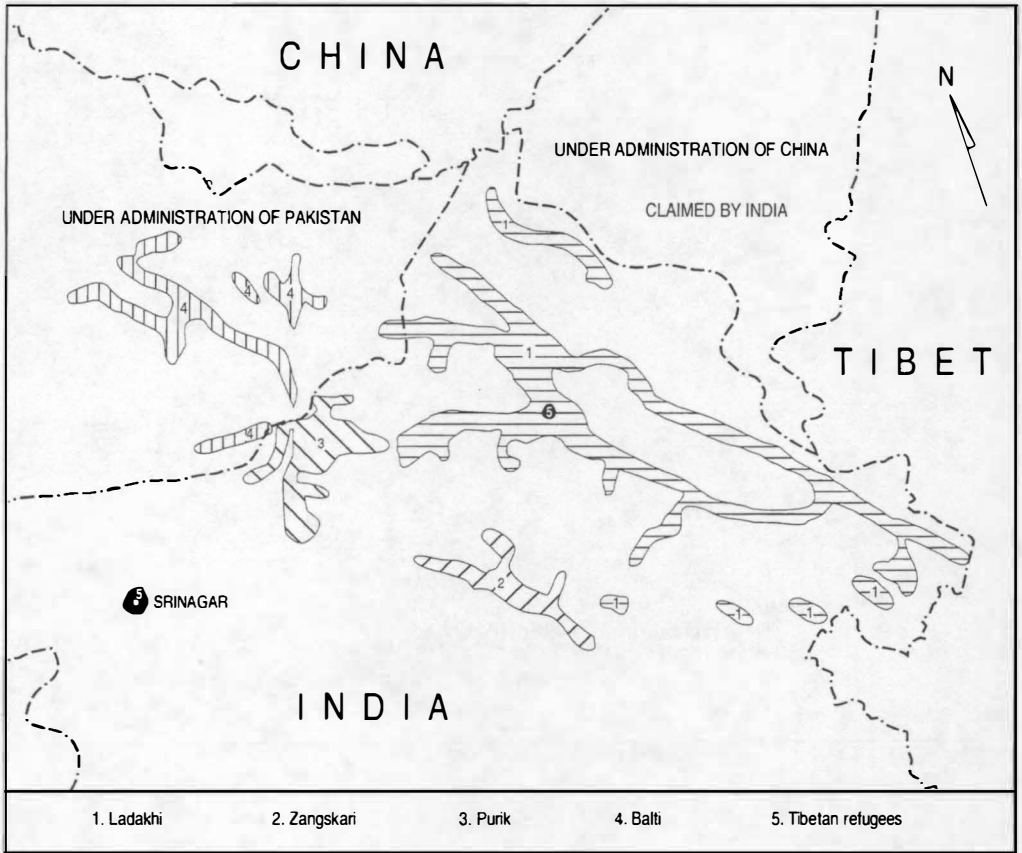
REFERENCES

All Bodo Students' Union, 1987, *Why separate state?* Kojkrahar: All Bodo Students' Union.

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

Baruah, Nagendra Nath, 1992, *A Trilingual Dimasha-English-Assamese dictionary*. Guwahati: Assam Publication Board.

Bauman, James J., 1976, An issue in the subgrouping of the Tibeto-Burman languages: Lepcha and Mikir. Paper presented at International Sino-Tibetan Conference, Copenhagen.

Benedict, Paul K., 1972, *Sino-Tibetan: a conspectus*. Cambridge: Cambridge University Press. 1976. Sino-Tibetan; another look. *Journal of the American Oriental Society* 96/2:167-197.

Bernot, Lucien, 1967, *Les Cak: contribution à l'étude ethnographique d'une population de langue loi*. Paris: Centre Nationale de la Recherche Scientifique.

Bodman, Nicholas C., 1987, The place of Lepcha in Sino-Tibetan – a lexical comparison. Paper presented at International Sino-Tibetan Conference, Vancouver.

Bradley, David, 1975, Nahsi and Proto-Burmese-Lolo. *Linguistics of the Tibeto-Burman area* 2/1:93-150.

1979a, *Proto-Loloish*. Scandinavian Institute of Asian Studies Monograph Series No.39. London/Malmö: Curzon Press.

1979b, *Lahu dialects*. Canberra: Australian National University Press.

1983, The linguistic position of Jinuo. In Chu, Chauncey et al., eds, *Proceedings of the Fourteenth International Conference on Sino-Tibetan Languages and Linguistics* 21-42. Taipei: Student Book Publishing Co.

forthcoming, Language policy for the Yi. In Stevan Harrell, ed., forthcoming volume on Yi Studies.

Brauns, Claus-Dieter and Lorenz G. Löffler, 1990, *Mru, hill people on the border of Bangladesh*. Basel: Birkhäuser.

Burling, Robbins, 1961, *A Garo grammar*. Pune: Linguistic Society of India.

1983, The Sal languages. *Linguistics of the Tibeto-Burman area* 7/2:1-32.

Caughley, Ross C., 1982, *The syntax and morphology of the verb in Chepang*. *PL*, B-84.

Census Of India, 1981. Census Department, Government of India.

Chang Kun and Betty S. Chang, 1978-1981, *Spoken Tibetan texts*. 4 vols. Taipei: Institute of History and Philology, Academia Sinica.

Dai Qingxia et al., 1985, *Outline of Achang* [in Chinese]. Beijing: Nationalities Publishing House.

Dai Qingxia, ed., 1992, *A Tibeto-Burman lexicon* [in Chinese]. Beijing: Central Institute of Nationalities Press.

Das Gupta, K., 1963, *An introduction to the Gallong language*. Shillong: Government of NEFA.

1968, *An introduction to Central Monpa*. Shillong: Government of NEFA.

1969, *Dafla language guide*. Shillong: Government of NEFA.

van Driem, George, 1987, *A grammar of Limbu*. Berlin: Mouton.

1992, *The grammar of Dzongkha*. Thimpu: Dzongkha Development Commission of the Royal Government of Bhutan.

1993, *A grammar of Dumi*. Berlin: Mouton.

forthcoming, *Languages of the Himalayas: an ethnolinguistic handbook*. Handbuch der Orientalistik. Leiden: Brill.

Duanghom, Srinuan, 1976, *An Mpi dictionary*. Bangkok: Indigenous Languages of Thailand Research Project.

Duffin, C.H., 1913, *A manual of the Pwo-Karen dialect*. Rangoon: American Baptist Mission Press.

Forrest, R.A.D., 1962, The linguistic position of Rong (Lepcha). *Journal of the Burma Research Society* 45/1:41-7.

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

Gai Xingzhi, 1986, *Outline of Jinuo* [in Chinese]. Beijing: Nationalities Publishing House.

Gilmore, D., 1898, *A grammar of Sgaw Karen*. Rangoon: American Baptist Mission Press.

Glover, Warren W., 1974, *Sememic and grammatical structures in Gurung (Nepal)*. Norman: Summer Institute of Linguistics.

Grierson, George A., 1903-1909, *Linguistic survey of India* vol.3, parts 1-3. Calcutta: Superintendant of Government Printing.
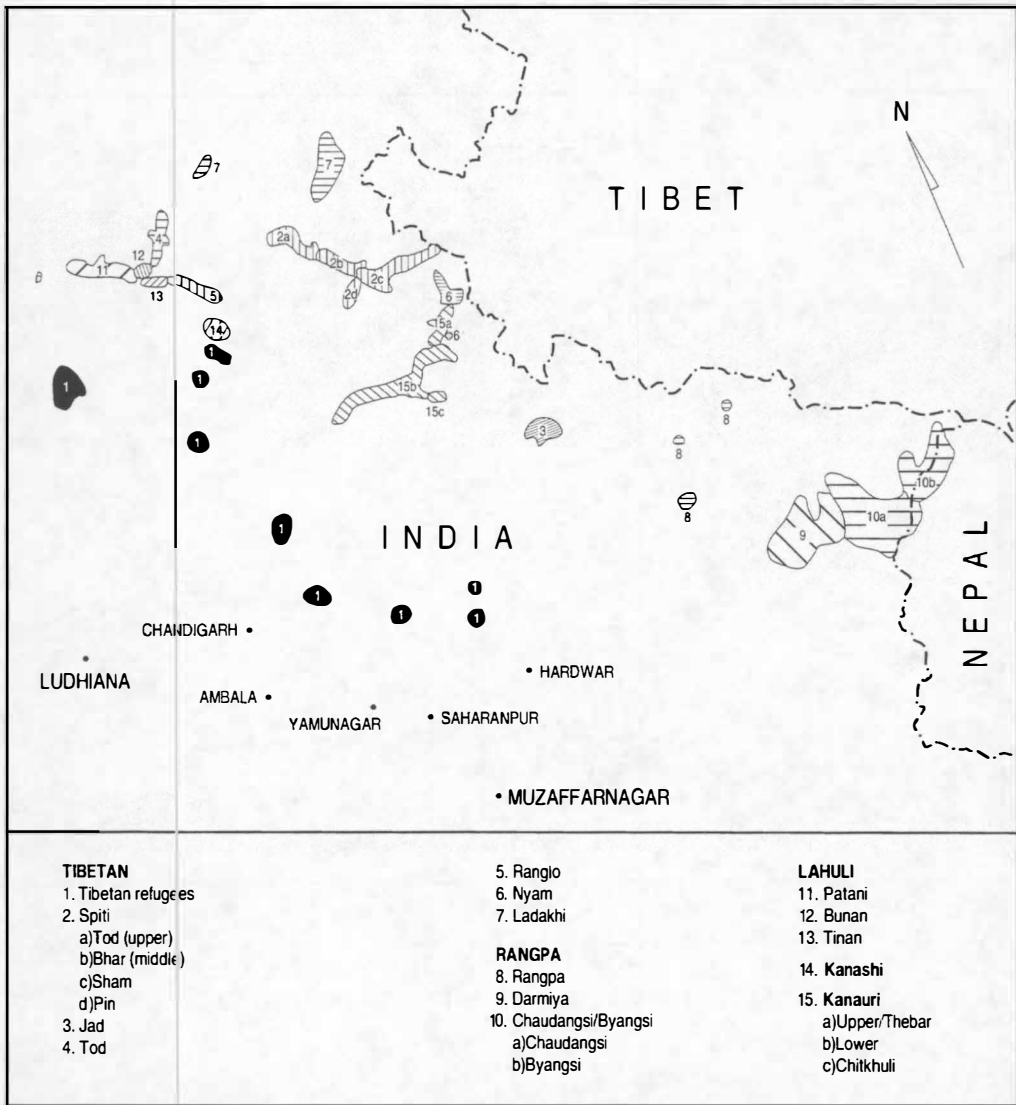
Grüssner, Karl H., 1978, *Arleng Alam, die Sprache der Mikir*. Wiesbaden: Steiner.

Hale, Austin, 1982, *Research on Tibeto-Burman languages*. Berlin: Mouton.

Hansson, Gerd, 1991, *The Rai of Eastern Nepal: ethnic and linguistic grouping findings of the Linguistic Survey of Nepal*. Kathmandu: Linguistic Survey of Nepal and Centre for Nepal and Asian Studies, Tribhuvan University.

He Jiren and Jiang Zhuyi, 1985, *Outline of Naxi* [in Chinese]. Beijing: Nationalities Publishing House.

Henderson, Eugénie J.A., 1965, *Tiddim Chin: a descriptive analysis of two texts*. London: Oxford University Press.

  forthcoming, *A dictionary of Blimaw (Bwe) Karen*, edited by A.J. Allott and with an introduction by David Solnit. London: School of Oriental and African Studies.

Hope, Edward R., 1974, *The deep syntax of Lisu sentences: a transformational case grammar*. *PL*, B-34.

Jones, R.B., 1961, *Karen linguistic studies: description, comparison, and texts*. University of California Publications in Linguistics 25. Berkley and Los Angeles.

Koshal, S., 1979, *Ladakhi grammar*. Delhi: Motilal Banarsidass.

Lewis, Paul, 1968, *Akha-English dictionary*. Ithaca: Cornell University Southeast Asia Program Data Paper Number 70.

Li Yongsui, 1992, Preliminary studies of the Sangkong language [in Chinese]. *Yuyan Yanjiu* 22/1:137-150.

Löffler, Lorenz, 1966a, Chakma und Sak, ethnolinguistische Beiträge zur Geschichte eines Kulturvolkes. *Internationales Archiv für Ethnographie* 50:72-115.

  1966b, The contribution of Mru to Sino-Tibetan linguistics. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 116:119-160.

Lorrain, R.A., 1951, *Grammar and dictionary of the Lakher or Mara language*. Gauhati: Government of Assam.

Lu Shaozun, 1983, *Outline of Pumi* [in Chinese]. Beijing: Nationalities Publishing House.

  1986, *Sketch grammar of Cuona Monpa* [in Chinese]. Beijing: Nationalities Publishing House.

Luce, Gordon H., 1985, *Phases of Pre-Pagán Burma*. 2 vols. Oxford: Oxford University Press.

Ma Xueliang, 1951, *Research on the Sani Yi Language* [in Chinese]. Beijing: Chinese Academy of Science.

Mainwaring, George Byres, 1876, *A grammar of the Rong (Lepcha) language as it exists in the Dorjeeling and Sikkim Hills*. reprinted 1971, New Delhi: Manjusri Publishing House.

Malla, Kamal P., 1985, *The Newari language: a working outline*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Marrison, George E., 1967, *The classification of the Naga languages of North-East India*. PhD thesis, University of London.

  1989, The Adi-Dafla languages of North-east India: a sketch. In Bradley, David, Eugénie J.A. Henderson and Martine Mazaudon, eds, *Prosodic analysis and Asian linguistics: to honour R.K. Sprigg*, 205-222. *PL*, C-104.

Matisoff, James Alan, 1982, *The grammar of Lahu*. Second printing. University of California Publications in Linguistics vol.75.

Michailovsky, Boyd, 1988, *La Langue Hayu*. Paris: Editions du CNRS.

Morse, Robert H., 1965, Syntactic frames for the Rvwàng (Rawang) verb. *Lingua* 15:338-369.

Morse, Robert H. and Betty Morse, 1966, Oral tradition and Rawang migration routes. In Ba Shin et al., eds, *Essays offered to G.H. Luce by his colleagues and friends in honour of his seventy-fifth birthday* vol.I, 195-204. Ascona, Switzerland: Artibus Asiae.

Morse, Steve, 1989, Five Rawang dialects compared plus more. In David Bradley, Eugénie J.A. Henderson and Martine Mazaudon, eds, *Prosodic analysis and Asian linguistics: to honour R.K. Sprigg*, 237-250. *PL*, C-104.

Moseley, Christopher, ed., 1994, *Atlas of the world's languages*. London: Routledge.

Nagano Yashuiko, 1983, *A historical study of the rGyarong verb system*. PhD dissertation, University of California Berkeley.

Nishi Yoshio, 1986, A classification of Tibetan dialects [in Japanese]. *National Minority Studies Museum Research Report* 11/4:837-901.

Nishida Tatsuo, 1973, *A study of the Tosu-Chinese vocabulary Tosu I-Yu* [in Japanese]. Kyoto: Shokado

Nishida Tatsuo and Sun Hongkai, 1990, *A study of the Baima-Chinese vocabulary Baima I-yu* [in Japanese and Chinese]. Kyoto: Shokado.

Ouyang Jueya, 1985, *Outline of Luoba*. [in Chinese]. Beijing: Nationalities Publishing House.

Pai, Pushpa, 1976, *Kokborok grammar*. Mysore: Central Institute of Indian Languages.

Polo, Marco, 1938, *The description of the world*, A.C. Moule and Paul Pelliot, eds. 2 vols. London: G. Routledge.

Qu Aitang and Tan Kerang, 1983, *Ali Tibetan language* [in Chinese]. Beijing: Academy of Social Sciences Publishing House.

Rai, Novel Kishore, 1978, *A basic Bantawa conversational course*. Kathmandu: author.

Reichle, Verena, 1981, *Bawm language and lore: Tibeto-Burman area*. Bern: Peter Lang.

Saikia, P.C., 1976. *The Dibongiyas: social and religious life of a priestly community*. Delhi: B.R. Publishing Corp.

de Sales, Anne, 1993, When the miners came to light: the Chantel of Dhaulagiri. In Gérard Toffin, ed. *Nepal, past and present*, 91-97. Paris: CNRS Editions.

Saxena, Anju, 1992, *Finite Verb Morphology in Tibeto-Kinnauri*. PhD dissertation, University of Oregon.

Shafer, Robert, 1974, *An introduction to Sino-Tibetan*. Wiesbaden: Harrassowitz.

Sharma, Devidatta, 1982, *Studies in Tibeto-Himalayan linguistics: a descriptive analysis of Pattani, a dialect of Lahul*. Hoshiarpur: Panjab University.

    1988, *A descriptive grammar of Kinnauri*. Delhi: Mittal.

    1989a, *Tribal languages of Himachal Pradesh (part one)*. Delhi: Mittal.

    1989b, *Tibeto-Himalayan languages of Uttarakhand (part one)*. Delhi: Mittal.

    1990, *Tibeto-Himalayan languages of Uttarakhand (part two)*. Delhi: Mittal.

    1992, *Tribal languages of Himachal Pradesh (part two)*. Delhi: Mittal.

Simon, Ivan M., 1976, The Khoa language. *Resarun* 2/3:8-10.

Sprigg, Richard Keith, 1983, Hooker's expenses in Sikkim: an early Lepcha text. *Bulletin of the School of Oriental and African Studies* 46/2:305-325.

Stargardt, Janice, 1990, *The ancient Pyu of Burma*. Cambridge: PACSEA.

Sun Hongkai, 1981, *Outline of Qiang* [in Chinese]. Beijing: Nationalities Publishing House.

    1982, *Outline of Dulong* [in Chinese]. Beijing: Nationalities Publishing House.

    1983a, A brief description of Luoba language (Idu) [in Chinese]. *Minzu Yuwen* 6:63-79.

    1983b, The nationality languages in the six valleys and their language branches [in Chinese]. *Yunnan Minzuxuebao* 3:99-273.
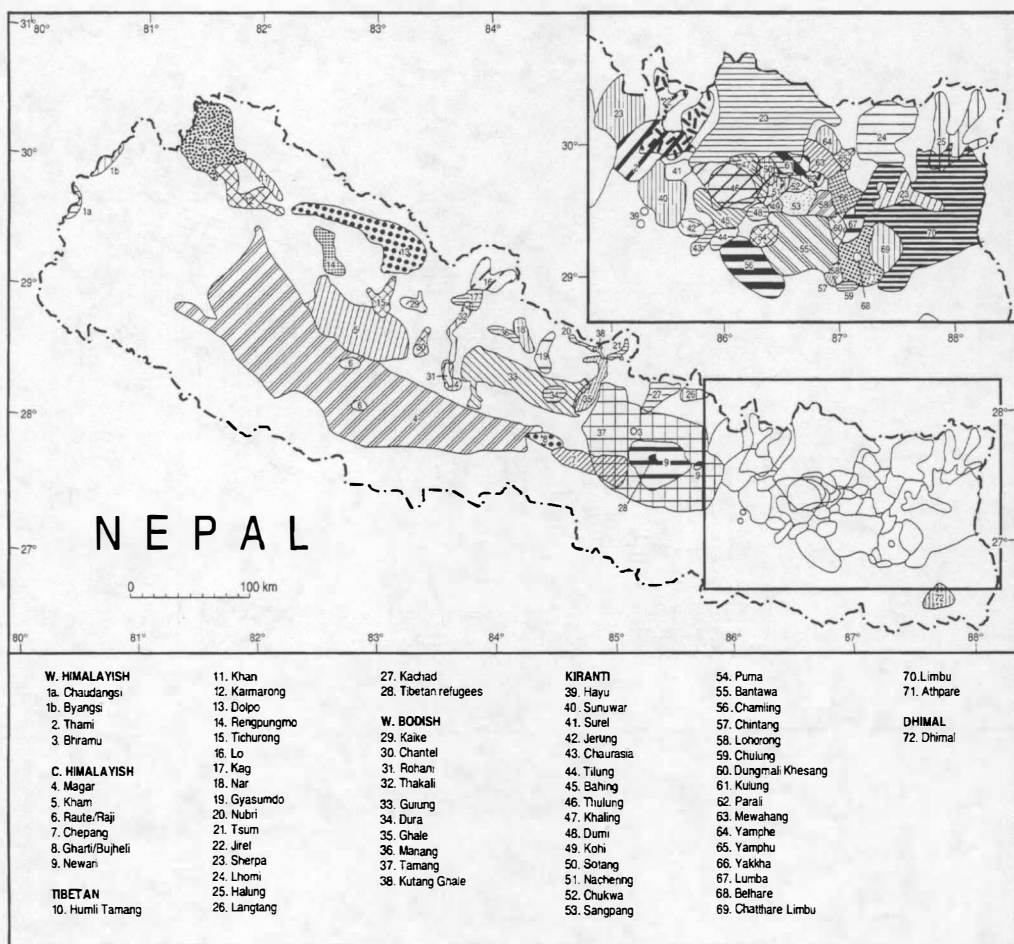
1991, *Tibeto-Burman language phonetics and vocabulary* [in Chinese]. Beijing: Chinese Academy of Social Sciences.

Sun Hongkai, et al., 1980, *The languages of the Monba, Luoba and Deng* [in Chinese]. Beijing: Chinese Academy of Social Sciences.

Sun Hongkai and Liu Lu, 1986. *Outline of Nu (Nusu language)* [in Chinese]. Beijing: Nationalities Publishing House.

Sun, Jackson T.-S., 1986, *Aspects of the phonology of Amdo Tibetan: Ndzorge Sæme Xāra dialect*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

1993, The linguistic position of Tani (Misingish) in Tibeto-Burman: a lexical assessment. *Linguistics of the Tibeto-Burman Area* 16/2:143-188.

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

Tian Desheng, et al., 1986, *Outline of Tujia* [in Chinese]. Beijing: Nationalities Publishing House.

Toba Sueyoshi, 1990, *Thami-English dictionary.* Unpublished MS.

1991, *A bibliography of Nepalese languages and linguistics.* Kathmandu: Linguistic Society of Nepal.

Vesalainen, Olavi and Marja Vesalainen, 1980, *Clause patterns in Lhomi. PL*, B-53.

Webster, Jeff (compiler), 1992, *A sociolinguistic survey of the Tibeto-Burman dialects of north Gorkha district, Nepal.* Kathmandu: author.

Weidert, Alfons and Bikram Subba, 1985, *Concise Limbu grammar and dictionary.* Amsterdam: Lobster Publications.

Wolfenden, Stuart N., 1929, *Outlines of Tibeto-Burman linguistic morphology.* London: Royal Asiatic Society.

Wurm, Stephen A., ed., 1996, *UNESCO atlas of endangered languages.* Paris: UNESCO.

Wurm, Stephen A. and Hattori Shiro, eds, 1981/1983, *Language atlas of the Pacific area. PL*, C-66/C-67.

Wurm, Stephen A., Peter Mühlhäusler and Darryl Tryon, eds, forthcoming. *Atlas of languages for intercultural communication.* Berlin: Mouton de Gruyter.

Wurm, Stephen A. et al., eds, 1987/1991, *Language atlas of China.* Hong Kong: Longmans [in English]. Beijing: Chinese Academy of Social Sciences [in Chinese].

Xu Lin and Zhao Yansun, 1984, *Outline of Bai* [in Chinese]. Beijing: Nationalities Publishing House.

Xu Xijian et al., 1984, *Outline of Jingpo (Caiwa language)* [in Chinese]. Beijing: Nationalities Publishing House.

Yu Daochuan and Y.R. Chao, 1930, *Love songs of the sixth Dalai Lama Tshangs-dbyangs=rgya-mtsho* [in Chinese]. Peiping: National Research Institute of History and Philology.

Zhang Jichuan, 1986, *Outline of Motuo Monpa.* (in Chinese) Beijing: Nationalities Publishing House.

Zoller, Claus P., 1983, *Die Sprache der Rangpas von Garhwal.* Wiesbaden: Harrassowictz.

CHINA

N

UNDER ADMINISTRATION OF CHINA

CLAIMED BY INDIA

UNDER ADMINISTRATION OF PAKISTAN

TIBET

SRINAGAR

INDIA

| 1. Ladakhi | 2. Zangskari | 3. Purik | 4. Balti | 5. Tibetan refugees |

MAP 2: TIBETO-BURMAN LANGUAGES OF KASHMIR

N

T I B E T

I N D I A

N E P A L

CHANDIGARH •

LUDHIANA

AMBALA •

YAMUNAGAR

• HARDWAR

• SAHARANPUR

• MUZAFFARNAGAR

TIBETAN	5. Rangio	LAHULI
1. Tibetan refugees	6. Nyam	11. Patani
2. Spiti	7. Ladakhi	12. Bunan
a)Tod (upper)		13. Tinan
b)Bhar (middle)	RANGPA	
c)Sham	8. Rangpa	14. Kanashi
d)Pin	9. Darmiya	
3. Jad	10. Chaudangsi/Byangsi	15. Kanauri
4. Tod	a)Chaudangsi	a)Upper/Thebar
	b)Byangsi	b)Lower
		c)Chitkhuli

MAP 3: TIBETO-BURMAN LANGUAGES OF HIMACHAL PRADESH AND UTTAR PRADESH

W. HIMALAYISH	11. Khan	27. Kachad	KIRANTI	54. Puma	70.Limbu
1a. Chaudangsi	12. Karmarong	28. Tibetan refugees	39. Hayu	55. Bantawa	71. Athpare
1b. Byangsi	13. Dolpo		40. Sunuwar	56. Chamling	
2. Thami	14. Rengpungmo	W. BODISH	41. Surel	57. Chintang	DHIMAL
3. Bhramu	15. Tichurong	29. Kaike	42. Jerung	58. Lohorong	72. Dhimal
	16. Lo	30. Chantel	43. Chaurasia	59. Chulung	
C. HIMALAYISH	17. Kag	31. Rohani	44. Tilung	60. Dungmali Khesang	
4. Magar	18. Nar	32. Thakali	45. Bahing	61. Kuiung	
5. Kham	19. Gyasumdo		46. Thulung	62. Parali	
6. Raute/Raji	20. Nubri	33. Gurung	47. Khaling	63. Mewahang	
7. Chepang	21. Tsum	34. Dura	48. Dumi	64. Yamphe	
8. Gharti/Bujheli	22. Jirel	35. Ghale	49. Kohi	65. Yamphu	
9. Newari	23. Sherpa	36. Manang	50. Sotang	66. Yakkha	
	24. Lhomi	37. Tamang	51. Nachering	67. Lumba	
TIBETAN	25. Halung	38. Kutang Ghale	52. Chukwa	68. Belhare	
10. Humli Tamang	26. Langtang		53. Sangpang	69. Chatthare Limbu	

MAP 4: TIBETO-BURMAN LANGUAGES OF NEPAL
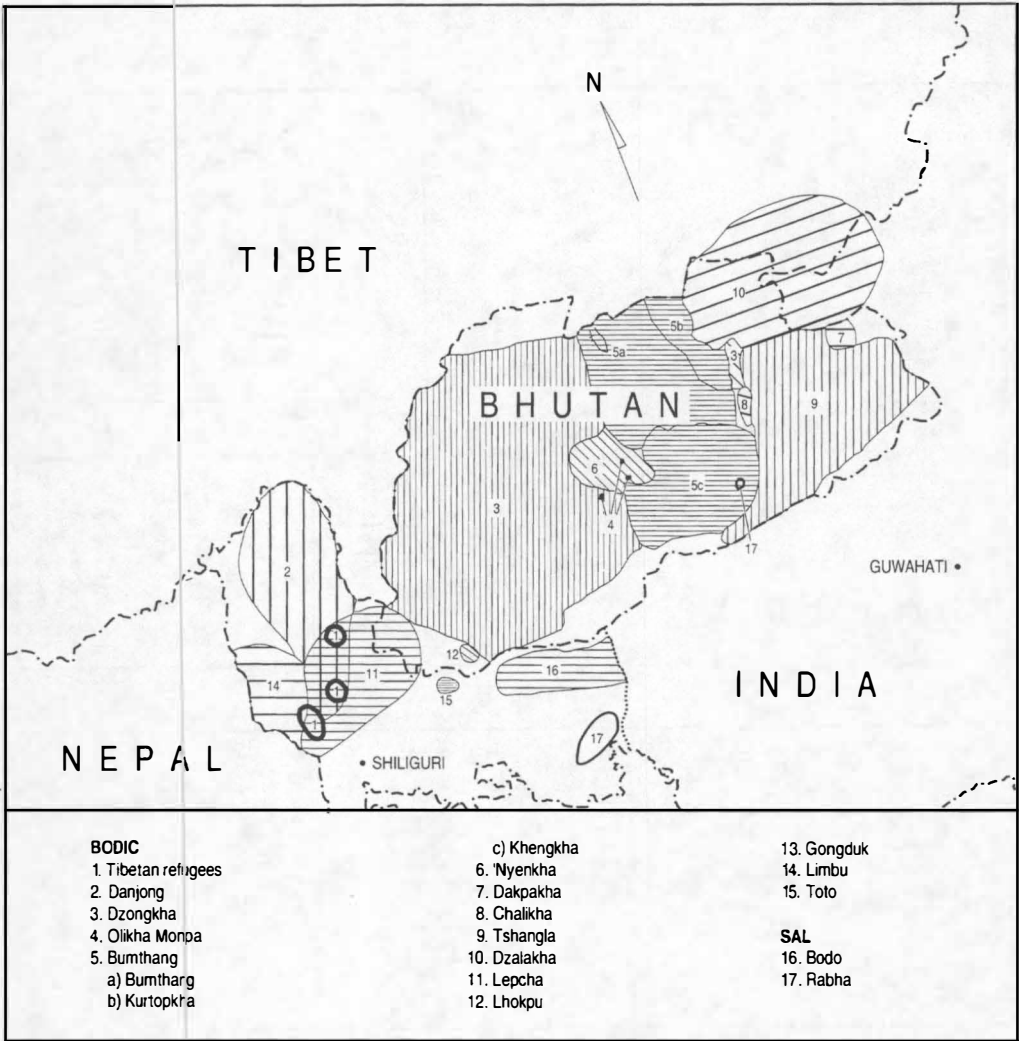
MAP 5: TIBETO-BURMAN LANGUAGES OF WEST BENGAL, SIKKIM AND BHUTAN

BODIC
A. Central
    1. Kham

B. Eastern
    2. Dzalakha
    3. Eastern Monpa
    4. Sherdukpen

C. Tshangla
    5. Tshangla

BURMESE-LOLO
    6. Lisu

CENTRAL
A. West Arunachal
    7. Sulung
    8. Bugun
    9. Hruso
    10. Dhammai
    11. Bangru

B. Adi-Mising-Nishi
    12. Nishi
    12a. Apa Tani
    13. Hill Miri
    14. Adi
    15. Mising

C. Digarish
    16. Idu
    17. Taraon

D. Keman
    18. Keman

SAL
A. Bodo-Garo
    19. Bodo
    20. Lalung
    21. Dimasa
    22. Hojai
    23. Kokborok

    24. Garo
    25. Atong
    26. Wanang
    27. Rabha
    28. Hajong
    29. Deori

B. Northern Naga
    30. Khienmungan
    31. Chang
    32. Phom
    33. Konyak
    34. Wancho
    35. Nocte

    36. Tangsa

C. Jinghpaw (Kadian)
    37. Jinghpaw (Kachin)

MAP 6: TIBETO-BURMAN LANGUAGES OF ARUNACHAL PRADESH,
NORTH ASSAM, NORTH NAGALAND AND MEGHALAYA

88°   90°   92°   94°   96°   30°

N

TIBET

BHUTAN

28°

MYANMAR

(BURMA)

26°

0   250 km

BANGLADESH

24°

Cancer

KUKI-CHIN NAGA
A. Southern Naga
  1. Ao
  2. Sangtam
  3. Lhota
  4. Yimchingru
  5. Ntenyi/Meluri
  6. Sema
  7. W. Ançami
  8. Chakhesang
  9. Mao
10. W. Rengma

11. Maram
12. Zeliangrong

B. Meithei (Manipuri)
13. Manipuri

C. Old Kuki
14. Rangkhol/Hallam/Langrong
15. Hmar
16. Anal
17. Kom
18. Chawte

19. Mayol
20. Tangkhul
21. Maring

D. Chin
22. N. Chin
23. C. Chin/Mizo

E.. Arleng (Karbi,Mikir)
24. Arleng

MAP 7: TIBETO-BURMAN LANGUAGES OF SOUTH ASSAM,
SOUTH NAGALAND AND MANIPUR

MAP 8: TIBETO-BURMAN LANGUAGES OF BANGLADESH, TRIPURA AND MIZORAM

TOWARDS A RECONSTRUCTION OF THE PROTO WEST HIMALAYISH AGREEMENT SYSTEM

ANJU SAXENA

1. INTRODUCTION[1]

West Himalayish (WH) languages display a range of verb agreement systems. Patterns range from the subject agreement system to remnants of a person-based split-ergative agreement system and a no-agreement system. The purpose of this paper is to present a description of the verb agreement systems in six WH languages (Kinnauri, PaTani, Tinani, Gahri, Darmiya and Rangpa). This description, it is hoped, will show the regularity in the verb agreement systems in WH languages, suggesting the possibility that verb agreement could be reconstructed for Proto West Himalayish (PWH). It will also suggest that the no-agreement system and the subject (and the object) agreement system are recent developments in the languages which have them, and that the split-ergative system (similar to the one reconstructed for Proto Tibeto-Burman (DeLancey 1989)) is older, probably reconstructable for PWH.

2. BACKGROUND

2.1 WH LANGUAGES: THEIR GENETIC AND GEOGRAPHICAL SITUATION

The WH group of languages belongs to the Tibeto-Burman (TB) language family. There have been several attempts to classify the TB languages, for example Shafer (1955, 1966), Benedict (1972), Thurgood (1985), and Nishi (1990). Classification of TB languages is still uncertain. Table 1 gives the classification of the WH subgroup based on our current knowledge. The postulation of Tibeto-Kinnauri as a separate branch is based on Benedict (1972), and the classification of WH is from Nishi (1990). The parenthesised languages under West Himalayish are my additions (see Saxena 1992 for details).

© Anju Saxena

TABLE 1: CLASSIFICATION OF THE WEST HIMALAYISH SUBGROUP

Tibeto-Burman

   Bodic

      Tibeto-Kinnauri

         Tibetan: Western, Central, Khams, Southern, Amdo, Monpa

         West Himalayish:

            (a) Kinnauri-PaTani, (Tinani)

            (b) Thebor-Gahri, Rangpa, Chaudangsi, (Darmiya)

WH languages considered in this paper are PaTani, Tinani, Gahri, Darmiya, Rangpa and standard Kinnauri (referred to here simply as Kinnauri). Kinnauri, PaTani, Tinani, and Gahri are spoken in Himachal Pradesh in India, and Darmiya and Rangpa are spoken in Uttar Pradesh in India.

2.2 DATA

While there is a number of individual Bodic languages for which documentation is lacking, the WH subbranch remains the most seriously underdocumented genetic unit within Bodic.

The available materials on Kinnauri are Bailey (1909), Joshi and Rose (1909), Neethivanan (1971), and D.D. Sharma (1988). Of the remaining languages of this group, we have only Konow (1909), Francke (1909), Zoller (1983), S.R. Sharma (1987), and D.D. Sharma (1989a,b). They are good attempts to describe some WH languages. However, with the exception of Zoller (1983), they miss some very important phonological and grammatical facts, and lack the kind of detailed information needed to do comparative and historical work.

For the present study, descriptions of Kinnauri and PaTani are based on the data which I collected during a fieldtrip to India in 1989-90. The description of Rangpa is from Zoller (1983). Descriptions of Tinani and Darmiya are based on the data provided in Sharma (1989a,b). And the description of Gahri is based on the data given in Sharma (1989a) and Francke (1909). The interpretation of the data from Sharma (1989a,b) and Francke (1909) is mine, except where mentioned. The morpheme and word boundaries in the examples below describe the analysis presented here. In some cases I have revised the free translations. The original transcription of the data is, however, retained.[2]

2.3 LITERATURE SURVEY

There has been some discussion concerning the development of verb agreement in TB (Konow 1909, Maspero 1947, Egerod 1973, Bauman 1975, Caughley 1982, DeLancey 1989). Konow (1909), Maspero (1947), Egerod (1973) and Caugley (1982) suggest that verb agreement in TB languages is a secondary development, whereas Bauman (1975) and DeLancey (1989) argue in favour of reconstructing verb agreement for PTB.

---

2   The following is an exception. Sharma (1989a,b) uses two symbols each to represent velar nasal, and palato-alveolar voiceless affricate. For the sake of clarity, I will use the symbol ŋ for velar nasal, and č for palato-alveolar voiceless affricate.

Konow (1909) claimed that verb agreement in TB is due to the influence of the Munda languages. TB and Munda languages show some resemblances in their verb agreement systems.

Maspero (1947) and Egerod (1973) rejected the Munda hypothesis, and argued that verb agreement in TB is due to the Indic influence.

In the same vein, Caughley (1982) suggested that Tibeto-Burman languages with verb agreement systems have developed these "through innovation or areal influence". He based his suggestion on the observation that TB languages display a wide range of verb agreement patterns.

Bauman (1975) and DeLancey (1989), on the other hand, argue in favour of reconstructing verb agreement for PTB. To quote DeLancey (1989:317),

> There is in fact one paradigm, definable both by morphological form and paradigmatic structure, which is attested in at least one representative of almost every branch of the family, and that this paradigm, at least, must therefore be reconstructed for their common ancestor, PTB.

DeLancey (1989:316) reconstructed a person-based split-ergative agreement system for PTB: "...in which agreement in a transitive clause is associated with person rather than function, so that the verb agrees with 1 or 2p. subjects or objects". Table 2 presents the PTB verb agreement schema reconstructed by DeLancey. It illustrates only the agreement suffixes. The agreement markers are listed in this table, depending on the persons of the subject and the object. The horizontal lines indicate the person of the object, and the vertical lines indicate the person of the subject.

TABLE 2: PROTO TIBETO-BURMAN VERB AGREEMENT SCHEMA

SUBJECT	OBJECT		
	1	2	3
1		-*n*	-*ŋ*
2	-*ŋ*		-*n*
3	-*ŋ*	-*n*	-*u*

2.4 ORGANISATION OF THE PAPER

The purpose of this paper, as mentioned above, is to present data which suggest that verb agreement could be reconstructed for PWH. Data from Gahri, Darmiya and Kinnauri indicate that the person-based split-ergative agreement system is older than other prevalent agreement systems in WH languages, and could perhaps be reconstructed for PWH.

The organisation of the paper is as follows. Section 3 presents a typology of the finite verb morphology in WH, concentrating on agreement morphology. Section 4 presents a preliminary schema of the PWH verb agreement system. In this section, first, PWH verb agreement schema for intransitive verbs is presented. Next, for transitive verbs, data are

presented which suggest the antiquity of the person-based split-ergative system. And, finally, development of the subject agreement markers in WH will be considered.

While describing the finite verb morphology in these languages, the terms 'subject', 'object', 'split-ergative agreement system', 'subject agreement system', 'no-agreement system' and 'auxiliary' will be used. The term 'subject' refers to the only core argument of intransitive verb, and the agentive argument of a transitive verb. The term 'object' refers to the patient argument of a transitive verb. The term 'split-ergative agreement system' refers to the person-based split-ergative agreement system, where (DeLancey 1989:318) "the agreement is sometimes with object, i.e. in an ergative pattern and sometimes with subject, with the choice determined by the person of the two arguments". In the 'subject agreement system' on the other hand, the agreement is with the subject of the clause, regardless of the person of the two arguments. 'No-agreement system' indicates that there is no agreement morphology on the verb. And, the term 'auxiliary' is used here to refer to copulas when they occur in non-copula constructions, and to a set of morphemes which occurs at the end in non-copula constructions in some WH languages. Morphemes belonging to the latter group have the status of independent morphemes, but they do not function as verbs in these languages.

3. TYPOLOGY OF THE FINITE VERB MORPHOLOGY IN WH

WH languages are clause-chaining languages, where the verb of the final clause has tense, aspect and agreement morphology. In this section I will present a brief description of the finite verb morphology in Kinnauri, PaTani, Tinani, Darmiya, Gahri and Rangpa, concentrating on their agreement morphology. Since in many TB languages the agreement morphology is a reanalysis of the pronominals used in those languages, a table describing the pronominals of that language will follow the table showing the agreement morphology.

3.1 KINNAURI

A final verb in Kinnauri consists of a verb stem, a tense marker and a subject agreement marker. In some cases markers of aspect, object agreement and honorificity are also suffixed to the verb. The final verb in Kinnauri has the structure

V-(OBJ)-TNS-SUB   or   V-(OBJ)-ASP AUX-TNS-SUB.[3]

Tables 3 and 4 describe the Kinnauri subject agreement markers and the pronominal paradigm, respectively. A blank in a slot in tables indicates that there is no overt agreement marker in the language for that category, and ---- indicates that the form is not available.

The subject agreement markers are the same in copula and non-copula constructions. There is, however, some variation among Kinnauri speakers concerning the third person singular non-honorific subject agreement marker in non-copula constructions. None of my Kinnauri informants, except one, use overt agreement marker for third person non-honorific

---

3   The descriptions of the abbreviations used in this study are as follows. ACC = accusative, AGR = agreement, ASP = aspect, AUX = auxiliary, COP = copula, DAT = dative, DEF = definite, DU = dual, ERG = ergative, EXC = exclusive, GEN = genitive, HON = honorific, IMPF = imperfective, INC = inclusive, INST = instrumental, LOC = locative, NOM = nominaliser, OBJ = object agreement, ORD = ordinary (-honorific), PERF = perfective, PL = plural, PROG = progressive, PST = past, SG = singular, SUB = subject agreement, TNS = tense, and V = verb.

singular subject agreement marker. But in one informant's speech[4], there is an alternation between Ø and -t/d, for example khya, khya-d '(he/she) saw', and dza, dza-d '(he/she) ate'. Such an alternation is restricted to a few verbs in the past tense, even in this informant's speech.

TABLE 3: KINNAURI SUBJECT AGREEMENT MARKERS

Person	Singular	Dual	Plural
1	-k	-č	-me
2	-n	-n	-n(o)
2 (HON)	-ñ	-č	-č
3	(-d)		
3 (HON)	-š	-š	-š

TABLE 4: KINNAURI PRONOMINAL PARADIGM

Person	Singular	Dual	Plural
1	gə	niši (EXC) kišaŋ (INC)	kišaŋ
2	kə	kəniš	kano
2 (HON)	ki	kiši	kino
3	do / hodo	----	dogo / hodogo
3 (HON)	honogo	----	honogo

Kinnauri marks object agreement on the verb, if the object is a first or a second person pronoun. The object agreement marker is -č (except for the verb 'give', see below for details). It is suffixed to the main verb. It occurs in all tenses and aspects. Clauses involving object agreement can have any person as their subject. Examples (1) and (2) illustrate the object agreement marker. Example (3) shows that the object agreement marker does not occur with third person objects.

(1) Ram-əs   əŋ-u   taŋ-č-e-š.
     Ram-ERG self-ACC see-OBJ-PST-3HON
     Ram saw me.

(2) Gə ki-nu   taŋ-č-o   du-k.
     I   you.HON-ACC see-OBJ-PROG   be-1SG
     I am watching you.

---

4   This informant (Arjun Negi) is from the Kalpa region in Kinnaur. According to my Kinnauri informants, -t/d is a peculiarity of the Kinnauri spoken in the Kalpa region.

(3)   *Ram-əs   tshətshats-(u) taŋ-e-š.*
       Ram-ERG girl-ACC   see-PST-3HON
       Ram saw the girl.

The verb 'give' in Kinnauri has two forms, *ker* and *ran*. Their occurrence depends on the person affected. The verb form *ker* 'give' occurs with first and second persons, and the verb form *ran* 'give' occurs with third person arguments. In such cases the object agreement marker *-č* does not occur.

(4)   *Arjun-əs   goldi-pəŋ   kitab ran-o-š.*
       Arjun-ERG Goldi-DAT book give-PST-3HON
       Arjun gave a book to Goldi.

(5)   *əma-s   əŋ-u   za-mu   ker-o-š.*
       mother-ERG self-DAT eat-NOM give.OBJ-PST-3HON
       Mother gave me food to eat.

3.2 PaTani[5]

A final verb in PaTani involves a verb stem, a tense marker, a subject agreement marker, and, optionally, an aspect marker.

Tables 5 and 6 describe the PaTani subject agreement markers and the pronominal paradigm, respectively. As Table 5 shows, there are two allomorphs of each of the agreement markers. These allomorphs are in free variation.[6]

TABLE 5: PaTani SUBJECT AGREEMENT MARKERS

Person	Singular	Dual	Plural
1	*-g(à)*	*-š(î)*	*-ñ(î)*
2 (+/- HON)	*-n(à)*	*-š(î)*	*-ñ(î)*
3 (+/-HON)		*-k(ù)*	*-r(è)*

TABLE 6: PATANI PRONOMINAL PARADIGM

Person	Singular	Dual	Plural
1	gè	nè-kù (EXC) hén-gù (INC)	nè-rè (EXC) hénə-rè (INC)
2 (-HON)	kà	kè-kù	kè-rè
2 (HON)	kè-nà	kèŋ-gù	kènə-rè kèn-dè
3 (+/-HON)	dù	dò-kù	dò-rè

Unlike Kinnauri, PaTani does not mark object agreement.

(6)   *Ràm-è   gì-bì   táŋ-à   thù.*
        Ram-ERG I-ACC see-PERF AUX.3SG
        Ram saw me.

(7)   *Ràm-è   kátu-bì   táŋ-à   thù.*
        Ram-ERG child-ACC see-PERF AUX.3SG
        Ram saw the child.

3.3 TINANI

A final verb in Tinani involves a verb stem, a tense marker, a subject agreement marker, and optionally an aspect marker. The copula construction has the structure COP-TNS-SUB, and the non-copula construction has the structure V-ASP AUX-SUB or V-TNS-SUB.

The subject agreement markers are regularly suffixed to the verb (except when the tense marker is -*min*; see below for details). Tables 7 and 8 describe the Tinani subject agreement markers and the pronominal paradigm, respectively. The pronominal paradigm is from Sharma (1989a:145-146). As in PaTani, in Tinani, each of the agreement markers has two allomorphs.

TABLE 7: TINANI SUBJECT AGREEMENT MARKERS

Person	Singular	Dual	Plural
1	-k / -g(a)	-ṣ(i)	-ñ(i)
2	-n(a)	-č(i)	-č(i)
3		-k(u)	-r(e)

<div align="center">TABLE 8: TINANI PRONOMINAL PARADIGM</div>

Person	Singular	Dual	Plural
1	*gye*	*iša* (EXC) *ñiši* (INC)	*ena* (EXC) *ñena* (INC)
2 (ORD)	*kə*	*kənca*	----
2 (HON)	*kenə*	*kenci*	*kena*
3 (+/-HON)	*du/do*	*do-ku*	*do-re*

The suffix *-min/men* in WH languages functions as a nominaliser. Examples of *-min* as a nominaliser follow. (The examples are the same in Kinnauri, PaTani and Tinani).

(8)   *za*   eat   :   *za-min*   food/eating

    tuŋ   drink   :   *tuŋ-min*   drinking

An important characteristic of the Tinani finite verb system is the use of the nominaliser *-min* as a tense marker. It occurs in copula as well as in non-copula constructions. In such constructions it is not followed by a copula. According to Sharma (1989a), it has a past tense interpretation.

(9)   *Gye ica seu za-min.*[7]
    I   one apple eat-NOM/TNS
    I ate an apple.

The suffix *-min* as a tense marker in Tinani occurs with all persons and numbers, but unlike other tense markers, it does not take agreement markers.

(10)   *Gye ri-riŋ   ica khoro hə-min.*
    I   sister-DAT one cap   take-NOM/TNS
    I brought a cap for (my) sister.

(11)   *Kə khyaŋ i-min?*
    you where go-NOM/TNS
    Where had you gone?

(12)   *Do-i   gye-riŋ bəgət rə-min.*
    he-ERG I-DAT food   give-NOM/TNS
    He gave me food.

(13)   *Do-re eki   ə-min.*
    he-PL yesterday come-NOM/TNS
    They came here yesterday.

3.4 DARMIYA

A final verb in Darmiya involves a verb stem and a tense marker. It optionally takes an aspect marker. The subject agreement markers occur in certain restricted enviornments.

---

7   Sharma (1989a,b) does not provide interlinear glossing. The glosses, provided here, are my additions.

The finite verb system in Darmiya is different from the systems found in Kinnauri, PaTani and Tinani in a number of ways. First, unlike these languages, Darmiya makes a two-way number distinction on nominals as well as on verbs. Second, Darmiya shows only traces of the subject agreement system. The agreement marker occurs with first person plural and second person (singular and plural) subjects. In the copula construction the agreement morphology occurs only in the past tense, but in the non-copula construction it occurs in all tenses. In the past tense the agreement marker is *-n,* and in the present and future tenses the agreement markers are *-n* and *-ni.* While *-n* occurs with first person plural and second person singular subjects, *-ni* occurs with second person plural subjects. Further, the agreement marker precedes the past tense marker, but follows the tense marker in present and future tenses. The verb paradigm, given below, is illustrative (*ga* 'do').

(14) | | 1SG | 2SG | 3SG |
|---|---|---|---|
| Present | *ga-di* | *ga-də-n* | *ga-da* |
| Future | *ga-ŋdi* | *ga-ŋdə-n* | *ga-ŋda* |
| Past | *ga-su* | *ga-n-su* | *ga-su* |

The agreement markers are regularly suffixed to the verb in such constructions, except when the verb ends in a nasal. In that case there is no agreement marker (see example (17)).

(15) *Gɛ-su   dilli   khərju ji daŋsu khə   rhe-n-su?*
you-ERG Delhi from I to   what bring-SUB-PST
What have you brought for me from Delhi?

(16) *Niŋ əphi=lən=əphina ga-ŋdə-n*
we   ourselves   do-FUT-SUB
We will do the work by ourselves.

(17) *Gɛ-su   ge   udi   taŋ-su?*
you-ERG where clothes put-PST
Where have you put the clothes?

Tables 9 and 10 describe the Darmiya subject agreement system and the pronominal paradigm, respectively. The pronominal paradigm is from Sharma (1989b:56-57).

TABLE 9: DARMIYA SUBJECT AGREEMENT SYSTEM

	1PL, 2SG, 2PL	1PL, 2SG, 2PL
Copula	V-*n*-TNS (past tense)	
Non-copula (transitive)		V-TNS-*n/ni* (present, future) V-*n*-TNS (past tense)

TABLE 10: DARMIYA PRONOMINAL PARADIGM

Person	Singular	Plural
1	*ji*	*niŋ*
2	*gɛ/gəe*	*gəni*
3	*o/u*	*usi*

3.5 GAHRI

The final verb system in Gahri is interesting for a number of reasons. First, it shows traces of three cycles of agreement morphology. These are the (1) split-ergative agreement system, (2) subject agreement system, and (3) no-agreement system. Second, a three-way number distinction is made on nominals, but a two-way number distinction is made on verbs in most cases. Third, the nominaliser *-men* functions as a tense marker in Gahri too, but unlike Tinani, in the Gahri data provided in Francke (1909), it takes subject agreement markers.

The suffix *-min/-men* in Gahri, as in other WH languages, functions as a nominaliser (example (18)). As in Tinani, in Gahri, it functions as a tense marker with all persons (example (19)).

(18)   *za*   :   *za-men*
        eat   eat-NOM
        eat   eating/food

(19)   *Gi-zi   khai   khyu-ti   thaŋ-men.*
        I-ERG black dog-DEF see-NOM/TNS
        I saw a black dog.

There is one major difference between the copula paradigms given in Francke (1909) and Sharma (1989a). Unlike Francke's paradigms, Sharma's paradigms almost completely lack agreement markers. For example, in Sharma (1989a:240) the agreement marker does not follow the tense marker *-men*.

(20)   *Gi-zi   lig-men.*
        I-ERG do-NOM/TNS
        I did (it).

But, in the data provided in Francke (1909) *-men* takes the subject agreement markers.

(21)   *Lig-men-gya.*
        do-PST-1SG
        (I) did (it).

Further, in Sharma (1989a) the present tense copula paradigm of *yen/hen* has one invariant form *hen*, for all persons and numbers (see Table 11). Similarly, the copula *ni* in the past tense is shown as having one invariant form *ni-n-za* for all singular subjects, which is not the case in Francke (1909). For the purpose of comparison, the copula paradigms from Sharma (1989a) and Francke (1909) are given below. Tables 11 to 14 describe the copula paradigms provided in Sharma (1989a).

TABLE 11: *Hen* (PRESENT TENSE)

Person	Singular	Plural
1	*hen*	*hen*
2	*hen*	*hen*
3	*hen*	*hen*

TABLE 12: *Ni/go* (PRESENT TENSE)

Person	Singular	Plural
1	*na*	*goig*
2	*ni*	*gwag*
3	*ni*	*gwag*

TABLE 13: *Ni/go* (PAST TENSE)

Person	Singular	Plural
1	*ni-n-za*	*go-i-ča*
2	*ni-n-za*	*gwa-n-ča*
3	*ni-n-za*	*gwa-n-ča*

TABLE 14: *Kya/hen* (FUTURE TENSE)

Person	Singular	Dual	Plural
1	*kya-ni /* *hen-ge-ni*	*kya-kheg*	*kya-kheg /* *heŋ-kheg*
2	*kya-ni /* *hen-ge-ni*	*khag*	*kya-khag /* *heŋ-khyag*
3	*kya-ni /* *hen-ge-ni*	*khag*	*kya-khag /* *heŋ-khyag*

Francke's copula paradigms are given in Tables 15 to 18.

TABLE 15: *Yen* (PRESENT TENSE)

Person	Singular	Dual and Plural
1	*yen-gya*	*yen-ni*
2	*yen-na*	*yen-ni*
3	*yen*	*yen*

TABLE 16: *Ni/goag* (PRESENT TENSE)

Person	Singular	Dual and Plural
1	*ni-a*	*goa-i-(g)*
2	*ni-na*	*goag-ni*
3	*ni*	*goa(g)*

TABLE 17: *Ni/goag* (PAST TENSE)

Person	Singular	Dual and Plural
1	*ni-za*	*goa-i-thsa*
2	*ni-za-na* *ni-n̠-za-na*	*goa-n̠-thsa-n̠i*
3	*ni-za, ni-n̠-za*	*goa-n̠-t sa*

According to Francke (1909), *ded* 'be' occurs in the "incomplete present tense".

TABLE 18: *Ded* (PRESENT TENSE)

Person	Singular	Plural
1		
2	*de-na*	*ded-ni*
3	*de*	*de(d)*

Based on the data provided in Sharma (1989a) and Francke (1909), the possible finite verb endings in the past tense non-copula constructions in Gahri are given in Table 19. These inflectional markers are suffixed to the verb. The verb endings are listed in the table, depending on the persons of the subject and the object. The horizontal lines indicate the person of the object, and the vertical lines indicate the person of the subject. For example,

-ki-za[8] in the second column indicates that it occurs when the subject is a first person pronoun and the object is a second person pronoun. And -ku-za in the third row indicates that it occurs with third person subject and first person object arguments.

TABLE 19: GAHRI NON-COPULA VERB INFLECTIONAL ENDINGS (PAST TENSE)

SUBJECT	OBJECT		
	1	2	3
1		-ki-za	-d -ki-za -men-(gya) -i tha-i-(g) -ki ta (sa)
2	-za		-za -ta (na) -tad (ni) -(n)-za-na -men-na -(n)-ča-ni -men-ni
3	-(ku)-za	-za	-ta -men -(n)-za

Among the various verb endings, -ča, -men, -d and -za function as the past tense markers. As seen earlier, -ča and -za function as the past tense markers in the copula constructions too. Ta and tha are "versatile" verbs (Matisoff 1969). Examples of ta as a lexical verb and as an auxiliary follow.

(22)   Gi-i   iṣring   ti   ta.
        I-GEN   sister   one   have
        I have a sister.

(23)   Gi   ingi-i   lən   ingi-zi   lik-a   ta.
        I   self-GEN   work   self-INST   make-?   AUX
        I myself do my work.

As Table 19 shows, two agreement markers can be affixed to one verb in Gahri (for example with second person subject and third person object-n and -na/ni in the final verb structure V-n-TNS-na/ni), which is not the case in any other WH language that we have examined so far.

Further, Gahri displays signs of at least three verb agreement cycles: (1) the split-ergative agreement system, (2) the subject agreement system, and (3) the no-agreement system. The split-ergative system is represented here by -ki and -ku. The subject agreement system is represented here by -gya and -g (first person) and -na, -ni (second person). Of the two

---

8   In Francke (1909) -i is an allomorph of -ki.

second person markers, -*na* occurs with second person singular subjects, and -*ni* occurs with second person plural subjects. As seen earlier, they also occur in copula constructions. Verbs involving -*men* in Sharma's data display the no-agreement system. Since the alternation of the various agreement markers is found in almost all cases, their choice is not a syntactic decision.

Table 20 provides the Gahri pronominal paradigm (Sharma 1989a:221).

TABLE 20: GAHRI PRONOMINAL PARADIGM

Person	Singular	Dual		Plural	
1	*gi*	*hiŋ*	(EXC)	*hiŋ-ži*	(EXC)
		eraŋ	(INC)	*eraŋ-ži*	(INC)
2 (ORD)	*hən*	*hən*	(INC)	*hən-ži*	(INC)
		nispi	(INC)		
2 (HON)	*ini*				
3	*təl*	*təl*	(EXC)	*təl-ži*	(EXC)
		nispi	(EXC)		

3.6 RANGPA

Rangpa has a two-way number distinction on nominals as well as on verbs. The subject agreement marker is regularly suffixed to the verb.

The WH languages which we have considered so far have -*k/ga* as the first person singular marker. Rangpa is unique in this respect. It has -*ŋ* as the first person singular agreement marker in the "general present tense" (Zoller 1983). The agreement markers in the various tenses are given in Tables 21 to 24. In the past and future tenses the third person subject agreement also involves change in the stem final vowel.

TABLE 21: RANGPA SUBJECT AGREEMENT MARKERS (DEFINITE PRESENT)

Person	Singular	Plural
1	-*ki*	-*ni*
2	-*n(i)*	-*ni*
3	-*ni*	-*ni*

TABLE 22: RANGPA SUBJECT AGREEMENT MARKERS (GENERAL PRESENT)

Person	Singular	Plural
1	-*ŋ*	-*ni*
2	-*n*	-*ni*
3	-*n*	-*ni*

TABLE 23: RANGPA SUBJECT AGREEMENT MARKERS (PAST TENSE)

Person	Singular	Plural
1	*-ki*	*-n*
2	*-n*	*-n*
3	*(-n)*	*(-n)*

TABLE 23: RANGPA SUBJECT AGREEMENT MARKERS (FUTURE TENSE)

Person	Singular	Plural
1		*-n*
2	*-n*	*-n*
3		*-n*

4. PRELIMINARY SCHEMA OF PWH VERB AGREEMENT SYSTEM

4.1 PWH VERB AGREEMENT SCHEMA (INTRANSITIVE)

The above description shows that verb agreement occurs regularly in WH languages.

Kinnauri, PaTani, Tinani, and Rangpa have the same set of agreement markers in the copula constructions, and in the non-copula constructions involving transitive and intransitive verbs. We do not have sufficient data of Gahri and Darmiya intransitive verbs to say whether or not the agreement markers in the non-copula constructions involving intransitive verbs are the same as in the copula construction. Table 25 summarises the WH agreement morphology in copula constructions.

TABLE 25: WH SUBJECT AGREEMENT MARKERS

	Kin	PaT	Tin	Dar	Gah (pres)	Gah (past)	Rang
1SG	*-k*	*-g(à)*	*-k / -g(a)*		*-gy(a)*		*(-ki)/(-ŋ)*
1DU	*-č*	*-š(ì)*	*-ṣ(i)*				
1PL	*-me*	*-ñ(ì)*	*-ñ(i)*	*-n*-TNS	*-ni*		*-n(i)*
2SG	*-n*	*-n(à)*	*-n(a)*	*-n*-TNS	*-na*	*-(n)*-TNS-*na*	*-n(i)*
HON	*-ñ*						
2DU	*-n*	*-s(ì)*	*-c(i)*				
HON	*-č*						
2PL	*-n(o)*	*-ñ(ì)*	*-č(i)*	*-n*-TNS	*-ni*	*-n*-TNS-*ni*	*-n(i)*
HON	*-č*						

3SG	(-d)					(n)-TNS	(-n(i))
HON	-š						
3DU		-k(ù)	-k(u)				
HON	-š						
3PL		-r(è)	-r(e)			-n-TNS	(n(i))
HON	-š						

In this table Kin, PaT, Tin, Dar, Gah and Rang refer to Kinnauri, PaTani, Tinani, Darmiya, Gahri and Rangpa, respectively. A blank in a slot indicates that there is no overt agreement marker in the language for that category. Kinnauri, PaTani, Tinani and Darmiya[9] have one set of agreement markers for all tenses. Gahri has separate sets of agreement markers for the past and the present tenses.[10] Rangpa is like Gahri to some extent. It has two first person singular markers -ŋ and -ki. The marker -ŋ occurs in the "general present tense", and -ki occurs in the past and "definite present tense" (Zoller 1983).

The table shows that verb agreement occurs regularly in these languages, and that the agreement markers are also similar. Such regularities in the agreement systems suggest that verb agreement has not developed individually in these languages. In fact, the agreement markers and their patterns are similar to the agreement markers found in TB languages outside WH, and, in some cases, to the agreement markers reconstructed for PTB.

Benedict (1972) and Bauman (1975) have reconstructed *ŋa and *naŋ as the first and second person singular pronouns, respectively, for PTB. There are, however, some TB languages (such as Bahing, Vayu and Mishmi), which have a voiced velar stop for first person singular (for example, Bahing and Vayu go), and a voiceless velar stop for second person singular, instead of a nasal. WH languages belong to the latter group. Examples of WH first person singular pronouns starting with a voiced velar stop, are gə (Kinnauri), gye (PaTani, Tinani) and gi (Gahri), and an example of a voiceless velar stop for second person singular is kə in Kinnauri, PaTani and Tinani.

The second person singular agreement marker in all the WH languages considered here is -n/na, suggesting the possibility of reconstructing *-na as the second person singular agreement marker for PWH.

Concerning the first person singular agreement marker, there are two markers that could each be potentially reconstructed for PWH: (1) -k/ga, which functions as first person singular agreement marker in all WH languages, and (2) -ŋ, the first person singular agreement marker in "general present tense" in Rangpa. DeLancey (1989) reconstructs *-ŋ as the first person singular agreement marker for PTB. It is plausible that -ŋ in Rangpa is a cognate of the PTB agreement marker, though the majority rule argues in favour of reconstructing -k/ga as the first person singular marker. Cognacy of -ŋ with the PTB first person singular marker is taken here to suggest *-ŋ as the PWH first person singular agreement marker, though -ga could equally well be reconstructed.

---

9   The agreement markers in the copula construction in Darmiya occur only in the past tense.
10   This description is based on Francke (1909).

Of the six WH languages considered here, Kinnauri, PaTani and Tinani make a three-way number distinction on nominal arguments as well as on verbs. Darmiya and Rangpa make a two-way number distinction – both on nominals and on verbs. Gahri, on the other hand, has a three-way number distinction on nominals, but mostly a two-way distinction on verbs.[11] The following pieces of evidence suggest reconstructing a three-way number distinction for PWH.

First, WH languages which make a three-way number distinction have cognate forms for first person singular, first person dual and first person plural agreement markers, suggesting their common origin (see Table 25). The first person singular has a -k/ga, the dual marker has an affricate (č- or š-), and first person plural has a nasal.

Second, voiceless affricate as a dual marker is also attested in TB languages outside WH. For example, -chi functions as first and second person dual markers in Vayu, and tayhca as the first person dual marker in Chepang. Lushai has -ce as the second person dual marker, and Thulung has -ci as the first and second person dual marker.

The frequent occurrence of -č as a dual marker suggests the possibility of reconstructing *-č as the dual marker – at least for Proto Bodic. Bauman (1975:103) goes a step further, and states that

> The dual marker can fairly easily be traced back to some sibilant plus high front vowel (#shi) [in PTB]. Such an element is present in all of the affix forms and some of the free pronoun forms.

Based on the observations made above, a preliminary schema of the PWH intransitive agreement morphology is given in Table 26.

TABLE 26: PWH AGREEMENT MORPHOLOGY (INTRANSITIVE)

Person	Singular	Dual	Plural
1	*-ŋ	*-či	*-ñi
2	*-na	*-či	*-ñi
3		*-či	*-ñi

WH languages which have a three-way number distinction, differ regarding their dual and plural markers in non-first persons (see Table 25). For example, PaTani has the same agreement marker for first and second persons, and maintains a three-way number distinction (-ši for first or second person dual and -ñi for first or second person plural), but in Tinani and Kinnauri the distinction between dual and plural is lost in the second person (-či is second person dual and second person plural marker in Tinani, and -č is second person dual (honorific), second person plural (honorific) and first person dual marker in Kinnauri). It is suggested here that separate dual and plural markers for various persons, and the honorificity marker in Kinnauri, are secondary developments in these languages. The agreement markers, which are not cognates of the agreement markers reconstructed for WH, are cognates of the number and the honorific markers on nominals in these languages. For example, the third person honorific marker -š is also the honorific marker on nominals in Kinnauri. Similarly,

---

[11]   This statement is true of the copula constructions in Gahri. It is not clear whether a two-way or a three-way number distinction is made in the noncopula constructions in Gahri.

the third person dual and third person plural agreement markers in PaTani and Tinani, -*ku* and -*re*, respectively, are also the third person dual and third person plural markers on nominals in PaTani. In Tinani, unlike PaTani, they occur as agreement markers, and not as number markers on nominal arguments. It is possible that Tinani has borrowed them as verbal affixes from PaTani.

Development of separate dual and plural markers in some WH languages could be an effort to avoid ambiguity. TB languages are zero anaphora languages, where verb agreements carry information concerning core arguments. The same-number markers for all persons were bound to create ambiguity in reference. The incorporation of person information in the verb agreement system avoided this ambiguity.

4.2 PWH VERB AGREEMENT SCHEMA (TRANSITIVE)

All the WH languages considered in this study have the subject agreement system, either as the sole verb agreement system in the language, or as one possible agreement system. Beside the subject agreement markers, Gahri and Darmiya also have agreement markers which represent signs of the split-ergative system, and the no-agreement system. In this section I will present data which suggest that the split-ergative system is older, and could perhaps be reconstructed for PWH. Gahri, Darmiya and Kinnauri provide crucial data in this regard.

The description of the WH verb agreement morphology is presented in §3. Here I will present only the set of data which is essential to explicate the problem. I will begin with Gahri. Table 27 shows the distribution of the two agreement markers -*ki/ku* and -*n* in Gahri.

TABLE 27: GAHRI VERB AGREEMENT SCHEMA

SUBJECT	OBJECT		
	1	2	3
1		-*ki*-TNS	-*ki*-TNS
2			-(*n*)-TNS-*na* -(*n*)-TNS-*ni*
3	-(*ku*)-TNS		-(*n*)-TNS

The agreement markers -*ki/-ku* and -*n* occur in the past tense. The marker -*ki/-ku* occurs in first becoming second person, third person, and third becoming first person, and -*n* occurs with some non-first person subjects. With second person subjects, two agreement markers can be affixed to one verb (V-*n*-TNS-*na/ni*). The marker -*n* occurs with some non-first person subjects, and -*na* and -*ni* occur with second person singular and second person plural subjects, respectively. The distribution of -*ki/-ku* and -*n* differs from the subject agreement markers (-*na* and -*ni*, here) in terms of the position class in which they occur. The former set of agreement markers precedes the tense marker, whereas the subject agreement markers follow the tense marker. The -*ki* and -*n* forms seem to represent the older paradigm, where -*ki* occurred with first becoming second person, third person and -*n* with second

becoming third person. In modern Gahri, -n has lost the split-ergative distribution, and it now occurs with non–first person subjects.

In Darmiya, the agreement marker -n, occurs with first person plural, second person singular and second person plural subjects in the copula and the non-copula constructions in the past tense, where it precedes the tense marker. It seems to be a cognate of the agreement marker -n in Gahri. Unlike the past tense, in the present and future tenses the agreement marker -n occurs with first person plural and second person plural subjects, and -ni occurs with second person plural subjects. The distribution and placement of -ni in Darimya is the same as in Gahri. Table 28 shows the distribution of the verb agreement markers in Darmiya.

TABLE 28: DARMIYA VERB AGREEMENT SYSTEM

	1PL, 2SG, 2PL	1PL, 2SG, 2PL
Copula	V-n-TNS (past tense)	
Noncopula (transitive)		V-TNS-n/ni (present, future) V-n-TNS (past tense)

The following inferences can be drawn from the Gahri and Darmiya data, presented here, concerning the PWH agreement system. First, the presence of two agreement markers on one verb in Gahri suggests that one of them could be a later development. Second, the ordering of the split-ergative markers and the subject agreement markers suggests the antiquity of the split-ergative markers. The split-ergative markers are affixed closer to the verb, and they are followed by the tense marker in most cases, whereas the subject agreement markers follow the tense marker. The subject agreement markers are also the final affixes. Third, the split-ergative markers and their distribution in Gahri and Darmiya are cognates of the forms and the agreement system reconstructed for PTB, suggesting the possibility that the split-ergative system could be reconstructed for PWH.

Reconstructing a split-ergative system for PWH not only accounts for the agreement markers, and their distribution in these languages, but also accounts for what may be considered an anomaly in the split-ergative system. As mentioned earlier, -n in Darmiya also occurs with first person plural arguments, which in a split-ergative system should have a first person agreement marker. DeLancey (1989:318) notes that "the second person forms are less consistent". Several languages, outside WH too, have n- form for the first person becoming the second person verb agreement marker (for example Lohorong Rai (Weidert 1985:918) and Limbu (Weidert and Subba 1985:59-61)). Reconstructing a split-ergative system for PWH will not only explain the regularities in the verb agreement system, but will also account for the "variable marking of particular 2p. forms" (DeLancey 1989:326) in Darmiya.

The object agreement marker -č in Kinnauri provides indirect evidence in favour of reconstructing the split-ergative agreement system for PWH. The occurrence of the object agreement marker in Kinnauri is restricted to the first and second persons only. This is especially significant for reconstructing a split-ergative agreement system. The positioning of the object agreement marker (V-OBJ-TNS-SUB, V-OBJ-ASP AUX-TNS-AUX) also shows

that this is not a recent development. It is plausible that the object agreement marker in Kinnauri was originally a first or second person marker which became a third person marker, and later got reanalysed as an object agreement marker as a consequence of the development of the subject agreement system (see below for details).

The above discussion suggests that the split-ergative system is the older system, and could perhaps be reconstructed for PWH. Table 29 presents a preliminary schema of the PWH verb agreement system. It represents the reconstructed consonants of the singular forms only.

TABLE 29: PWH VERB AGREEMENT SCHEMA (TRANSITIVE)

SUBJECT	OBJECT		
	1	2	3
1		*-k / *-n	*-k
2			*-n
3	*-k		

The split-ergative agreement markers in Gahri and Darmiya are cognates, but this is not the case with the Kinnauri object agreement marker (-č). The latter does not resemble the putative PWH agreement marker. There are at least two feasible explanations concerning the development of the modern form in Kinnauri. First, the form is a secondary development in Kinnauri. It is possible that the Kinnauri object marker underwent a change, before it got realigned as an object marker (see below for details). Second, -č is a cognate of the older agreement marker -k, which got realised as -č as a result of palatalisation, and the vowel was lost in the process. At this stage there is no way to prove or disprove either of these alternatives. Data from other dialects of Kinnauri will be decisive in this regard.

The questions that can now be raised are: assuming that PWH had a split-ergative system; how did the split-ergative markers become the non-final suffixes, and what accounts for the regularity with which subject agreement markers occur in the WH languages?

It is possible that the subject agreement system in WH is a consequence of the reanalysis of the older copulas as tense markers. The subject agreement markers, as mentioned earlier, are regularly suffixed to the copulas. Once the language started using copulas as tense markers, copulas (along with the subject agreement markers) started occurring at the end of a finite verb, giving rise to a combination of the split-ergative system and the subject agreement system. This development can schematically be shown as follows.

        Stage I   V-SUB
        Stage II   V-SUB COP-SUB
        Stage III   V-SUB-TNS-SUB

Stages II and III are found in Kinnauri, and III is found also in Gahri and Darmiya. It is plausible that once this system got stabilised, the older split-ergative marker, in Kinnauri, realigned itself to the subject agreement system, whereas traces of the older split-ergative system are still found in Gahri and Darmiya.

Sunwari, a language belonging to the Kiranti group of the TB language family, provides evidence in favour of this suggestion. Genetti (1988) describes the Sunwari verb agreement system. It shows striking resemblance with the Gahri and Darmiya data presented above. In Sunwari, as in Gahri and Darmiya, two agreement markers can be affixed to one transitive verb. Such verbs have the structure V-TNS-AGR1-AGR2. Two separate sets of agreement markers occur in the position classes AGR1 and AGR2. The agreement markers which occur in the final position are the same (with minor differences) as the agreement markers with intransitive verbs. The distribution of AGR1 "is somewhat idiosyncratic without a clear synchronic organisational pattern" (Genetti 1988:81). Interestingly, the first person singular object marker in AGR1 position is *yi* in Sunwari.

It is not clear at this stage if the stage II (i.e. V-SUB COP-SUB) should be reconstructed for PWH, or if it is a later development. It is, however, obvious that this structure is not a recent development in WH. The consistent subject agreement system in WH, and the cognate morphology suggest early development of this structure in the history of WH. The fact that subject agreement is regularly found in WH languages, and not found in neighbouring TB languages such as Tod, a variety of Tibetan, suggests that even if the subject agreement system is a secondary development, it is not a complete innovation that each of these languages underwent independently. Seeds of the subject agreement system were present in PWH.

REFERENCES

Bailey, T.G., 1909, A brief grammar of the Kinnauri language. *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 63:661-687.

Bauman, James, 1975, Pronouns and pronominal morphology in Tibeto-Burman. PhD dissertation, University of California, Berkeley.

Benedict, Paul K., 1972, *Sino-Tibetan: a conspectus*. Cambridge: Cambridge University Press.

Caughley, Ross, 1982, *The syntax and morphology of the verb in Chepang. PL*, B-84.

Delancey, Scott, 1989, Verb agreement in Proto Tibeto-Burman. *Bulletin of the school of Oriental and African Studies* 52/21:315-333.

Egerod, Søren, 1973, Review of *Sino-Tibetan: a conspectus*, by Paul K. Benedict. *Journal of Chinese Linguistics* 1:498-505.

Francke, A.G., 1909, Tabellen der Pronomina und Verba in den drei Sprachen Lahoul's: Bunan, Manchad, und Tinan. *Zeitschrift der Deutschen Morgenländischen Gesellschaft.* 63:65-87.

Genetti, Carol, 1988, Notes on the structure of the Sunwari transitive verb. *Linguistics of the Tibeto-Burman Area* 11/2:62-92.

Joshi, Pandit Tika Ram and H.A. Rose, 1909, *A grammar and dictionary of Kanawari*. Calcutta: Baptist Mission Press.

Konow, Sten, 1909, The Tibeto-Burman sub-family. In G.A. Grierson, ed. *Linguistic Survey of India*, vol.I, 53-80. Calcutta, India: Superintendent of Government Printing. [Reprint: 1967, 1973. Delhi: Motilal Banarasidass]

Maspero, Henri, 1947, Notes sur la morphologie du tibeto-birman et du Munda. *Bulletin de la Société de Linguistique de Paris* 44:155-185.

Neethivanan, J., 1971, *Survey of Kanauri in Himachal Pradesh*. Calcutta: Office of the Registrar General.

Nishi, Yoshio, 1990, The distribution and classification of Himalayan languages (Part I). *Bulletin of the National Museum of Ethnology* 15/1:265-337 [in Japanese].

Saxena, Anju, 1991, Tone in PaTani and Central Tibetan: Parallel developments? *Linguistics of the Tibeto-Burman Area* 14/1:129-135.

1992, Finite verb morphology in Tibeto-Kinnauri. PhD dissertation, University of Oregon, Eugene.

Shafer, Robert, 1955, Classification of Sino-Tibetan languages. *Word* 11:94-111.

1966, *Introduction to Sino-Tibetan*. Wiesbaden: Otto Harrassowitz.

Sharma, D.D., 1988, *A descriptive grammar of Kinnauri*. Delhi: Mittal Publications.

1989a, *Tribal languages of Himachal Pradesh*. 2 vols. Delhi: Mittal Publications.

1989b, *Tibeto-Burman languages of Uttarakhand*. 2 vols. Delhi: Mittal Publications.

Sharma, Suhnu R., 1987, Non-Tibeto-Burman features in PaTani. *Linguistics of the Tibeto-Burman area* 10/1:126-129.

Thurgood, Graham, 1985, Pronouns, verb agreement systems, and the subgrouping of Tibeto-Burman. In Graham Thurgood, James A. Matisoff and David Bradley, eds *Linguistics of the Sino-Tibetan area: the state of the art,* 376-400. *PL*, C-87.

Weidert, Alfons, 1985, Paradigmatic typology and its application to verb agreement analysis. In Pieper and Stickel, eds *Studia Linguistica Diachronica et Synchronica,* 903-936. Berlin: Mouton de Gruyter.

Weidert, Alfons and B. Subba, 1985, *Concise Limbu grammar and dictionary*. Amsterdam: Lobster Publications.

Zoller, Claus Peter, 1983, *Die Sprache der Rang Pas von Garwhal: Grammatik, Texte, Wörterbuch*. Wiesbaden: Otto Harrassowitz.

SEMANTICALLY RELATED VOWEL GRADATION IN SUNWAR AND CHEPANG

R.C. CAUGHLEY

1. INTRODUCTION

This study examines the relationship between vowel gradation in words of various grammatical classes, and certain semantic features, for the Chepang language of Nepal.[1] Marlene Schulze (1987) has described such a vowel gradation for a subclass of adverbials in Sunwar.[2] As I will show, for Chepang the use of vowel gradation extends beyond the adverbial class. Before discussing the situation in more detail, however, it is necessary to define some of the terms involved, since they tend to be used with a variety of meanings.

2. ONOMATOPOEIA

In common usage onomatopoeic forms are those that are sound-imitative, such as 'crack', 'bang' and 'tinkle' in English. Sound-imitative roots can be found in words of different grammatical classes – adjectives, nouns, verbs and adverbs, as for example:

---

[1] For a more complete description of the Chepang language see Caughley (1982) and Caughley (forthcoming).

[2] The phonemic inventories of Sunwar and Chepang languages are:

		Sunwar					Chepang			
Stops	p	t	ṭ	k	'		p	t	k	ʔ
	ph	th	ṭ	kh			ph	th	kh	
	b	d	ḍ	g			b	d	g	
	bh	dh	ḍh	gh			bh	dh	gh	
Affricates			ts						c	
									j	
									ch	
									jh	
Fricatives		s	sh	h				s	h	
			dz							
Nasals	m	n	ŋ				m	n	ŋ	
Liquid		l						l		
Flap		r						r		
Glides	w	y					w	y		
Vowels	i, ii	u, uu					i	u		
	e, ee	o, oo					e	ə		
	a, aa						a	o		
Tone	' before word = high tone (for Sunwar)									

95

'a banging door'

'It went off with a bang.'

'He banged the door.'

'The cracker went off bang!'

In many languages, including those of South Asia, sound-imitative roots occur most commonly as a subset of the class of adverbs. That is, they function as modifiers of verbs, or forms derived from verbs, but they are often distinguished phonologically and/or syntactically from other adverbs. This adverbial subclass is often called 'onomatopoeic' (e.g. Emeneau 1969) However an examination of the data shows that, while many roots in this subclass are indeed sound-imitative, a considerable number have no reference at all to sounds. To use some of Emeneau's own examples (from Kota, Emeneau 1969):

dag dag	(flame) to burn brightly	
taṛ taṛ	(sun) rises or sets	
tan taṇk	(skin) feels smooth and satiny	
gam gam	to be fragrant	

It will be noted that the first two of these examples involve visually sensed events, while the third example has to do with the sense of touch and the fourth to the sense of smell. None of these can be related easily to sounds in any respect. Emeneau (1969:284), of course, was aware of this and states: "We are dealing only in the most marginal way with blatantly sound-imitative forms (like the English 'choochoo' or the like). Perhaps it would be more just to say the class denotes various types of sensation...".

Partly for this reason words of this category have been given a variety of names by different authors – including 'ideophones', 'expressives' and 'phonaesthemes'. Because of the close association of the term 'onomatopoeic' with sound-imitation, I will henceforth use the term 'ideophone' for this special subset of adverbs which includes both sound-imitative and non-sound-imitative roots. Matisoff (1989), following Japanese tradition, divides 'reduplicated sound symbolic expressions' into two major classes: *giseigo* 'sound-imitative', and *gitaigo* 'attitudinals or imitate-attitude words'. As I will show further below, such a dichotomy does not well represent the situation in Chepang (and possibly not for many languages). All ideophones are, in a more general sense, 'attitudinal' in that they normally occur in speech when the speaker has emotionally involved himself in what he is talking about – that is probably why such expressions are usually absent from unemotively expressed speech, and from writing. In other words ideophones are used to help the speaker convey his feelings and impressions concerning the situation.

Chepang ideophones fall into at least three major classes: sound-imitative, visual imitative and emotive (this last being closest to the *gitaigo* attitudinals).

Visual-imitative expressions may be divided into two subclasses, representing movement and static images respectively. In the first of these subclasses are words like *phiŋ* 'take off like a rocket/get away smartly', where the plosive initial represents the sudden take off and the continuant final the more gradual disappearance into the distance. Another example is *pyolololo* which represents a trotting motion (as of a jackal). In neither case is any sound necessarily involved, rather the sound pattern of the word reflects the visual pattern of movement. In the second subclass are expressions like *kriniŋ kranaŋ* 'stand upright together – one small person and one large' where the contrast in vowels repesents the contrast in size,

and *grugu thugu* which represents the visual impression of people crouched or huddled around food or a fire.

3. SOUND SYMBOLISM

Closely associated with the category described above is what has been called 'sound symbolism', in which certain sounds (or the physical gestures that produce them) are associated with particular physical features of the concept referred to. In fact, as Gregerson (1984, fn.4) points out, 'sound iconism' would be a better term for this relationship because of the non-arbitrariness of the sign-referent relation.

In particular, as Gregerson and others have shown, the size of the vocal cavities used to produce the sounds may be directly related to the physical size of the referent. Moreover, the occurrence of such a relationship is very widespread – it is found in Africa as well as in Asia (Gregerson 1984). Sapir (1929) raises the question whether the basis of this symbolism is acoustic or kinesthetic or both. It seems likely, however, that the starting point at least for such iconism is acoustic, since small cavities, like small objects, tend to produce high-frequency sounds. For a complex cavity such as the vocal tract the position is, of course, more complicated. Nevertheless a vowel sound made with the tongue in a high forward position does have most of its energy concentrated in the high frequency section of the spectrum and the best imitation of the 'clinking' of a coin or the 'tinkling' of a small bell is made with the oral cavity reduced in size.

Semantically determined vowel gradation is one expression of this sound iconism; another expression of this is found in consonant variation (Gregerson 1984, Schultze 1987). This paper seeks to examine the first of these – semantically related vowel gradation.

4. SUNWAR

According to information given in Schulze (1987:64) ideophones in Sunwar are themselves divided into two kinds:

(1) Words which correspond very closely to ideophones as defined above – a subset of adverbials, many of which are sound-imitative, and which are syntactically marked. Words of this subclass (which Schultze calls 'onomatopoeic') are always followed by *'pa* (< *'pa.tsa* 'to do'). Many of her examples are sound-imitative:

shap shap 'pa phiiktsa	to sweep with a swish
dok dok 'pa kruptsa	empty a big bottle

The examples also include reduplicative representation of repeated action:

'koroŋ 'pa dim.tsa	(for a tree) to fall of its own accord
'koroŋ 'koroŋ 'pa dim.tsa	(for two trees) to fall of their own accord)

Both of these features involve non-arbitrary relationships of sign and referent.

(2) Another subset of ideophones Schultze terms 'intensifiers'.

These differ from the 'onomatopoeic' subset in that:

i.   they can occur readily in formal speech, where the 'onomatopoeic' words are less appropriate;

ii.   they are more consistent in form and meaning than the onomatopoeic set, which tends to be subject to ideolectical variation;

iii. the significance of reduplication, and the phonological forms it can take, are more restricted;

iv.   unreduplicated forms are always monosyllabic, in contrast to those of the onomatopoeic set, which are often polysyllabic (as for the example 'koroŋ given above).

Both subsets exhibit a magnitude symbolism (or iconism) related to vowel position and also to consonant type.

The relation of magnitude to vowel position is given as: decreasing vowel height or fronting correlates with increasing magnitude of reference.

| Vowel height correlation with magnitude | $i < e < a;\quad u < o < a$ |
| Vowel fronting correlation with magnitude | $i < u;\qquad e < o$ |

The referent may be either the subject participant or the object or else the magnitude may relate to the intensity of an action or a condition. Most examples of vowel gradation occur in contrastive pairs, though occasionally sets of three are found. In combination with consonant variation, however, sets of four, five or more can occur.

Examples are:

Intensifiers occurring with:

i.   *thaaktsa* 'cut off a strip of bark' are (Schulze pers.comm.):

pre	(very small piece)
bre	(small piece)
bro	(big piece)
bra	(very big piece)

Intensifiers occurring with:

ii.   *hemtsa* 'fall off with a bump' are (Schulze 1987:73):

brel	(very small fruit, etc.)
brol	(ripe plums)
bral	(large fruit such as apples)

5. CHEPANG

Chepang has a large set of ideophones which are distinguished from other adverbs through being almost always followed by *tə* 'such, in such a way'. The set of non-ideophonic adverbial roots is, in fact, very small indeed – the only important ones are *ʔanə*

'much' and *bələ* 'a little', with the function of adverbs (and adjectives) being carried out normally by words derived from what are syntactically verb roots, or by ideophones. For example, the attributive verb root *jok-* 'fast, vigorous' can be used as an adverb or adjective in the following way:

Adverb	*ʔowʔ jok.to jyal.ʔa*	He went away fast.
Adjective	*ʔowʔ jok.ʔo manta.leʔ*	He is a vigorous person.

An ideophonic adverb, corresponding approximately to *jok-*, would be the word *phiŋ* given above, as in:

Ideophonic	*ʔowʔ phiŋ tə jyal.ʔa*	He got clean away.

As has already been pointed out, for Chepang, as for other languages, many (indeed most) ideophones are not sound-imitative. Many are visual-imitative, with the sound pattern (especially reduplication) of the ideophone imitating the visual pattern of the referent. For example:

Pyolololo tə wah.naʔ.	(The jackal) trots off.
Ŋənəna tə wah.naʔ.	(The elephant) walks swaying.
Ŋici ŋacya tə cyuŋʔ.naʔ.cə.	The two are sitting up straight (one small, one large).

Moreover, magnitude-related vowel gradation occurs freely in all kinds of ideophones, including those that are non-sound-imitative, as for example:

grigi thigi tə parʔ.ti mu.naʔ	sit crouched around (small persons)
grugu thugu tə parʔ.ti mu.naʔ	sit crouched around (medium sized persons)
grəgə thəgə tə parʔ.ti mu.naʔ	sit crouched around (large sized persons)

and

kiŋcidiŋ tə cyuŋʔ.naʔ	sit crouched (very small persons)
keŋcedeŋ tə cyuŋʔ.naʔ	sit crouched (small persons)
kuŋcuduŋ tə cyuŋʔ.naʔ	sit crouched (medium persons)
kəŋcədəŋ tə cyuŋʔ.naʔ	sit crouched (large persons)

Note that in these examples there is a magnitude–vowel position relation similar to that of Sunwar:

least $i < e < u < ə$ greatest

5.1 VERB ROOT VOWEL GRADATION

The above examples are taken from the ideophone set. What is highly unusual (and perhaps unique) about Chepang, however, is that non-sound-imitative verb roots also exhibit magnitude-related vowel gradation, as the following examples show:

krin-	pick out small burrs, etc.
krən-	pick out medium burrs, etc.
kran-	pick out large burrs, etc.

sim?-	subside slightly (fever)
syum?-	subside somewhat (fever)
sem?-	subside much (fever)
səm?-	subside completely (fever)

A sentence using an example from the second set would be:

?ow?.kay? ?əy.?o sim?.?a.thəy   His fever subsided slightly.

Here we have the gradations:

smallest $i < ə < a$ largest object referent

and

least $i < yu < e < ə$ greatest degree of the state referred to.

Another set is:

phep-	have very small blister
phəp-	have small blister
phop-	have large blister

This gives the gradation:

smallest $e < ə < o$ largest subject referent.

This gradation is consistent with the fact that, in the dialect of Chepang from which this set comes, /o/ is a low back vowel whereas /ə/ is mid central in position. There are many sets of verb roots showing this type of sound iconism in Chepang, usually pairs or sets of three. Some very common words show vowel gradation, such as mi- 'very small' and məy- 'small'.

Although there is a large number of sets of this sort the process is apparently no longer productive.

6. SUMMARY

Chepang, when compared with Sunwar, shows a considerably increased use of vowel gradation in formal language to indicate magnitude of referent or intensity of action, where these two features are related probably through sound imitation – a large object tends to produce a louder sound than a smaller one when it falls (or is struck).

Sunwar has its set of intensifiers which are indeed a class of ideophone, but are unusual in that they have become more fixed in usage and acceptable in formal speech. It is worth noting that, while Sunwar evidently does not have this vowel gradation in actual verb roots, the intensifiers that do have such gradation are often clearly related to verbs.

In contrast, for Chepang, vowel gradation is found not only in ideophones but also quite commonly in verb roots. These, of course, can occur freely in formal, unemotive speech. Such a feature, if productive, could be a prolific source of word families (see Matisoff 1978:16ff.) since it gives rise to sets of roots that are very closely related, both phonologically and semantically.

In conclusion then, semantically related vowel gradation or magnitude symbolism/iconism does appear to be a universal phenomenon, with this universality extending even to the particular way in which magnitude and vowel position are related (small objects with high and front vowels, large objects with low and back vowels). However, languages vary considerably in the extent to which this phenomenon can be used in more formal speech.

REFERENCES

Caughley, Ross C., 1982, The syntax and morphology of the verb in Chepang. *PL*, B-84.

forthcoming, *A dictionary of Chepang, Negap. PL*, C-129.

Emeneau, Murray B., 1969, Onomatopoetics in the Indian linguistic area. *Language* 45:274-299.

Gregerson, Kenneth, 1984, Pharynx symbolism and Rengao phonology. *Lingua* 62:209-238.

Matisoff, James Alan, 1978, Variational semantics in Tibeto-Burman: the organic approach to linguistic comparison. In F.K. Lehman, ed. *Occasional papers of the Wolfenden Society on Tibeto-Burman Linguistics* vol.6. Philadelphia: Institute for the Study of Human Relations.

1989, Tone, intonation and sound symbolism in Lahu: loading the syllable canon. *Linguistics of the Tibeto-Burman Area.* 12/2:147-163.

Sapir, Edward, 1929, A study of phonetic symbolism. *Journal of Experimental Psychology* 12:225-239.

Schulze, Marlene, 1987, Intense action adverbials in Sunwar: a verb intensifier system. *Linguistics of the Tibeto-Burman Area.* 10/1:63-85.

THE VERBAL MORPHOLOGY OF KULUNG

Gerard J. Tolsma

1. KULUNG AND ITS DIALECTS

Kulung (autonym *kulu riŋ*) is a Rāī language spoken by an estimated 7,000 people in about ten villages along the upper reaches of the Hoṅgu river (a tributary of the Dudhkosī in Solu Khumbu District of Sagarmāthā Zone in eastern Nepal. Rāī languages belong together with Limbu and Yakkha to the Kiranti branch of the Tibeto-Burman language group. Kiranti languages are characterised by an elaborate system of verbal endings, which is said to have been the result of a historical process that traditionally has been called 'complex pronominalising', that is the incorporation of personal pronouns in the verb system. Kiranti languages have been provisionally divided into Eastern Kiranti languages, that is languages spoken east of the Sālpā pass, and the rest (van Driem 1990). Limbu is classified as an Eastern Kiranti language and the majority of the Rāī languages, with the exception of Lohorung, Yakkha and Yamphu, belong to the rest. A more detailed picture of Kiranti phylogeny is suggested by Michailovsky's (1994) Kiranti sound laws. The closest relative of Kulung appears to be Khaling (Khāliṅ).

The main Kulung-speaking villages are Chemsīṅ and Cheskām. According to Kulung oral tradition the ancestor of the Kulung, *kʰəp*, had two sons, *cʰemsi* and *tamsi,* who founded the settlements Chemsīṅ and Cheskām respectively. The language in those two villages is considered by the Kulung to be the most original form of the language. From there, the offspring of *cʰemsi* and *tamsi* are said to have gone downstream at both sides of the Hoṅgu river to the villages that are now called Lujām, Gudel, Cācāluṅ, Nāmluṅ, Pelmāṅ, Buṅ, Chekmā, and Sātdi.

The Kulung living in the village Sotāṅ call their language *sɔttɔ riŋ* 'Sotāṅ language', but consider themselves to be Kulung and easily understand the Kulung language spoken in Chemsīṅ and Cheskām. The dialect spoken in Sotāṅ shows the result of historical phonological processes. For example the Kulung phoneme /u/, which has an allophone [ü] before the vowel /e/, was further fronted in the Sotāṅ dialect to merge with the phoneme /i/, so that Kulung *tu:-e* [tüː-e] 'it is' has become *ti:-e* in the Sotāṅ dialect. Other processes include secondary vowel length as a result of consonant cluster reduction. As a dialect study will be part of my forthcoming *Grammar of Kulung*, I will provide only the above example. Further downstream in the village Hulu, people refer to themselves as Nachering, but their language may also be considered as a dialect of Kulung with further dialectal developments. It is interesting that the oral tradition of the migration of the Kulung people is supported by

linguistic evidence. The present study is based on research conducted in the Kulung-speaking area in the summer of 1993 in the village Cheskām.[1]

2. PHONOLOGY

The inventory of Kulung phonemes is given below. Vowel phonemes are described in terms of the distinctive features height, backness, rounding and length. A total of fourteen vowel phonemes are found.

Vowel phonemes

	front unrounded	central unrounded	back rounded
high-close	*i/i:*		*u/u:*
mid-close	*e/e:*	*ə/ə:*	
mid-open			*ɔ/ɔ:*
low-open	*a/a:*		

Kulung distinguishes six places of articulation of consonant phonemes: glottal, dorso-velar, palatal, dental, and bilabial. These are produced by seven manners of articulation.[2]

---

[1]   The following abbreviations are used:

1	first person
2	second person
3	third person
→	indicates the direction of a transitive relationship
AG	agent of a transitive verb
DU	dual
EXC	exclusive
INC	inclusive
N1SG	non-first personal singular
NEG	negative
NPT	non-preterite
NSG	non-singular
PAT	patient of a transitive verb
PF	prefixal slot
PL	plural
PT	preterite
S	subject of an intransitive verb
SF	suffixal slot
SG	singular
Vi	intransitive verb
Vtr	transitive verb
Σ	verb stem

[2]   The phoneme symbols /c/ and /j/ represent the affricates [tɕ ~ ts] and [dʑ ~ dz] respectively.

Consonant phonemes

	bilabial	dental	palatal	dorso-velar	labial-velar	glottal
STOPS AND AFFRICATES						
voiceless unaspirated	*p*	*t*	*c*	*k*		*ʔ*
voiceless aspirated	*pʰ*	*tʰ*	*cʰ*	*kʰ*		
voiced unaspirated	*b*	*d*	*j*	*g*		
FRICATIVES						
voiceless grooved		*s*				
NASALS						
voiced	*m*	*n*		*ŋ*		
VIBRANTS						
voiced trill		*r*				
LATERAL APPROXIMANTS						
voiced		*l*				
CENTRAL APPROXIMANTS						
voiced		*y*			*w*	*h*

3. VERBAL MORPHOLOGY

An inflected verb form in Kulung consists of a verb stem to which affixes are attached, which express person and number agreement with one or two actants. According to the fixed pattern of paradigmatic verb stem alternation a number of different conjugation types can be identified. Up to now I have postulated four different conjugation types for both intransitive and transitive verbs. The number of verb stems found in each of the conjugation types varies from one to seven. The different verb stems do not have any meaning, but are the result of historical phonological processes. Personal pronouns may be added to verb forms for emphasis, but are not compulsory. Kulung distinguishes ten pronominal categories. Unlike most of the Kiranti languages, Kulung makes no formal distinction between a third person dual and plural. Instead it has a third person non-singular category. The personal pronouns are given below:

I	1SG	*kɔŋ*
we	1DU.INC	*kas*
we	1DU.EXC	*kaska*
we	1PL.INC	*keː*
we	1PL.EXC	*keːka*
you	2SG	*an*
you	2DU	*anci*
you	2PL	*anni*
s/he, it (here)	3SG	*ŋkə*
s/he, it (a little distant)	3SG	*naːkə*
s/he, it (distant)	3SG	*muŋkə*
s/he, it (down)	3SG	*nuːkə*
s/he, it (up)	3SG	*tɔːkə*
they (here)	3NSG	*ŋkəs*
they (a little distant)	3NSG	*naːkəs*
they (distant)	3NSG	*muŋkəs*
they (down)	3NSG	*nuːkəs*
they (up)	3NSG	*tɔːkəs*

The third person pronouns are demonstrative pronouns that have geography-dependent meaning. The pronouns *ŋkə* 's/he, it (here)' and *ŋkəs* 'they (here)' have proximal meaning, while the pronouns *naːkə* 's/he, it (a little distant)' and *naːkəs* 'they (a little distant)' have medio-distal meaning. The pronouns *muŋkə* 's/he, it (distant)' and *muŋkəs* 'they (distant)' have distal meaning. The pronouns *nuːkə* 's/he, it (down)' with the corresponding non-singular form *nuːkəs* 'they (down)' and *tɔːkə* 's/he, it (up)' with the corresponding non-singular form *tɔːkəs* 'they (up)' are elevation-dependent, meaning respectively 'at a lower altitude than the reference point' and 'at a higher altitude than the reference point'. Patients of transitive verbs and subjects of intransitive verbs appear in the absolute case marked by Ø, and agents of transitive verbs occur in the ergative case marked -*a*. A fixed set of personal endings is found for the intransitive, transitive and reflexive paradigms in both preterite and non-preterite tenses. An overview of the affixes of the intransitive and transitive conjugation is given below.

3.1 INTRANSITIVE CONJUGATION

Below the personal affixes are given of the intransitive conjugation of the intransitive verb *imma* (stem *ims-* ~ *im-*) 'to sleep'. The verb has two different stems. The stem *im-* occurs in negative preterite forms, and the stem *ims-* appears elsewhere. The endings of the intransitive and transitive paradigms are also shown diagrammatically in Tables 1 and 2.

	NON-PRETERITE	PRETERITE
1SG	*ims-ɔ:* Σ_1-1SG.NPT	*ims-ɔ* Σ_1-1SG.PAT.S.PT
	ims-ɔ:-nɔ Σ_1-1SG.NPT-NEG.NPT	*maŋ-im-ŋa* NEG.PT-Σ_2-1SG.NEG.PT
1DU.INC	*ims-ci* Σ_1-DU	*ims-a-ci* Σ_1-PT-DU
	ims-ci-nɔ Σ_1-DU-NEG.NPT	*maŋ-im-ci* NEG.PT-Σ_2-DU
1DU.EXC	*ims-ci-ka* Σ_1-DU-EXC	*ims-a-ci-ka* Σ_1-PT-DU-EXC
	ims-ci-ka-nɔ Σ_1-DU-EXC-NEG.NPT	*maŋ-im-s-ka* NEG.PT-Σ_2-DU-EXC
1PL.INC	*ims-ya* Σ_1-1PL.PAT.S.NPT	*ims-i* Σ_1-1PL.PAT.S.PT
	ims-ya-nɔ Σ_1-1PL.PAT.S.NPT-NEG.NPT	*maŋ-im-yi* NEG.PT-Σ_2-1PL.PAT.S.PT
1PL.EXC	*ims-ya-ka* Σ_1-1PL.PAT.S.NPT-EXC	*ims-i-ka* Σ_1-1PL.S.PT-EXC
	ims-ya-ka-nɔ Σ_1-1PL.PAT.S.NPT-EXC-NEG.NPT	*maŋ-im-yi-ka* NEG.PT-Σ_2-1PL.PAT.S.PT-EXC
2SG	*ims-e* Σ_1-N1SG.PAT.S	*ims-a-Ø* Σ_1-PT-N1SG.PAT.S
	ims-Ø-nɔ Σ_1-N1SG.PAT.S-NEG.NPT	*maŋ-im-na* NEG.PT-Σ_2-2SG.PT
2DU	*ims-ci* Σ_1-DU	*ims-a-ci* Σ_1-PT-DU
	ims-ci-nɔ Σ_1-DU-NEG.NPT	*maŋ-im-ci* NEG.PT-Σ_2-DU
2PL	*ims-ni* Σ_1-2PL.PAT.S	*ims-a-nni* Σ_1-PT-2PL
	ims-ni-nɔ Σ_1-2PL-NEG.NPT	*maŋ-im-ni* NEG.PT-Σ_2-2PL
3	*ims-e* Σ_1-N1SG.PAT.S	*ims-a-Ø* Σ_1-PT-N1SG.PAT.S
	ims-Ø-nɔ Σ_1-N1SG.PAT.S-NEG.NPT	*maŋ-im-Ø* NEG.PT-Σ_2-N1SG.PAT.S

TABLE 1: THE AFFIXES OF THE INTRANSITIVE CONJUGATION

	NPT	PT
1SG	Σ_1-$\mathfrak{I}$:	Σ_1-$\mathfrak{I}$
	Σ_1-$\mathfrak{I}$:-$n\mathfrak{I}$	man-Σ_2-ηa
1DU INC	Σ_1-ci	Σ_1-a-ci
	Σ_1-ci-$n\mathfrak{I}$	man-Σ_2-ci
1DU EXC	Σ_1-ci-ka	Σ_1-a-ci-ka
	Σ_1-ci-ka-$n\mathfrak{I}$	man-Σ_2-s-ka
1PL.INC	Σ_1-ya	Σ_1-i
	Σ_1-ya-$n\mathfrak{I}$	man-Σ_2-i
1PL.EXC	Σ_1-ya-ka	Σ_1-i-ka
	Σ_1-ya-ka-$n\mathfrak{I}$	man-Σ_2-i-ka
2SG	Σ_1-e	Σ_1-a-$\emptyset$
	Σ_1-$\emptyset$-$n\mathfrak{I}$	man-Σ_2-na
2DU	Σ_1-ci	Σ_1-a-ci
	Σ_1-ci-$n\mathfrak{I}$	man-Σ_2-ci
2PL	Σ_1-ni	Σ_1-a-nni
	Σ_1-ni-$n\mathfrak{I}$	man-Σ_2-ni
3	Σ_1-e	Σ_1-a-$\emptyset$
	Σ_1-$\emptyset$-$n\mathfrak{I}$	man-Σ_2-$\emptyset$

3.2 TRANSITIVE CONJUGATION

A full paradigm of the verb *kerma* 'to hit' is given below in both non-preterite and preterite tenses. The negative forms are listed below the corresponding affirmative forms.

	NON-PRETERITE	PRETERITE
1SG→2SG	*ker-yan* Σ-1SG→2.NPT	*ker-na* Σ-2SG.PT
	ker-yan-nɔ Σ-1SG→2.NPT-NEG.NPT	*maŋ-ker-na* NEG.PT-Σ-2SG.PT
1SG-→2DU	*ker-yan-ci* Σ-1SG→2.NPT-DU	*ker-n-ci* Σ-2SG.PT-DU
	ker-yan-ci-nɔ Σ-1SG→2.NPT-DU-NEG.NPT	*maŋ-ker-n-ci* NEG.PT-Σ-2SG.PT-DU

1SG→2PL	*ker-yan-ni* ∑-1SG→2.NPT-2PL	*ker-Ø-ni* ∑-PT-2PL
	ker-yan-ni-nɔ ∑-1SG→2.NPT-2PL-NEG.NPT	*maŋ-ker-ni* NEG.PT-∑-2PL
1NSG.3→2SG	*ker-e* ∑-N1SG.PAT.S	*ker-a-Ø* ∑-PT-N1SG.PAT.S
	ker-Ø-nɔ ∑-N1SG.PAT.S-NEG.NPT	*maŋ-ker-na* NEG.PT-∑-2SG.PT
1NSG.3→2DU	*ker-ci* ∑-DU	*ker-a-ci* ∑-PT-DU
	ker-ci-nɔ ∑-DU-NEG.NPT	*maŋ-ker-ci* NEG.PT-∑-2DU
1NSG.3→2PL	*ker-ni* ∑-2PL	*ker-a-nni* ∑-PT-2PL
	ker-ni-nɔ ∑-2PL-NEG.NPT	*maŋ-ker-ni* NEG.PT-∑-2PL
1SG→3SG	*ker-ɔ:* ∑-1SG.NPT	*ker-Ø-u* ∑-PT-3PAT
	ker-ɔ:-nɔ ∑-1SG.NPT-NEG.NPT	*maŋ-ker-ŋa* NEG.PT-∑-1SG.NEG.PT
1SG→3NSG	*ker-ɔ:-ci* ∑-1SG.NPT-3NSG	*ker-Ø-u-ci* ∑-PT-3PAT-3NSG
	ker-ɔ:-ci-nɔ ∑-1SG.NPT-3NSG-NEG.NPT	*maŋ-ker-ŋa* NEG.PT-∑-1SG.NEG.PT
1DU.INC-→3	*ker-c-u* ∑-DU-3PAT	*ker-a-c-u* ∑-PT-DU-3PAT
	ker-c-u-nɔ ∑-DU-3PAT-NEG.NPT	*maŋ-ker-ci* NEG.PT-∑-DU
1DU.EXC→3	*ker-c-u-ka* ∑-DU-3PAT-EXC	*ker-a-c-u-ka* ∑-PT-DU-3PAT-EXC
	ker-c-u-ka-nɔ ∑-DU-3PAT-EXC-NEG.NPT	*maŋ-ker-s-ka* NEG.PT-∑-DU-EXC
1PL.INC→3	*ker-am* ∑-1PL→3.NPT	*ker-Ø-u-m* ∑-PT-3PAT-PL.AG
	ker-am-nɔ ∑-1PL→3.NPT-NEG.NPT	*maŋ-ker-i* NEG.PT-1PL.PAT.S.PT
1PL.EXC→3	*ker-am-ka* ∑-1PL→3.NPT-EXC	*ker-Ø-u-m-ka* ∑-3PAT-PL.AG-EXC

	ker-am-ka-nɔ ∑-1PL→3.NPT-EXC-NEG.NPT	*maŋ-ker-i-ka* NEG.PT-∑-1PL.PAT.S.PT-EXC
2SG→3SG	*ker-ə* ∑-2SG.3→3.NPT	*ker-∅-u* ∑-PT-3PAT
	ker-ə-nɔ ∑-2SG.3→3.NPT-NEG.NPT	*maŋ-ker-na* NEG.PT-∑-2SG.PT
2SG→3NSG	*ker-ə-ci* ∑-2SG.3→3.NPT-3NSG	*ker-∅-ci* ∑-PT-3NSG
	ker-ə-s-nɔ ∑-2SG.3→3.NPT-3NSG-NEG.NPT	*maŋ-ker-na* NEG.PT-2SG.PT
2DU→3	*ker-c-u* ∑-DU-3PAT	*ker-a-c-u* ∑-PT-DU-3PAT
	ker-c-u-nɔ ∑-DU-3PAT-NEG.NPT-NEG.NPT	*maŋ-ker-n-ci* NEG.PT-∑-2SG.PT-DU
2PL→3	*ker-n-u-m* ∑-2PLA-3PAT-PL.AG	*ker-a-n-u-m* ∑-PT-2PL.AG-3PAT-PL.AG
	ker-n-u-m-nɔ ∑-2PL.AG-3PAT-PL.AG-NEG.NPT	*maŋ-ker-ni* NEG.PT-∑-2PL
3SG→3SG	*ker-ə* ∑-2SG.3→3.NPT	*ker-∅-u* ∑-PT-3PAT
	ker-ə-nɔ ∑-2SG.3→3.NPT-NEG.NPT	*maŋ-ker* NEG.PT-∑
(3→3)NSG	*ker-ə-ci* ∑-2SG.3→3.NPT-3NSG	*ker-∅-ci* ∑-PT-3NSG
	ker-ə-s-nɔ ∑-2SG.3→3.NPT-3NSG-NEG.NPT	*maŋ-ker-ci* NEG.PT-∑-3NSG
2SG.3→1SG	*ker-ɔ:* ∑-1SG.NPT	*ker-ɔ* ∑-1SG.PAT.S.PT
	ker-ɔ:-nɔ ∑-1SG.NPT-NEG.NPT	*maŋ-ker-ŋa* NEG.PT-∑-1SG.NEG.PT
2DU→1SG	*ker-ɔ:-ci* ∑-1SG.NPT-DU	*ker-∅-ɔ-ci* ∑-PT-1SG.PAT.S.PT-DU
	ker-ɔ:-ci-nɔ ∑-1SG.PAT.S-NEG.NPT	*maŋ-ker-ŋa-ci* NEG.PT-∑-1SG.NEG.PT-DU
2PL→1SG	*ker-ɔ:-ni* ∑-1SG.NPT-2PL	*ker-ɔ-ni* ∑-1SG.PAT.S.PT-2PL
	ker-ɔ:-ni-nɔ ∑-1SG.NPT-2PL-NEG.NPT	*maŋ-ker-ŋa-ni* NEG.PT-∑-1SG.NEG.PT-2PL

3→1DU.INC	*ker-ci*	*ker-a-ci*
	∑-DU	∑-PT-DU
	ker-ci-nɔ	*maŋ-ker-ci*
	∑-DU-NEG.NPT	NEG.PT-∑-DU
2.3→1DU.EXC	*ker-ci-ka*	*ker-a-ci-ka*
	∑-DU-EXC	∑-PT-DU-EXC
	ker-ci-ka-nɔ	*maŋ-ker-s-ka*
	∑-DU-EXC-NEG.NPT	NEG.PT-∑-DU-EXC
3→1PL.INC	*ker-ya*	*ker-i*
	∑-1PL.PAT.S.NPT	∑-1PL.PAT.S.PT
	ker-ya-nɔ	*maŋ-ker-i*
	∑-1PL.PAT.S.NPT-NEG.NPT	NEG.PT-∑-1PL.PAT.S.PT
2.3→1PL.EXC	*ker-ya-ka*	*ker-i-ka*
	∑-1PL.PAT.S.NPT	∑-1PL.PAT.S.PT-EXC
	ker-ya-ka-nɔ	*maŋ-ker-i-ka*
	∑-1PL.PAT.S.NPT-NEG.NPT	NEG.PT-∑-1PL.PAT.S.PT-EXC

3.3 MORPHEMIC ANALYSIS OF VERBAL AFFIXES

To account for the Kulung conjugational endings, one prefixal and seven suffixal slots must be posited.

(1) Prefixal slot

(i) Negative preterite morpheme in prefixal slot PF1

   man- ~ mam- ~ maŋ-   NEG.PT

The preterite negative prefix *man- ~ mam- ~ maŋ-* indexes negation and preterite tense in all persons, for example *maŋ-ker-na* (NEG.PT-hit-2SG.PT) 'I didn't hit you'. This morpheme has the following allomorphic distribution:

/man-/ → [mam-]/— ⎰ bilabial stops and bilabial nasal: [b-, p-, pʰ-, m-]

/man-/ → [maŋ-]/— ⎰ vowel [V-]
   glottal central approximant [h-]
   dorso-velar stops [g-, k-, kʰ-]
   dorso-velar nasal [ŋ-]
   labial-velar central approximant [w-]

Note that verb stems beginning with a vowel are always automatically preceded by a subphonemic hiatus glottal stop.

TABLE 2: THE AFFIXES OF THE TRANSITIVE CONJUGATION

patient →, agent ↓

	1SG	1DU.INC	1DU.EXC	1PL.INC	1PL.EXC	2SG	2DU	2PL	3SG	3NSG
1SG						Σ-yan / Σ-na	Σ-yan-ci / Σ-n-ci	Σ-yan-ni / Σ-Ø-ni	Σ-ɔː / Σ-Ø-u	Σ-ɔː-ci / Σ-Ø-u-ci
1DU.INC									Σ-c-u / Σ-a-c-u	
1DU.EXC						Σ-e / Σ-a-Ø	Σ-ci / Σ-a-ci	Σ-ni / Σ-a-nni	Σ-c-u-ka / Σ-a-c-u-ka	
1PL.INC									Σ-am / Σ-Ø-u-m	
1PL.EXC						Σ-e / Σ-a-Ø	Σ-ci / Σ-a-ci	Σ-ni / Σ-a-nni	Σ-am-ka / Σ-Ø-u-m-ka	
2SG	Σ-ɔː / Σ-Ø-ɔ								Σ-ə / Σ-Ø-u	Σ-ə-ci / Σ-Ø-ci
2DU	Σ-ɔː-ci / Σ-Ø-ɔ-ci								Σ-c-u / Σ-a-c-u	
2PL	Σ-ɔː-ni / Σ-Ø-ɔ-ni		Σ-ci-ka / Σ-a-ci-ka		Σ-ya-ka / Σ-i-ka				Σ-n-u-m / Σ-a-n-u-m	
3SG	Σ-ɔː / Σ-Ø-ɔ	Σ-ci / Σ-a-ci		Σ-ya / Σ-i		Σ-e / Σ-a-Ø	Σ-ci / Σ-a-ci	Σ-ni / Σ-a-nni	Σ-ə / Σ-Ø-u	Σ-ə-ci / Σ-Ø-ci
3NSG	Σ-ɔː / Σ-Ø-ɔ	Σ-ci / Σ-a-ci		Σ-ya / Σ-i		Σ-e / Σ-a-Ø	Σ-ci / Σ-a-ci	Σ-ni / Σ-a-nni		Σ-ə-ci / Σ-Ø-ci

(2) Suffixal slots

The suffixes assigned to the suffixal slots, diagrammed in Table 3, are explained and illustrated by examples taken from the transitive paradigm of the verb *kerma* 'to hit'.

(i) Tense morphemes in suffixal slot SF1

    a ~ Ø   PT

The tense marker *-a ~ Ø* (PT) occurs in the affirmative forms of the preterite tense, for example *ker-a-c-u-ka* (hit-PT-DU-3PL-EXC) 'we.DU.EXC hit him'. It has a regular zero allomorph *Ø* before a vocalic affix, that is *vocalis ante vocalem corripitur*, and a paradigmatically conditioned zero allomorph *Ø* in the transitive forms 1SG→2PL, 2SG→3NSG and (3→3)ᴺˢᴳ. The corresponding non-preterite forms of these forms, however, are marked for non-preterite tense by a portmanteau morpheme.

TABLE 3: SUFFIXAL SLOTS AND SLOT FILLERS

SF1	SF2	SF3	SF4	SF5	SF6	SF7
	-ɔ: 1SG.NPT					
	-ɔ 1SG.PAT.S.PT					
	-ŋa 1SG.NEG.PT					
	-ya 1PL.PAT.S.NPT		*-u* 3PAT	*-ci ~ -s* 3NSG		
a ~ Ø PT	*-i ~ -yi* 1PL.PAT.S.PT	*-ci ~ -c ~ -s* DU	*-ə* 2SG.3→3.NPT	*-m* PL.AG	*-ka* EXC	*-nɔ* NEG.NPT
	-e ~ Ø N1SG.PAT.S		*-am* 1PL→3.NPT	*-ni ~ -nni* 2PL		
	-na ~ -nna ~ -n 2SG.PT					
	-yan 1SG→2.NPT					
	-n 2PL.AG					

(ii) Person morphemes in suffixal slot SF2

Suffixal slot SF2 is defined by the relative position of -ɔː (1SG.NPT), -ŋa (1SG.NEG.PT), -yan (1SG→2.NPT) and -na ~ -nna ~ -n (2SG.PT) in the suffixal string, because they follow the preterite tense marker -a ~ Ø, but precede the dual morpheme -ci ~ -c ~ -s. Of the remaining morphemes in the suffixal slot the following can be said: the tensed first person plural patient/subject morphemes, -ya (non-preterite) and -i ~ -yi (preterite), precede the exclusive morpheme -ka in suffixal slot SF6. The non-first person -e ~ Ø precedes the third person non-singular morpheme -ci ~ -s (3NSG) in suffixal slot SF5 and the second person plural agent suffix -n occupies a position anterior to the third person patient morpheme-u in suffixal slot SF4. On the basis of the semantic common denominator of person agreement, these suffixes have been assigned to this slot.

-ɔː	1SG.NPT
-ɔ	1SG.PAT.S.PT
-ŋa	1SG.NEG.PT
-ya	1PL.PAT.S.NPT
-i ~ -yi	1PL.PAT.S.PT
-e ~ Ø	N1SG.PAT.S
-na ~ -nna ~ -n	2SG.PT
-yan	1SG→2.NPT
-n	2PL.AG

The first person singular non-preterite actant morpheme -ɔː indexes a first singular actant in the non-preterite tense, for example ker-ɔː-ci-nɔ (hit-1SG.NPT-DU-NEG.NPT) 'youDU don't hit me'.

The first singular preterite patient/subject morpheme -ɔ indexes a first singular subject or patient morpheme in the preterite tense, for example ker-ɔ (hit-1SG.PAT.S.PT) 'youSG hit me'.

The first negative preterite singular morpheme -ŋa occurs in negative preterite forms with a first singular actant, for example maŋ-ker-ŋa-ci (NEG.PT-hit-1SG.NEG.PT-DU) 'youDU didn't hit me'.

The non-preterite first plural patient/subject morpheme -ya indexes a first plural patient or subject in the non-preterite tense, for example ker-ya-ka (hit-1PL.PAT.S.NPT-EXC) 'you, s/he, they hit usPAT.EXC'.

The preterite first plural patient/subject morpheme -i ~ -yi indexes a first plural patient or subject in the preterite tense, for example ker-i (hit-1PL.PAT.S.PT) 's/he (they) hit us'. This suffix has an allomorph -yi after stems ending in a vowel or nasal.

The non-first person singular patient/subject morpheme -e ~ Ø indexes a non-first singular patient or subject, for example ker-e (hit-N1SG.PAT.S) 'we hit you'. This suffix occurs only in intransitive forms with a non-first person singular subject and in 1DU.EXC→2SG, 1PL.EXC→2SG and 3→2SG forms; that is, it does not occur in forms in which involvement of non-first singular patient or subject is indicated by another morpheme, viz. in preterite 1SG→2SG forms where the second person singular is indexed by the morpheme -na (2SG.PT). Moreover, the suffix -e is epenthetic and elides to Ø before the

non-preterite negative suffix -*nɔ*, the conditional suffix -*dɔ* and the reported speech suffix -*tʰe*. The non-first person singular patient/subject morpheme -*e* ~ *Ø* also has a zero allomorph *Ø* after the preterite marker -*a* as well as before the vocalic suffixes -*ɔ:* (1SG.NPT), -*ə* (2SG.3→3.NPT) and the third person patient morpheme -*u*. Furthermore, this morpheme has a paradigmatically conditioned zero allomorph *Ø* in the third person negative preterite subject.

The second person singular morpheme -*na* ~ -*nna* ~ -*n* indexes a second singular actant in the preterite tense, for example *ker-na* (hit-2SG.PT) 'I hit you'. The suffix has an allomorph -*nna* after stems ending in a vowel and an allomorph -*n* when followed by the dual suffix -*ci*.

The 1SG→2 portmanteau morpheme -*yan* indexes the transitive relationship between a first singular agent and a second person patient in the non-preterite tense, for example *ker-yan-ci* (hit-1SG→2.NPT-DU) 'I hit you^DU'.

The second person plural agent morpheme -*n* (2PL.AG) indexes a second plural agent in 2PL→3 forms, with the exception of the negative preterite 2PL→3 form, for example *ker-a-n-u-m* (hit-PT-2PL.AG-3PAT-PL.AG) 'I hit you^PL'.

(iii) The dual morpheme in suffixal slot SF3

Suffixal slot SF3 is defined by the relative position of the dual morpheme -*ci* ~ -*c* ~ -*s* in the suffixal string; that is, the suffix in this slot may follow the person morphemes in suffixal slot SF2, but precedes the third person patient marker -*u* in suffixal slot SF4.

The dual morpheme -*ci* ~ -*c* ~ -*s* indexes dual actant number, for example *ker-c-u* (hit-DU-3PL) 'we^DU.INC hit him'. This suffix has an allomorph -*c* when followed by the third person patient suffix -*u* and the paradigmatically conditioned allomorph -*s* in any negative preterite first person dual exclusive form.

(iv) Third person patient morphemes in suffixal slot SF4

This slot contains the third person patient marker -*u* and portmanteau morphemes indexing a third person patient. Suffixal slot SF4 is defined by the relative position of the third person patient marker -*u* in the suffixal string. On the basis of positional criteria alone, however, the three portmanteaus could be assigned to any slot following the tense slot SF1, but preceding the third person non-singular morpheme -*ci* ~ -*s* in suffixal slot SF5. The three portmanteaus are assigned to suffixal slot SF4, therefore, solely on the basis of formal and semantic criteria.

-*u*	3PAT
-*ə*	2SG.3→3.NPT
-*am*	1PL→3.NPT

The third person patient morpheme -*u* indexes a third person patient, for example *ker-Ø-u* (hit-PT-3PAT) 'I hit him'.

The 2SG.3→3 portmanteau morpheme -*ə* (2SG.3→3.NPT) indexes a transitive relationship between a second person singular or third person agent and a third person

patient in the non-preterite tense, for example *ker-ə-ci* (hit-2SG.3→3.NPT-NSG) 'you hit them'.

The 1PL→3.NPT portmanteau morpheme *-am* indexes a transitive relationship between a first person plural agent and a third person patient in the non-preterite tense, for example *ker-am-nɔ* (hit-1PL→3.NPT-NEG.NPT) 'we[PL.INC] don't hit him (them)'.

(v) Number morphemes in suffixal slot SF5

Suffixal slot SF5 is defined by the relative position of the plural agent morpheme *-m* in the suffixal string, because this morpheme precedes the exclusive morpheme *-ka* in suffixal slot SF6, but follows the third person patient marker *-u* in suffixal slot SF4. Of the remaining morphemes in the suffixal slot the following can be said: the third person non-singular morpheme *-ci* and the second plural actant morpheme *-ni ~ -nni* follow the third person patient morpheme *-u* in suffixal slot SF4 and the first singular non-preterite actant morpheme *-ɔ:* in suffixal slot SF2 respectively, and precede the negative non-preterite morpheme *-nɔ* in suffixal slot SF6.

-ci	3NSG
-m	PL.AG
-ni ~ -nni	2PL

The third person non-singular morpheme *-ci ~ -s* indexes non-singular number of a third person actant, for example *ker-ɔ:-ci* (hit-1SG.NPT-3NSG) 'I hit them'. This morpheme has the allomorph *-s* in the negative non-preterite 2SG→3NSG and (3→3)[NSG] forms.

The plural agent morpheme *-m* indexes plural number of a first or second person agent, for example *ker-n-u-m* (hit-2PL.AG-3PAT-PL.AG) 'you[PL] hit him'.

The second person plural morpheme *-ni ~ -nni* indexes a second person plural actant in intransitive forms and in 1→2PL, 2PL→1SG and 1NSG.3→2PL forms, for example *ker-ni* (hit-2PL) 'I hit you[PL]'. The allomorph *-nni* occurs after the tense marker *-a* and after stems ending in a vowel.

(vi) The exclusive morpheme in suffixal slot SF6

Suffixal slot SF6 is only occupied by the morpheme *-ka*. The morpheme follows the plural agent morpheme *-m* in suffixal slot SF5 and precedes the negative non-preterite morpheme *-nɔ* in suffixal slot SF7.

-ka	EXC

The exclusive morpheme *-ka* indexes exclusive actant except in 1→2 forms, for example *ker-ci-ka* (hit-DU-EXC) 'you didn't hit us[PL.EXC]'. The Kulung paradigm does not distinguish the various 1NSG→2 forms from the corresponding 3→2 forms.

(vii) The negative non-preterite morpheme in suffixal slot SF7

The last suffixal slot is occupied by the negative non-preterite morpheme *-nɔ* (NEG.NPT).

-nɔ	NEG.NPT

The non-preterite negative morpheme *-nɔ* (NEG.NPT) indexes negation in the non-preterite tense, for example *ker-ɔː-nɔ* (hit-1SG.NPT-NEG.NPT) 'he doesn't hit me'.

REFERENCES

van Driem, George, 1990, The fall and rise of the phoneme */r/* in Eastern Kiranti: sound change in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 53/1:83-86.

Michailovsky, Boyd, 1994, Manner vs place of articulation in the Kiranti initial stops. In Hajime Kitamura, Tatsuo Nishida and Yasuhiko Nagano, eds. *Current issues in Sino-Tibetan Linguistics*, 766-772. Osaka.

Tolsma, Gerard J., forthcoming, *Grammar of Kulung*.

TRIPLICATED VERBAL ADJUNCTS IN BANTAWA[1]

NOVEL KISHORE RAI AND WERNER WINTER

Among the elements specifying properties of verbal actions in Bantawa, a 'complex pronominalising' Tibeto-Burman language of eastern Nepal, forms deserve special attention in which a nucleus (kernel) K is repeated twice so that a triplet KKK (followed by a suffix -*wa*) results. Examples are:

(a)	*mükmükmükwa khap-*	weep profusely	: *mük*	eye
(b)	*kakkakkakwa let-*	burn with a red glow	: *kak*	live coal
(c)	*cekcekcekwa wa ta-*	rain continuously	: *cek-*	?
(d)	*nunnunnunwa dutt-*	feel soft touch	: *nun-*	?

The nuclei differ in status: in the case of (a) and (b) they exist as free forms or as monolexemic stems as part of the lexicon of contemporary Bantawa; in the case of (c) and (d), they do not – only the triplets extended by -*wa* are found as actually attested words.

The distinction made here between forms with lexicalised and with unlexicalised nuclei is of course one based on current use in the language of at least one native speaker; yet it seems appropriate to apply it consistently to all triplets which one of us, Novel Kishore Rai, has been able to collect from his own dialect of Bantawa (as spoken at Rabi, Panchthar).

In the examples below, each triplet + -*wa* is represented by 'X'.

1. This group, comprising triplets relatable to simple nuclei recurring as free forms or as monolexemic stems, consists of the following items:

>(1)   *con-* 'be eager to do something' in:[2]
>
>*?O cha konma   conconconwa cüyang.*
>this child to.move.about   X   is.CONT
>The child is eager to move about.

---

[1]   This is a corrected and slightly amended version of a paper which appeared in *Linguistic Fiesta* (Festschrift Hisao Kakehi) in 1990. Permission has been received from the publishers for this *PL* version.

[2]   The practical orthography used in this paper is that employed in Werner Winter – Novel Kishore Rai, *An analytical dictionary of Bantawa* (to appear).

© Novel Kishore Rai and Werner Winter   119

(2)   *cük-* 'save; be stingy' in:

    Khamma cükcükcükwa müye!
    money   X   do
    Use the money like a miser!

(3)   *chek-* 'come to one's senses' in:

    ʔO cha   chekchekchekwa cepyang.
    this child X   speaks.CONT[3]
    What the child says makes sense.

(4)   *chop-* 'be dry' in:

    ʔÜng cha   dham yüngma chopchopchopwa lisang.
    my   child falls to.say   X   I.became
    I was afraid that my child would fall.

(5)   *chük-* 'pinch' in:

  a.   *ʔO bungga chükchükchükwa nam.*
      this flower X   smells
      This flower smells very bad.

  b.   *ʔOda chükchükchükwa cung luyang.*
      here X   cold feels.CONT
      It is very cold here.

(6)   *dek-* 'cover; block' in:

  a.   *Kuiroʔa dekdekdekwa lisa.*
      by.cloud X   became
      The cloud caused it to be very dark.

  b.   *Nam lunta kiya dekdekdekwa lisa.*
      sun se   and X   became
      The sun set and it became very dark.

(7)   *hak-* 'be warm' in:

  a.   *ʔO khimhutda   hakhakhakwa luyang.*
      this in.house.interior X   feels.CONT
      It is very warm inside the house.

  b.   *ʔO tit   hakhakhakwa luyang!*
      this cloth X   feels.CONT
      This material is very warm.

(8)   *hük* 'wind' in:

  a.   *Khana hükhükhükwa mancüde!*
      thou   X   do.not.do!
      Don't move about aimlessly.

---

[3]   Abbreviations used in these examples are: ABSTR (abstract), CONT (continuous), ERG (ergative), NSG (non-singular) and REFL (reflexive).

b. *Mo bhürüdu   hükhükhükwa luyang.*
that on.top.of.hill X   feels.CONT
It is very windy up on the hill.

(9)   *kak* 'live coal' in:

Moko mi   kakkakkakwa letyang.
that fire X   burns.CONT
The fire is burning a deep red.

(10)   *kup-* 'keep warm; incubate' in:

Moko kupkupkupwa ?imsa.
he/she X   slept
He/she slept nice and warm.

(11)   *küp-* 'cut into pieces' in:

?Üng buk   küpküpküpwa tukyang.
my   belly X   hurts.CONT
My stomach hurts very badly.

(12)   *küt-* 'affect by being rotten' in:

?Am khowa kütkütkütwa namyang.
thy   wound X   smells.CONT
Your wound smells abominably.

(13)   *khok-* 'be aged' in:

?O müna khokkhokkhokwa thung.
this man X   coughs
This man keeps coughing.

(14)   *khük-* 'be bitter' in:

?O puyupma khükkhükkhükwa luyang.
this cucumber X   feels.CONT
This cucumber tastes very bitter.

(15)   *mak-* 'be black' in:

?ayü nampük makmakmakwa ya?ang.
today night X   is.CONT
It is pitch dark tonight.

(16)   *mok-* 'kick' in:

?Ü   tang tuk   kiya moko mokmokmokwa ten.
his/her head hurts and he/she X   falls.down
He/she has a headache and collapses.

(17)   *mük* 'eye' in:

Mo müna mükmükmükwa khapyang.
that man X   weeps.CONT
The man is weeping profusely.

(18)   *nop-* 'touch' in:

    ʔÜ   komaʔa   mo cha nopnopnopwa mü.
    his/her grandmother.ERG that child X   did
    His/her grandmother treated the child with much affection.

(19)   *ngep-* 'brand' in:

    ʔÜng buk ngepngepngepwa tuk.
    my   belly X   hurts
    I have severe stomach pain.

(20)   *rak* 'heat' in:

    Mi rakrakrakwa letyang.
    fire X   burns.CONT
    The fire is very strong.

(21)   *rep-* 'become strong' in:

    ʔO cha   repreprepwa kama mü.
    this child X   work does
    This child words very hard.

(22)   *seng-* 'clean' in:

    Moko khim   sengsengsengwa katyang.
    that   house X   appears.CONT
    That house looks very clean.

(23)   *sip-* 'cease; be closed' in:

    a.   *ʔAyi sipsipsipwa nam nuyang.*
        today X   sun is.good.CONT
        The sunshine is very weak today.

    b.   *Tuwa   kiya moko sipsipsipwa lisa.*
        became.sick and he/she X   became
        He/she fell ill and became very weak.

(24)   *süng* 'wood' in:

    Papapmawo singa süngsüngsüngwa li.
    snail's   horn X   becomes
    The snail's tentacles stand up firmly.

(25)   *tük-* 'wipe' in:

    ʔÜngka khim tüktüktükwa sengnga.
    I   house X   I.cleaned
    I cleaned the house thoroughly.

(26)   *them-* 'be lost' in:

    ʔÜngka themthemthemwa konnga.
    I   X   I.walked.about
    I walked about completely lost.

2. The list of words in this group that is of triplets the nucleus of which is not found as a free form or as a monolexemic stem, is much longer:

(27)   *cap-* '?' in:

?Üng kopa   capcapcapwa bima   ?ümrü.
my   grandfather X   to.walk are.able
My grandfather can walk very well.

(28)   *cek-* '?' in:

?Ayü cekcekcekwa wa   tayang.
today X   rain comes.CONT
Today it keeps raining.

(29)   *cin-* '?' in:

Moko tukkaba cincincinwa bana.
that   patient X   arrived
The patient arrived in a very poor state of health.

(30)   *cip-* '?' in:

?Am tangda   cipcipcipwa ?a?wa ya?ang.
thy   on.head X   oil   is.CONT
Oil is dripping from your head.

(31)   *chak-* '?' in:

Mo khokmawo   ngalüng chakchakchakwa ya?ang.
that old.woman's face   X   is.CONT
That old woman's face is shining.

(32)   *cheng-* '?' in:

?O   tit   chengchengchengwa ya?ang.
this cloth X   is.CONT
This material is very transparent.

(33)   *chok-* '?' in:

?Üng tit   chokchokchokwa katyang.
my   cloth X   appears.CONT
Your cloth looks very clean.

(34)   *chom-* '?' in:

?Am tit   chomchomchomwa katyang.
thy   cloth X   appears.CONT
Your cloth looks very white.

(35)   *dap-* '?' in:

Chaci   dapdapdapwa loma yakmaci   li.
children X   to.tell to.be.NSG becomes
Children must be told what to do in a strict way.

(36)   *deng-* '?' in:

Moko   dengdengdengwa di   dhatyang?
he/she X   what beats.CONT
What is he beating so loudly?

(37)   *dop-* '?' in:

Buktanghutda   dopdopdopwa ya?ang.
in.cave.interior   X   is.CONT
It is very dark inside the cave.

(38)   *dük-* '?' in:

Mangpa dükdükdükwa dhuntungu.
shaman X   shivers.CONT
The shaman is shivering continously.

(39)   *gom-* '?' in:

Moko müna gomgomgomwa bi.
that   man   X   walks
The man walks in a sluggish way.

(40)   *ghe-* '?' in:

Mo bukkhundima   ghegheghewa biyang.
that pregnant.woman   X   walks.CONT
The pregnant woman walks with her stomach protruding.

(41)   *ghek-* '?' in:

?Ummak?a set   ghekghekghekwa ?i.
by.beer   dying   X   laughs
A drunkard laughs continuously.

(42)   *ghok-* '?' in:

Hongkudangka ghokghokghokwa ca?wa ta.
from.river   X   water comes
Abundant water comes from the river.

(43)   *hip-* '?' in:

    a. *Moko hiphiphipwa thung.*
        he/she X   coughs
        He/she coughs in an irritating way.

    b. *?Ungka?a khana hiphiphipwa mitna.*
        I.ERG   thou   X   I.remember.thee
        I will think of you all the time.

(44)   *hok-* '?' in:

Khana ca?wa hokhokhokwa manthokde!
thou   water X   do.not.pour
Don't pour the water in such a wasteful way!

(45)   *jük-* '?' in:

   Moko ma?a   dhiru kiya jükjükjükwa liyang.
   he/she by.fever met   and X   became.CONT
   He had an attack of fever and became very ill.

(46)   *jhong-* '?' in:

   M   himsale jhongjhongjhongwa lot.
   that lunatic X   runs
   The lunatic runs around wildly.

(47)   *kep-* '?' in:

   ?Ayü   kepkepkepwa nam nuyang.
   today X   sun is.good.CONT
   The sun is shining nicely today.

(48)   *kük-* '?' in:

   Moko hongku kükkükkükwa datyang.
   that river X   appears.CONT
   The river looks a deep blue.

(49)   *khek-* '?' in:

   Moko khekkhekkhekwa ?iyang.
   he/she X   laughs.CONT
   He/she keeps laughing.

(50)   *khuk-* '?' in:

   Moko buwale khukkhukkhukwa patyang.
   that owl   X   cries.CONT
   The owl keeps hooting.

(51)   *lak-* '?' in:

   Dena khana laklaklakwa tü?iyang?
   why thou X   laughest.CONT
   Why are you laughing all the time?

(52)   *lek-* '?' in:

   a. *Moko bhürü lekleklekwa dat.*
       that hill X   appears
       The hill can be seen clearly.

   b. *?Üngka?a moko lekleklekwa sintung.*
       I.ERG   that X   I.know.it
       I know that very well.

(53)   *leng-* '?' in:

   ?Am tit   lenglenglengwa datyang.
   thy cloth X   appears.CONT
   Your cloth seems quite transparent.

(54)   *luk-* '?' in:

    ʔÜng kümaʔa lukluklukwa dhuntang.
    I   by.fear X   I.trembled
    I was very much afraid.

(55)   *lum-* '?' in:

    Nosaʔa   süngtang lumlumlumwa khuyu.
    he/she.ERG tree   X   carried
    He/she carried the tree with difficulty.

(56)   *mek-* '?' in:

    a. *Mekmekmekwa ʔüng ʔipma yiyang.*
        X   my sleep comes.down.CONT
        I feel very sleepy.

    b. *Namda mekmekmekwa ʔan ʔipma yi.*
        in.sun X   our sleep comes.down
        We feel very sleepy in the sun.

(57)   *müng-* '?' in:

    Müngmüngmüngwa ʔummak manhade!
    X   beer   do.not.distribute
    Don't be to generous with the beer!

(58)   *nun-* '?' in:

    ʔAm chuk nunnunnunwa duttungyang.
    thy hand X   I.feel.CONT
    I feel the soft touch of your hand.

(59)   *ngak-* '?' in:

    ʔAm ngalüng ngakngakngakwa katyang.
    thy face   X   appears.CONT
    Your face looks very red.

(60)   *pe-* '?' in:

    Mo chaʔa ʔü nicha   pepepewa khuyu.
    hat child-ERG his/her younger.sibling X   carried
    The child carried its younger brother/sister with difficulty.

(61)   *phe-* '?' in:

    ʔÜngkaʔa lüngga phephephewa khuyung.
    I.ERG   thatch X   I.carried.it
    It was easy for me to carry the thatch.

(62)   *phom-* '?' in:

    Sumpuk phomphomphomwa bhungyang.
    dry.grass X   is.piled.up.CONT
    Dry grass is easily piled up.

(63)   *phuk-* '?' in:

       ʔO bungga phukphukphukwa nam.
       this flower X   smells
       The flower has a very pleasant smell.

(64)   *rang-* '?' in:

       Khana rangrangrangwa di   tükhunyang?
       thou   X   what carriest.CONT
       What is the huge thing you are carrying?

(65)   *rok-* '?' in:

       ʔAyü rokrokrokwa wa ta   yakyang.
       today X   rain comes is.CONT
       Today it has been raining all day.

(66)   *suk-* '?' in:

  a. *Kutiwaʔa suksuksukwa namsu.*
      dog.ERG X   smells
      The dog sniffs eagerly.

  b. *Mo ma   suksuksukwa khawa.*
      that woman X   wept
      The woman kept weeping.

(67)   *sük* '?' in:

       Khana dena süksüksükwa tüʔiyang?
       thou why X   laughest.CONT
       Why do you keep laughing to yourself.

(68)   *tak-* '?' in:

       ʔÜng maʔa   taktaktakwa münga.
       my   mother.ERG X   does.to.me
       My mother treats me very strictly.

(69)   *tam-* '?' in:

       Mosowo mükwa tamtamtamwa lisa.
       his/her   tears   X   became
       His/her tears were about to flow.

(70)   *tet-* '?' in:

       ʔOko cha tettettetwa cep.
       this   child X   speaks
       This child speaks very clearly.

(71)   *tong-* '?' in:

       ʔÜngka tongtongtongwa süng khunnga.
       I   X   wood I.carry
       I carry a heavy load of wood.

(72)   *thuk-* '?' in:

a.   *Saho?wa hukthukthukwa namyang.*
       leather   X   smells.CONT
       The leather smells very bad.

b.   *?Üngka cunga?a thukthukthukwa dhunnga.*
       I   by.cold X   I.tremble
       I shiver very badly from the cold.

(73)   *yap-* '?' in:

       ?Oko kodali yapyapyapwa lisa.
       this   hoe   X   became
       This hoe has worn very thin.

(74)   *yük-* '?' in:

       Yükyükyükwa tit   manthupde!
       X   cloth do.not.sew!
       Do not sew the cloth when you are trembling badly!

(75)   *yüng-* '?' in:

       Tit   pera kiya yüngyüngyüngwa lisa.
       cloth   tore and   X   became
       The material has become transparent by wear and tear.

3. The forms given as nuclei in §§1–2 are all characterised by finals that are admissible, in terms of Bantawa morphophonemics, in preconsonantal position. This is only natural as the nuclei occur only before consonants in the items discussed here, viz. in the interior of triplets or before *-wa*. The consonants found before other consonants (other than *-s-*) are *p, t, k, m, n* and *ng*, provided they follow immediately upon a vowel; in bases of the types CVCs- and CVCt-, the consonantal increments *-s-* and *-t-* are deleted before a further consonant. The list of consonants incurred in position before non-derivational consonants is thus shorter than that of consonants and consonant clusters admissible in prevocalic position for roots and extended bases: it is the prevocalic variants which provide the data that allow us to identify /s/, /n/, /l/, and /r/ as root-final morphophonemes of Bantawa – before consonants (other than *-s-*) /s/ is lost without a trace, and /n/, /l/, and /r/ merge to yield *-n-*.

It follows that the nuclei listed in §§ 1–2 cannot be equated directly with (unextended) roots or (extended) bases of the Bantawa lexicon: conclusive evidence will have to come rather from forms other than the triplets – only the combination of information about preconsonantal and prevocalic variants can lead to a safe identification of the morphophonemic shape of the configurations underlying the nuclei enumerated.

Thus unextended roots can be said to be reflected by, for example, the nuclei (4) *chop-*, (18) *nop-*, (5) *chük-*, (25) *tük-*, (26) *them-*, (22) *seng-*. On the other hand, for, for example, (2) *cük-*, (6) *dek-*, (7) *hak-*, (10) *kup-* and (19) *ngep-*, prevocalic variants make it necessary to derive the nuclei from the extended bases *cükt-*, *dekt-*, *hakt-*, *kupt-* and *ngept-*.

4. A few of the nuclei in §§ 1–2 are identical in shape with free forms of the Bantawa lexicon: (8) *hük* 'wind', (9) *kak* 'live coal', (17) *mük* 'eye', (20) *rak* 'heat' and (24) *süng* 'wood'. The remaining nuclei require an identification with verb stems, that is with bound forms. The fact that both *hük* and an extended verb stem *hükt-* would be reflected by a preconsonantal nucleus *hük-* opens the possibility to consider some, if not all, of the nuclei listed in this paragraph as derived from extended bases in -*t*-. There are two advantages to such a view: on the one hand, all triplets in Group One could be taken to consist of triplicated verb stem plus a deverbative suffix -*wa*; on the other, the rather awkward semantic distance between, say, *mük* 'eye' and (17) *mükmükmükwa* glossable by 'with many tears' would be lessened.

A suffix -*t*- added to a primary noun serves to shift it to the category of the verb (the function of Bantawa -*t*- can be described in most general terms as that of creating an opening for an additional argument or of increasing the valency of a form by one step); the basic meaning of *mükt-* would thus be 'do something with respect to the eye' (English *to eye* is not a bad parallel). The unreduplicated verb base *mükt-* 'to cast an eye upon' and the triplet *mükmükmükwa* can then be taken to show two different specifications of the use of the denominative verb base *mükt-*.

A similar argument can be brought to bear on *süngsüngsüngwa*, which can be translated roughly by 'standing erect'; again the semantic distance to the noun *süng* 'wood' is striking.

If one considers the fact that the older meaning of Kiranti **siŋ* appears to have been 'tree', a hypothetical **süngt-* would call for a hypothetical translation 'to tree', which would form the basis for an expression denoting 'stand upright like a tree'; this verb stem, with a meaning 'to remain motionless', is found in expressions such as *Manceppang süngtancin!* (NEG.speak.ABSTR X.REFL) 'Keep quiet!'.

The situation found in the case of *mükt-* has a further parallel in another form of our short list. If the argument put forward with regard to *mükmükmükwa* is appropriate, it should apply also to (8) *hükhükhükwa*, a form which one may want to render by 'with wind'. Again a monolexemic verb form exists alongside the noun *hük* 'wind'; again, as in the case of *mükt-* 'to cast an eye upon', it shows a specification of 'to do something with respect to wind' in that it denotes 'to produce wind; to fan'. Once more the unspecified notion expressed by the denominative in -*t*- is the one from which the meaning of the triplet can most readily be derived; it has to remain an open question whether this fact reflects a chronology of semantic changes.

5. Triplets included in Group Two all have in common that they cannot convincingly be aligned with monolexemic entities in the lexicon of present-day Bantawa. This means that the nuclei in §2 cannot be compared with variants outside the triplets, in particular not with prevocalic ones. As pointed out in §3, the nuclei in §2 cannot be derived from underlying morphophonemic configuration, at least not on the basis of synchronic Bantawa data. In the absence of a documentation of earlier stages of this language the only possibility left for a deeper analysis of the nuclei is a comparison with forms found in other Kiranti languages; this line of investigation will, however, not be pursued in the present paper.

6. If the approach just alluded to should prove successful, it would mean that part of the items in Group Two could be transferred to Group One, albeit on a pre-Bantawa level. This result would imply that a larger percentage of the forms attested in triplets could be related to items from the normal (pre-)Bantawa lexicon. It can, however, be claimed even at this point that in all likelihood there would remain a residue of items that would not find a place in such a normal lexicon. This claim can be based on the following observation:

If one considers roots of the shape CVC- as basic for the native Bantawa inventory of forms, and if one furthermore takes the prevocalic root initials to be representative of the components of the phonemic system, then this system can be said to contain the following items:

p	$\underline{t}$	t	c	k	$\textglotstop$
p^h	$\underline{t}^h$	t^h	c^h	k^h	
b	$\underline{d}$	d			
b^h	$\underline{d}^h$	d^h			
m		n		η	r l y w
	s		h		

Simple inspection of the list in §2 shows that six of the fifty items assembled here have an onset that is not included in the inventory of phonemes as represented in the basic lexicon: (39) gomgomgomwa, (40) ghegheghewa, (41) ghekghekghekwa, (42) ghokghokghokwa, (45) jükjükjükwa, and (46) jhongjhongjhongwa. The onsets, though not part of the system of basic Bantawa phonemes, supplement this system remarkably well; this fact may be taken to provide support against their being replaced by closely similar basic phonemes; further support may be assumed to derive from the fact that g, gh, j and jh do occur in the regular inventory of Nepali and thus in Bantawa loanwords taken over from this language.

The fact that g, gh, j and jh thus are not in any way outlandish to speakers of Bantawa still does not make forms that contain these sounds part of the basic lexicon of the language. It seems useful to separate such items from the normal inventory of lexemes by labelling them paralexemes.

7. Paralexemes then would be items whose phonological shape differs from forms in the basic lexicon. In our case, jhongjhongjhongwa and so on then have to be identified as paralexemes or as derivatives based on paralexemic nuclei. (It may be mentioned in passing that the application of the same criterion would also lead to the identification of fairly numerous borrowed items as paralexemic entities – a result which offers considerable advantages for further analysis.)

Paralexemic status has to be ascribed to a number of entities incurred in word lists assembled for present-day Bantawa on the strength of the phonological shape. Cases in question are kagak 'crow', ghangghangma 'spider', and jharak 'all', as well as gAjihangpa 'bride(groom)'s father' and gAjihangma 'bride(groom)'s mother', with the latter two forms also containing the vowel -A-, a mid-central vowel alien to the basic lexicon of Bantawa, though of high frequency in words borrowed from Nepali. (Again, there appears to be no indication that the mid-central vowel was ever replaced by either the high-central or the low-

central vowel, both of which form part of the phonemic system of the basic lexicon of Bantawa, and again it can be pointed out that mid-central *A* filled a gap in this system:

	front	central	back
high	*i*	*ü*	*u*
mid	*e*		*o*
low		*a*	

Thus, *gAji-*, though showing in sequence three sounds alien to the basic phonemic system of Bantawa, apparently still had a good chance to survive as such since the three sounds all could serve as symmetry-furthering complements of the system.)

Ghangghangma and *gAjihangpa/gajihangma* give evidence of the fact that paralexemes may be subject to word-formation processes normally affecting the basic lexicon: *ghangghangma* shows an extension by *-ma*, a suffix not uncommon in terms denoting small animals; *gAjihangpa/gAjihangma* are best analysed as compounds with *hang* as their second lexeme, an honorific term which as a free form is applied to 'king'. The dividing line between paralexemes and lexemes is thus rather tenuous: items from outside the lexicon proper can be used to extend and enrich it.

Still, it is worth noting that parts of the basic lexicon are highly resistant to such additions. Among the close to 600 verb stems listed in Novel Kishore Rai's dissertation,[4] only three have an onset that would lead one to consider including them among paralexemes: *jhomt-* 'attack', *jhus-* 'instigate' and *jhüs-* 'show one's teeth' – for 99.5% of all verb stems the initial consonant is part of the basic inventory of Bantawa phonemes.

8. An onset *g-* is found also in *gwamm-*, used to refer to the sound made when jumping. The extrasystemic initial consonant occurs in an item which shows in addition a violation of present-day Bantawa phonotactic rules: nowhere in the basic vocabulary does one incur words with initial consonant clusters. However, *gwamm-*, with its double deviation from the normal pattern, itself forms part of a set of some twenty configurations characterised by onsets C_1C_2a- followed by a sequence -C_3C_3-; only the continuants *-l-*, *-y-* and *-w-* can fill the position -C_2-. For -C_3-, *-y-*, *-p-*, *-t-*, *-k-*, *-m-* and *-ng-* are attested. Thus, even if an item such as *pwakk-*, referring to the sound made when hitting something, contains only sounds that can be taken to represent basic phonemes of Bantawa, the phonotactics of the form make it advisable to include *pwakk-* (as well as the other members of the set) in the category of paralexemes.

9. An item such as *gonom* 'stupid' shows, apart from its *g-*, another violation of normal root and base structure: rather than CVC- or CVC*t-*/CVC*s-* we find here an unreducible pattern CVCVC. The pattern recurs in, for example, *kagak* 'crow' and *jharak* 'all', that is in other items from the class of paralexemes identifiable as such by *-g-* and *jh-*. All three of these forms of the type CVCVC are further characterised by identity of the two vowels. The formula for the pattern can thus be rewritten to read CV_2CV_2C. This pattern too, recurs in

---

4   The additional forms introduced in §§7–11 are taken from Novel Kishore Rai's dissertation (*A descriptive study of Bantawa*, Pune 1985), a revised version of which is currently in preparation.

forms with sounds from the inventory of basic phonemes of Bantawa such as *bethem* 'idiot', *bhorok* 'frog', and *puyup* 'cucumber'; again, as in the case of words discussed in §7, derivationally or compositionally extended forms with a first element with stem-internal vowel harmony are found in *papapma* 'snail', *therengpa* 'thunder' and *chürütbung* 'a kind of orchid'.

10.   The combined evidence of extrasystemic sounds, of consonant clusters in initial position, and of vowel harmony within stems thus enables us to identify a substantial body of forms as belonging to a category which has here been called paralexemes. This phonological and phonotactic evidence can be supplemented by evidence from morphology such as has been presented in the main body of the present paper.

Triplication has been shown to occur with nuclei which in part could be assigned to the class of paralexemic entities on phonological grounds (cf. §6). If one wants to consider triplets an internally cohesive class of forms, then at least most of the items in Group Two should be included among the paralexemes, if not for the nuclei as such, then at least as far as the triplets are concerned. The same classification would seem to be in order for the triplets of Group One; here we would then have a recurrence of the phenomenon described in §§7 and 9, albeit with a difference in direction: forms may be transferred from one component of the opposition 'lexemes – paralexemes' to the other, which makes it possible that a lexemic nucleus may be expanded into a paralexemic triplet.

11.   Up to this point, our argumentation has been entirely in terms of form. It seems, however, possible to arrive at useful generalisations about semantic properties of paralexemes of native origin in Bantawa.

The verbal adjunct discussed in §§1–2 all serve to specify the manner of the verbal action alluded to; in extreme cases the verb itself is semantically near-empty. The specification usually refers to repetition or intensity of the action. As repetition is one means of reinforcing an action, repetition may be taken to be basic in this context. If so, the use of triplets may be taken to be highly iconic. (It should be noted here that the triple nucleus cannot be reduced to a double one, so that while KKK is acceptable, KK is not, and not even K can be used by itself without a serious difference in connotation.)

Iconicity of a slightly different type is used in forms of the type discussed in §8. Again the manner in which an action is performed is focused upon; now it is essentially the sound which accompanies the action that is of interest. One may want to interpret variations such as *pwakk-* (associated with breaking): *bhwakk-* (associated with falling down) or *twakk-* (hitting): *thwakk-* (breaking: *dhwakk-* (hitting) as pointing to traces of a system of consonant gradation used to signal degrees of intensity, and one may view the doubling of C_3 to C_3C_3 as signalling intensity as such and thus not very different functionally from triplication. However, it hardly seems safe to venture beyond a more general statement that in forms of this group, too, the connotative value of an expression seems to be more significant than the denotative one.

It is this general claim that may open the way toward an inclusion of forms of the type CV_1CV_1C. To claim that *bethem* or *gonom* or *jharak* are 'onomatopoeic' formations does not do much good: what is it that is imitated by these forms? What is, however, much more

easily claimed – and probably much more readily accepted as a claim – is that the use of paralexemic labels seems to evoke something apart from the mere identification of an entity in a linguistically grasped universe: persons, animals, things, actions and states so named are not just perceived as recognisable phenomena, but are at the same time appraised and evaluated subjectively. Viewed this way, items as different in form as members of the Bantawa classes KKK, $C_1C_2VC_3C_3$- and CV_1CV_1C (as well as of some classes not discussed here) can effectively be treated as belonging to one category – a category definable not only in terms of deviations from normal phonological, phonotactic, and morphological patterns of Bantawa, but also in terms of a shared functional-semantic dimension. Ideally, paralexemes should be characterised by both formal deviations and what may somewhat loosely be called functional overload. As can be seen clearly from the forms listed in §2, the degree of formal deviation may be reduced without a form becoming deprived of its status as a paralexeme: most nuclei in Group Two contain only sequences of sounds which recur in the basic vocabulary, and yet (67) *süksüksükwa* and (41) *ghekghekghekwa* apparently do not differ at all on the level of meaning and functional overload. To be sure, the pattern KKK keeps *süksüksükwa* safely inside the category of paralexemes, but if one were to decide that not only the triplets KKK should be recognised as paralexemes, but also the nuclei K, then *sük-* would not show any deviation from normal patterns of Bantawa root structure, and still one would probably be very reluctant to separate *sük-* from *ghek-*.

We have to conclude then that in spite of formal non-deviation, functional considerations may make us decide in favour of inclusion in the category of paralexemic entities. There would then be a cline from maximal deviation to non-deviation, with the domain of the paralexeme (as determined on functional grounds) extending well into the area of non-deviation. If so, we would be in a position to align Group One with Group Two – the fact that the nuclei in Group One are all non-deviating, while in Group Two at least some normal nuclei are clearly deviating, would not force us to consider KKK formations in Group One as categorically different from KKK formations in Group Two.

The claim made here is then that there is a gradual transition from the extrasystemic to the intrasystemic within the category of paralexemes, which makes it possible for entities from the category of lexemes to become incorporated in the functionally defined category of paralexemes.

The same lack of an insurmountable borderline between lexemes and paralexemes is found clearly in cases of a transfer in the opposite direction. Consider forms of the type CV_1CV_1C in §9 and §11: we have little difficulty in accepting the notion of a functional overload for *gonom* 'stupid', *bethem* 'idiot', or *jharak* 'all', but to maintain the same for *kagak* 'crow', *bhorok* 'frog', or *puyup* 'cucumber' is not that simple.

It seems necessary to draw a very general conclusion: as lexemes and paralexemes are coexisting parts of the lexical inventory of the speakers of a language, they will be in a position to transfer items from the one domain to the other if so desired. Such transfers may be occasional, or they may become conventionalised in the language: in the second case, lexicalisation may lead to elimination of a form from the category of paralexemes, paralexicalisation may have the opposite result. *Kagak* 'crow' then once was a (functionally defined) paralexeme in Bantawa, but no longer is, while some of the nuclei in Group Two may well have been part of the domain of lexemes in pre-Bantawa but now have ceased to belong there.

12. The investigation of what has been called here paralexemes in Bantawa has shown that we have before us a rich array of forms with their own formal characteristics, which serve to add an important functional dimension to the linguistic assessment of the speaker's universe and which provide a large reservoir for qualitative and quantitative changes in the lexicon of the language, changes which in part can be reconstructed, in part directly observed. Change, however, works both ways: paralexemes may become part of the normal lexicon, and the normal lexicon may feed the domain of the paralexemes.

What has been noted here about developments in Bantawa is both language-specific and of more general interest. The domain of the paralexemic coexists with that of the lexemic in other languages too; both domains will interact there too, because both of them are at the disposal of the competent speaker. The extent to which, say, the paralexemic is activated may well vary, but a language which could not make use of special expressions with a functional overload would lack an important ingredient for human beings' coping with the world around them, and for interacting successfully with others. What at first may seem to be marginal to language thus turns out to be a highly significant component, and to investigate it is well worth the effort of linguists.

THE POSSESSIVE OF EXPERIENCE IN BELHARE

BALTHASAR BICKEL

1. INTRODUCTION[1]

Bodily, emotional and cognitive states are likely to constitute a particular form class in linguistic coding. Well-known instances of this are dative subjects ('experiencer subjects') in South Asia (Masica 1976, Gupta & Tuladhar 1979, Verma & Mohanan 1990) or uncontrolled states/events ('experiential clauses') in Papuan languages (Reesink 1983, Foley 1986). What is common in such construction types is that the experiencing person (the one who feels or thinks) is marked as an oblique relation (in form of a case or verbal role marker). At the same time, however, the experiencer has subject properties to a certain degree. Depending on the language, the experiencer is treated like a regular subject (of intransitives and of active transitives) in cross-clausal coreference, verb agreement, reflexivisation, nominalisation, relativisation, and so on. The following examples are from Nepali[2] (Indo-Aryan; Wallace 1985:137) and Amele (Madang-Adalbert Range Stock; Roberts 1987:300), respectively. In both cases the experiencer, which is encoded as dative desinence in Nepali and as an undergoer marking affix in Amele, is accessible for cross-clausal same-subject marking:

(1) a. *Din bhari dāurā kāṭ-era   ma-lāī   tirkha   lāg-cha.*
       day full wood cut-SEQ.SS 1SG-DAT thirst strike-3SG.NPT
       After cutting wood the whole day, I am getting thirsty.

---

[1] I am indebted to the people of Belhārā and especially to Bimala Pa (Lekh Bahādur Rāī) for the generous hospitality and friendship which made all research possible. I also extend my thanks to William Foley and Sabine Stoll for most helpful discussions of an earlier draft and to Matthis Bickel, Eve Danziger, Sjors van Driem and Karen Ebert for help with the terminology and idiomaticity of experiences in English. All remaining flaws and mistakes are of course my own responsibility.

[2] Nepali is transliterated according to indological tradition except that, following van Driem (1987), mute *a* is not written even if it is not deleted by a *virām*. In Himālayan languages '*c*' and '*j*' represent alveolar (e.g. in Belhare) or (lamino-)postalveolar (e.g. in Limbu or Nepali) affricates.

---

b. *Ija   bim-ig   wen   te-i-a.*
1SG   come.up-1SG.SS   hunger   1SG.U-3SG.A-TODAY'S.PT
I came up and I became hungry.

Although dative subjects are typical for South Asia as a linguistic area, the Kiranti languages of eastern Nepal deviate from this pattern. Rather, these languages encode the experiencer as a possessor of the subject. The subject noun itself denotes an experience, or, more in line with Southeast Asian languages, the *"receptacle* or *arena"* (Matisoff 1986:8) where a physiological or psychological experience unfolds. The verb, which agrees with the noun, expresses the way in which the experiencer is affected. In some cases the verb is downgraded to an empty auxiliary. The pattern is illustrated by the following examples from Thulung (Western Kiranti; Allen 1975:99), Camling[3] (Central Kiranti), Limbu (Eastern Kiranti; van Driem 1987, s.v. *yu:ma?*) and Belhare (also Eastern Kiranti). The past inflection is used here for present states because the lexical *Aktionsart* of the verbs involved includes not only a stative but also an inchoative reading. Thus, a literal translation of, say, example (5) would be 'her mind became hurt', which potentially implies that it still hurts. On a stative interpretation, the translation would be 'her mind was hurt' without possible implications about the subject's present state.

(2)   *A-bhrem   liūra.*
1SG.POSS-laziness   affected
I am lazy.

(3)   *M-bulma   la-e.*
3SG.POSS-anger   AUX-PT
He is angry.

(4)   *A-lɛm   yu:s-ɛ.*
1SG.POSS-kidney   activated-PT
I don't have the courage.

(5)   *U-niũa   tug-he.*
3SG.POSS-mind   hurt-PT
She feels offended.

I call this construction 'possessive of experience' construction. It does not seem to be widespread in the languages of the world. The only well-documented[4] parallel I am aware of is found in some Papuan and Austronesian languages of New Guinea and Irian Jaya (see McElhanon 1975, 1977, 1992). As an alternative to the dative subject construction in example (1b), Amele has also possessive constructions like (6) (Roberts 1987:176). Example (7) is from Kâte (Finisterre-Huon Stock; McElhanon 1992:242) and (8) from Mangap Mbula (Austronesian; Bugenhagen 1990:202).

(6)   *Uqa   gema-g   be-i-a.*
3SG   liver-3SG.POSS   come.up-3SG-TODAY'S.PT
He became angry.

---

3   I am indebted to Karen H. Ebert, Zürich, for giving me access to her work in progress.
4   Masica (1976:164) mentions that Persian has "a few expressions of puzzling structure involving adjectives + suffixed possessive pronouns + 3SG 'be': e.g. *gorosn-am-e* = 'hungry-my-is' = 'I'm hungry'." In a survey on experiencers in South Asia, Hook (1990:329) cites an example from Shina (Dardic; Pakistan) as involving a "genitive-of-experience": *kitāp bodi par-ē mei gaar val-ar-ēgi* (book much read-ard my dizziness bring-CAUS-3SG.F.PT) 'I felt dizzy from reading the book (so) much'.

(7)   *Maŋ-ne*   *biaŋke-ka?*
       thoracic.cavity-3SG.POSS be.well-3SG.PRES
       She is happy.

(8)   *Ni-ŋ*   *i-saana.*
       body-1SG.POSS 3SG-go.bad
       I am exhausted.

In this paper I explore the possessive of experience in Belhare. In §2 I shall first define the construction and then (§3) discuss the lexical properties of the nouns and verbs involved. Section 4 is devoted to the syntactic properties of the possessive of experience. I shall demonstrate that the possessor in these constructions has subject properties. This is compared to the syntactic behaviour of the rare dative subject and unaccusative constructions that occur in Belhare (§5). Section 6 summarises the findings and puts them in a typological perspective.

The abbreviations used are:

A	actor	NOML	nominaliser
ADD	additive	NPT	non-past
CAUS	causal (case)	NS	non-singular
CIT	citation form	PERF	perfect
COM	comitative	POSS	possessive
CONTR	contrastive	PT	past
COP	copula	Q	interrogative
DIR	directive	REP	report marker
DISC	discovery	RES	resultative
DL	dual	S	subject of intransitives
EXC	exclusive	SEQ	sequential
F	feminine	SIM	simultaneous
ID	identifier	SS	same subject
IMPERS	impersonal	SUB	subjunctive
INC	inclusive	TEL	telic
INV	inverse	U	undergoer
IPFV	imperfective	(X)	eideme motivated by
M	masculine		meaning (category) X
MED	mediative	I, II, III, etc.	noun class labels

2. THE POSSESSIVE OF EXPERIENCE AND ITS DIAGNOSIS

The test that identifies a possessive of experience and separates it from ordinary possessives is found in a construction that renders 'as if' statements. In these constructions, the verb appears in the subjunctive past, marked by *-a* (as in example (9a)), or in the subjunctive non-past, signalled by zero (as in (9b)).

(9) a.   *Namniŋ   khar-a-ŋ-ha-e?wa*   *lui-?-ŋa.*
         last.year   go-SUB.PT-EXC-NOML-like   feel-NPT-EXC
         It seems to me as if I went last year. (but I am not sure any more)

    b.   *Cippa   nus-i-ha-e?wa*   *lu-het-na.*
         a.bit   cure-1PL-NOML-like   feel-IPFV-EXC
         I think it is getting better. (lit. I feel like one who is getting better.)

In example (9b), the person marker -*i*, literally a sign for first person plural inclusive,[5] has impersonal reference. It is a general rule in Belhare to use inclusive markers for impersonal reference. For instance, (10) can be used with reference to speaker and addressee as well as with non-specific reference.

(10)   *Lik-na   ka-ŋ-piu-ʔ-ni.*
        enter-CIT   INC.U-3NS.A-allow-NPT-NEG
        They don't allow us to enter. *or:* Entering is not allowed.

The impersonal use of the first person plural inclusive marker -*i* in example (9b) has been generalised with uncontrollable predicates such as weather expressions (11a) or, indeed, statements about bodily or emotional states with a possessive of experience (11b, c):

(11) a.   *Wet   ta-i-ha-eʔwa   cog-yu.*
        rain   rain-IMPERS-NOML-like   do-NPT
        It looks as if it is raining.

    b.   *A-hi   hond-i-ha-eʔwa   lu-het-na.*
        1SG.POSS-shit   appear-IMPERS-NOML-like   feel-IPFV-EXC
        It is as if I had to shit.

    c.   *A-niũa   tug-i-ha-eʔwa   lu-het-na.*
        1SG.POSS-mind   hurt-IMPERS-NOML-like   feel-IPFV-EXC
        It feels like I am sad.

The use of -*i* in (11) contradicts all agreement rules unless -*i* is synchronically analysed as a specialised marker for impersonal reference. The distribution of this marker serves as syntactic test for possessive of experience constructions: a possessor is in 'experience function' if and only if it is compatible with impersonal -*i* in the construction type exemplified by (11).

3. LEXICAL CHARACTERISTICS

The definition of the possessive of experience in the preceding section identifies the range of expressions collected in the following table. It is these terms that are used in possessive of experience constructions. In brackets I indicate the rough translation that would appear if a possessive of experience construction were to be put into English. For instance, *uniũa habhe*, literally 'his mind began to cry', translates idiomatically as 'he feels sorry, pity, sympathetic' or 'he is desperate (about all the things he has to do)'. Notice that I do not attempt here any full-fledged semantic analysis. A detailed inquiry into Belhare emotion semantics would presuppose anthropological analyses far beyond the scope of this chapter.

In the table some nouns are compatible with more than one verb. Along with, say, *aniũa tahe* (literally 'my mind became positively activated'), which denotes a state of happiness and joy, there is *aniũa tise* (literally 'my mind became pleased') which implies an object towards which the good feeling is directed, that is something that I like. I have classified the terms according to verbal valences. First, there are valence fields with body parts, or more accurately 'parts of a person'.[6] Next come fields that include body products (such as breath, sweat and excrements) and more and more purely experiential phenomena ('psychological'

---

5   See Bickel (1995) for a morpheme analysis of Belhare verb inflection.
6   Also from a cross-linguistic perspective, person is "a better candidate for the unique beginner in this domain" than body (Wilkins 1996:271).

experiences such as envy, detestation or fear as much as 'physiological' ones such as nausea, thirst or hunger).

TABLE: VERBS AND NOUNS IN THE BELHARE POSSESSIVE OF EXPERIENCE CONSTRUCTION

thotma 'stiff'	*laŋ* 'leg, foot', *muk* 'arm, hand', *sakmari* 'neck'
suma 'to itch'	*niūa* 'mind' ('fed up'), *laŋ* 'leg, foot', *muk* 'arm, hand', and other person parts
papma 'tangled'	
munma 'to forget'	*niūa* 'mind'
tima 'pleased'	
hapma 'to cry'	*niūa* 'mind' ('to feel sorry, pity, sympathetic, desperate')
rima 'to spin'	*niŋ* 'personality' ('dizzy, plastered'), *thona* 'decency,wisdom' ('plastered, disoriented, misbehaving')
ama 'to fall'	*laua* 'vital soul' ('sick and troubled' because of a shock, fear or anger)
khaama 'satiated'	
thomma 'distended'	*phok* 'belly'
yuma 'sleepy'	
yama 'to pity'	*mik* 'eye'
simma 'to tingle, paresthetic'	*laŋ* 'leg, foot', *muk* 'hand, arm' ('to be asleep' or 'to have pins and needles')
tonma 'aroused'	*li* 'penis', *si* 'vagina'
limma 'taste-sensitive'	*mun* 'mouth'
tukma 'to hurt'	*niūa* 'mind' ('offended, sad'), *sua* 'muscles (as a whole)' ('tired [of working, walking]'), *phok* 'belly', *khawa* 'wound', *mik* 'eye', *taŋghek* 'head', *yam* 'body' ('sick'), *nari* 'nose' (also 'feel like a cuckold'), and other person parts
tama 'activated; to come'	*niūa* 'mind' ('pleased, happy'), sua 'muscles (as a whole)' ('relaxation after effort'), *dʌsa* 'bad luck', *leppha* 'tongue' ('know to talk'), *sakma* 'breath' ('relief after danger')
pama 'to grow'	*sakma* 'breath' ('exhausted'), *nari* 'nose' ('fed up; stuck up')
honna 'perceptible, to appear'	*sakma* 'breath' ('relief after danger'), *hakliūa* 'sweat' ('hot') *khakmarak* 'mucus', *chepma* 'urine' ('need to urinate'), *hi* 'shit' ('need to shit'), *lalik* 'sperm', *phipma* 'fart', *mi?wa* 'tears', *khawa* 'wound', *chiat* 'spit' ('need to spit; despise'), *bʌl* 'power', *remsumik* 'envy', *yokma* 'embarrassment, fear, shameful joy'
katma 'activated; to come up'	*hikikpa* 'hiccup', *gauppa* 'burp', *hagamba* 'yawn', *hakchiūa* 'sneeze', *cipma* 'detest', *kipma ~ kitma* 'fear', *suma* 'weariness, tiredness', *pepma* 'nausea', *rek ~ ris* 'anger', *lamma* 'appetite', *salamma* 'appetite for meat', *iŋalamma* 'appetite for beer', *ŋoīsima* 'shame', *man* 'honour (of deities/ancestors)'
luma 'perceptible'	*waepma* 'thirst', *sak* 'hunger
pokma 'to rise'	*chom* 'desire'
uŋma 'to come down'	*retma* 'laughter' ('have to laugh')
lama 'to return'	*cik* 'detest' (archaic)

Left margin labels (bottom to top):
+ subcategorised for experience function
experiences
body products
parts of a person
+ regular nouns

Half of the nouns are ordinary lexemes that also appear outside the construction. This holds for most person-part terms including *niũa* 'mind' as in example (12), *thona* 'decency, wisdom' as in (13) and *niŋ* 'personality; name' as in (14).

(12) a. *Niũa-lo   mun dhup-ma khe-yu.*
     mind-COM talk talk-CIT must-NPT
     One has to make up one's mind.

   b. *U-niũa   ri-yu,   ri-sa-bu   cai-t-u: "emu-gari*
     3SG.POSS-mind turn-NPT turn-SS-REP eat-NPT-3U how-ABL

     omakg-e   tʌrkari-e   yum thikka aĩ-t-u?"...
     lentil.sauce-LOC vegetable-LOC salt correctly pour.in-NPT-3U
     He thought it over and over as he ate: "How does he put the right amount of salt into the *dāl* and the vegetable?"...

   c. *Un-chik-ŋaha niũa-cha   adhero-bu lis-e.*
     3-NS-GEN   mind-ADD   dark-REP be-PT
     And also their minds became grim.

   d. *N-niũa-e   emu mii-ka?*
     2SG.POSS-mind-LOC how think-NPT.2
     What do you think about that?

(13) a. *Na   u-thona   ŋ-wa-ni.*
     DEM 3SG.POSS-decency NEG-be-NEG
     This one has no sense of decency.

   b. *N-thona   yuŋŋ-ha-eʔwa   ceg-a   ai!*
     2SG.POSS-decency be-NOML-LIKE speak-IMP.SG EMPH
     Speak like a decent person!

(14) *A-niŋ   ser-he-m-ga!*
     1SG.POSS-personality kill-PT.3U-2PL.A-2
     You got on my nerves!

Among person-part terms only *yam* 'body' cannot be used outside the possessive of experience construction. It has become something of a negative polarity item and is specialised for a 'sick body'. Etymologically, *yam* seems to derive from a general term for 'body' as still attested in closely related Limbu (van Driem 1987). Most terms for body products occur outside the construction. Exceptions are *hagamba* 'yawn', *hikikpa* 'hiccup' and *gauppa* 'burp', but at least two of these (*hikikpa* and *gauppa*) are derived from ideophones (from *hikik* and *gauk* respectively). As for the experiential terms, it is only *waepma* 'thirst', *sak* 'hunger', *bʌl* 'power' and *lamma* 'appetite' (including its derivatives *salamma* and *iŋalamma* for specific kinds of appetite or hunger) that are used outside the possessive of experience construction, for instance in the following clause.

(15) *Waepma-a si-hai-ʔ-ŋa.*
     thirst-CAUS die-TEL-NPT-EXC
     I will die of thirst.

Taken together, the nouns in the table make up a 'category squish'. The property of being subcategorised for the experience function increases from top to bottom. At the top of the

table there are ordinary nouns. The terms at the bottom are specific experience terms, most of which are only used in possessive of experience constructions.

3.1 LEXICAL COMPOSITION

With nouns that are not strictly subcategorised for the experience function, the experiential meaning is clarified by the constructional meaning of the possessive of experience. In addition to this, the specific lexical structure signals experiential meaning. This works in several ways (see Matisoff 1986), which I propose to organise in terms of a major division between compositional and non-compositional collocations.

In non-compositional collocations, the verb, or 'psycho-mate' (Matisoff 1986), occurs with only one noun and is 'morphanic' in Matisoff's sense. This is the case for *tima* 'pleased', *munma* 'to forget', *khatma* 'satiated', *thomma* 'distended', *yuma* 'sleepy', *yama* 'to pity' and *limma* 'to have the sensation of taste, to be taste-sensitive'. The noun does not add information that is not already contained in the verb. The noun stands, however, in a relation to a full-fledged lexical item. The nouns used in example (16) also occur with the lexical meaning 'mind' (*niũa*), 'stomach, belly' (*phok*), 'eye' (*mik*) or 'mouth' (*mun*).

(16) a. *Na   u-niũa   muī-kha   ma²i.*
DEM   3SG.POSS-mind forget-NPT.NOML person
He is a forgetful person.

b. *A-phok   khas-e.*
1SG.POSS-stomach satiated-PT
I have had enough.

c. *A-mik   yus-e.*
1SG.POSS-eye sleepy-PT
I am sleepy.

d. *A-mik   ya-yu.*
1SG.POSS-eye pity-PT
I feel pity (so I can't kill).

e. *M-mun   lim-yu   i?*
2SG.POSS-mouth taste.sensitive-NPT Q
Can you taste anything? (to somebody who is having a cold)

I have called such signs 'eidemes' (Bickel 1995). They are semantically empty but potentially related to full morphemes. On an alternative analysis (suggested by Reh (1993) for Lwo languages in eastern Africa), the nouns are analysed as case markers for indicating an experiencer role. This analysis is not convincing for Belhare since the number of nouns in collocations like (16) is not much lower than the number of collocations itself. This would be case allomorphy beyond a reasonable degree. Moreover, non-compositional collocations are rather rare in Belhare. In contrast to the pattern in (16), the following collocation types are compositional. The types are distinguished mainly by whether it is the noun (type i.a and i.b) or the verb (type ii.a and ii.b) that has a general, rather than a specifically experiential meaning. A third type (iii) is represented by figurative collocations built on metaphors and metonymies.

Type (i.a) In one pattern there is a general noun and a specialised 'psycho-mate', that is a verb subcategorised for a possessive of experience. In these collocations, the semantic contribution of the noun is to localise the experience. The experience itself is denoted by the verb. This is the case with *tukma* 'to hurt', *thotma* 'stiff', *tonma* 'aroused, horny', *suma* 'to itch and *simma* 'to tingle, to have the sensation of paresthesia' (as one is 'having pins and needles' in the leg or arm), for instance:

(17) a. *Unchi-sua-bu*   *tuk-khar-e.*
          3NS.POSS-muscles-REP hurt-TEL-PT
          They became tired (of walking).

      b. *A-phok*   *tug-he.*
          1SG.POSS-belly hurt-PT
          My stomach aches.

      c. *A-laŋ*   *thot-khar-e.*
          1SG.POSS-leg stiff-TEL-PT
          My leg muscles became stiff.

      d. *N-li*   *toī-yu*   *i?*
          2SG.POSS-penis aroused-NPT Q
          Do you get an erection?

      e. *A-laŋ*   *sims-e.*
          1SG.POSS-leg paresthetic-PT
          My leg is asleep.

      f. *A-niũa*   *su-yu.*
          1SG.POSS-mind itch-NPT
          I am fed up.

Another example of this type is the experience verb *luma* 'perceptible'. However, this verb combines with nouns that are lexically experiential, albeit not strictly subcategorised for an experience function. This is the case with *waepma* 'thirst' and *sak* 'hunger', for instance:

(18)   *Ŋ-waepma*   *lus-e*   *i?*
        2SG.POSS-thirst perceptible-PT Q
        Are you thirsty?

Collocations with general nouns and specialised experience verbs seem to be the common collocation type also in Southeast Asian languages as discussed by Matisoff (1986).

Type (i.b) The next type is similar to the preceding one except that the verb's experience meaning is part of a polysemy structure. As a consequence, the noun helps construct the experience function insofar as it disambiguates the verb. The verbs *tama* and *katma*, for instance, are systematically ambiguous between an experiential meaning 'activated in a positive way, perceptible' and a motion meaning 'to come (from an unknown place)' and 'to come up', respectively. The meanings are disambiguated by the property of motion verbs to require a special allomorph of the imperfective marker in the non-past. With *tama* and *katma* in the experiential reading, the imperfective is formed by the regular suffix *-het,* just as with any other verb. Like all experiential verbs, the aspect of *katma* (as well as of *tama*) is ambiguous between inchoative and stative (see example (19a)). This contrasts with the case when the verbs are taken in their motion meaning. In this use they behave morphologically like any other motion verb and require the special suffix *-ket* to mark imperfective aspect in

the non-past. In agreement with other motion verbs, there is no ambiguity in the aspectual interpretation of the form, as in (19b).

(19) a. *U-salamma*   *katd-het.*
     3SG.POSS-appetite.for.meat activated-IPFV
     He is (getting) hungry for the meat.

   b. *Kat-ket.*
     come.up-IPFV
     She is coming up.

Moreover, the experiental meaning of *tama* in (20a) and (20b) is independently established by its use in (21a) and (21b), where the verb means 'activated, enabled'.

(20) a. *A-niũa*   *ta-he.*
     1SG.POSS-mind activated-PT
     I am pleased.

   b. *A-sua*   *ta-he.*
     1SG.POSS-muscles activated-PT
     I felt relaxation (after effort).

(21) a. *Riŋ*   *n-tai-ʔ-ni.*
     language NEG-activated-NPT-NEG
     He cannot speak. (of a mute person)

   b. *Nepali-cha ta-yu.*
     N.-ADD   activated-NPT
     He also speaks Nepali.

The verb *honma* 'to appear; perceptible' has a similar polysemy structure and often refers to the mere perceptibility of a body product:

(22) a. *A-hakliũa*   *hond-he.*
     1SG.POSS-sweat perceptible-PT
     I am hot.

   b. *A-sakma*   *hond-he.*
     1SG.POSS-breath perceptible-PT
     I felt relief (after danger).

   c. *A-chepma*   *hond-he.*
     1SG.POSS-urine perceptible-PT
     I have to urinate.

In other cases, however, the meaning deviates from this pattern and *honma* is to be taken in its more common meaning 'to appear'. The following examples imply that the body product is already visible.

(23) a. *A-lalik*   *hond-he.*
     1SG.POSS-sperm appear-PT
     I came (sexually).

   b. *U-miʔwa*   *hond-he.*
     3SG.POSS-tear appear-PT
     She cries.

c. *Hab-i-cha   mi?wa   ŋ-hoĩ- ?-ni.*
    cry-1PL-ADD   tear   NEG-appear-NPT-NEG
    Even when we cry, there are no tears.

Notice that, in contrast to *katma* and *tama,* the experience meaning of *honma* cannot be predicted by the possessive of experience construction. The pattern in both (22) and (23) is of possessive of experience constructions, yet *honma* means 'perceptible' in (22) and 'to appear' in (23).

Type (ii.a)   As noted above, some nouns are strictly subcategorised for the experience function. In these cases, the verbal 'psycho-mate' has a general meaning. This is the mirror-image effect of what we observed with general nouns and specialised verbs, that is in the collocation type (i). The verbs *pokma* 'to rise', *honma* 'to appear' and *lama* 'to return', are used with nouns having intrinsically an experience meaning, for instance:

(24) a. *Khat-ma a-chom   pog-yu.*
         go-CIT   1SG.POSS-desire rise-NPT
         I would like to go.

   b. *U-remsumik   hond-he.*
       3SG.POSS-envy appear-PT
       He is jealous.

   c. *U-yokma   hoĩ-yu.*
       3SG.POSS-embarrassment appear-PT
       He's embarrassed. (said of a child who has found a big amount of money and
       is embarrassed about it; can also be said of somebody being intimidated)

The noun *laua* 'vital soul' in the expression *ulaua arhe,* literally 'his soul fell', is a border-line case. On one hand, the noun is virtually restricted to this collocation and therefore seems to designate directly the experience of the (both psychological and physiological) sickness one is believed to have after a shock (due to a physical accident or due to sudden fear or anger). On the other hand, in Belhare psychology *laua* is often talked of as a full-fledged person part,[7] which would rather suggest that the experience meaning is generated by metonymy (cf. the lexical composition type (iii) discussed below). On such an account, the experience of sickness is referred to by the affected person part.

Type (ii.b)   Experiential nouns also collocate with verbs that have themselves a strictly experiential meaning. With respect to the semantic characteristics of the noun this is a subtype of the preceding collocation structure. It is observed with *katma*, which denotes the mere presence or activation of an experience (if it is not used as a motion verb; see above). In an expression such as in example (25), then, the semantic contribution of the verb is very low.

(25)   *A-suma   kar-he.*
       1SG.POSS-weariness activated-PT
       I am tired out.

This last type is well represented in Camling, where most possessive of experience constructions involve a specialised experience noun and a semantically empty auxiliary as

---

7   See Hardman (n.d.) for an elaborate discussion of similar psychological categories among the
     Lohorung, a linguistically closely related Kiranti group. On the concept of falling (*ama*) involved here,
     cf. Bickel (1996).

psycho-mate. In example (3) above (*mbulma lae* 'he is angry') *la-* merely serves to host inflection. Outside such collocations, *la-* functions as auxiliary with loan-words that are integrated into Camling by derivational means, for example *mil-ba la-* (agree-INTEGRATOR AUX-) from Nepali *milnu* 'to agree'.

Type (iii)   In the preceding types the experiential meaning of the possessive of experience constructions follows at least partly from the experiential meaning of one of its constituents, viz. of the verb (type i), of the noun (type ii.a) or of both (type ii.b). This is different from the following phenomenon, which, following a suggestion by McElhanon (1975), can be called 'idiom generation'. Here, the experience meaning is an effect only of the possessive of experience construction itself. It is not pre-established by lexical semantics. Examples involve the verbs *hapma* 'to cry', *pama* 'to grow', *papma* 'tangled', *tukma* 'to hurt', *rima* 'to spin' and *honma* 'to appear':

(26) a. *U-niũa*   *hab-yu.*
         3SG.POSS-mind   cry-NPT
         He is desperate.

   b. *U-nari*   *pas-e.*
         3SG.POSS-nose   grow-PT
         She became stuck up.

   c. *A-niũa*   *pap-khar-e.*
         1SG.POSS-mind   tangled-TEL-PT
         My thoughts became tangled.

   d. *Han-chi-nari*   *n-tuu-ʔ-ni*   *i?*
         2-DL.POSS-nose   NEG-hurt-NPT-NEG   Q
         Don't you feel like cuckolds?

   e. *N-niŋ*   *ri-yu.*
         2SG.POSS-personality   spin-NPT
         You will be drunk.

   f. *U-chiat*   *kolo*   *hoĩ-yu.*
         3SG.POSS-spit   CONTR   appear-NPT
         But he despises it.

In none of these instances are there independent grounds on which the verb could be assigned a specific experience meaning. The verb *tukma* 'to hurt' does have a general experience meaning (see example (17b) above), but in (26d) it is part of a metaphorical idiom and is not to be taken in its literal meaning. Notice that idiom generation is not restricted to metaphorical derivations as in (26a) to (26e). In some cases, the idiom is generated by a metonymic shift from the physical phenomenon that frequently goes with a certain emotion to the emotion itself. This is exemplified by (26f).

This figurative type of collocation is not so common in Belhare, nor indeed in Kiranti languages in general. Also Limbu seems to rely only occasionally on idiom generation. An example was given in (4). Together with (27), it seems to exhaust the possibilities (van Driem 1987, s.v. *sɔːmmaʔ* and *tiŋmaʔ*, respectively).

(27) a. *Ku-na-ːn*   *sɔːr-ɛ.*
         3SG.POSS-face-ABS   ooze.down-PT
         He is frowning.

b. *Kɛ-dhɛʔl   sɔːr-ɛ.*
2SG.POSS-spit ooze.down-PT
You're drooling.

c. *Kɛ-lɛː-n   tiŋd-ɛ-iː?*
2SG.POSS-penis-ABS flower-PT-Q
Do you have an erection?

In contrast to this, possessive of experience construction in Papuan and Austronesian languages involve mainly constructional metaphors and metonymies, based on complex culture-specific ways of creating idioms (McElhanon 1975, 1977).

3.2 A NOTE ON SEMANTIC FLUCTUATION BETWEEN VERB AND NOUN

Whether the terms in the table denote parts of a person, body products or experiences, they are all treated the same way and are compatible with the possessive of experience construction. Not surprisingly then, there is some historical fluctuation between these notional sub-domains. We have already observed *yam*, originally 'body', which has become a negative polarity item restricted to a sick body. The noun does not designate the experience itself since the 'mate' is a specialised experience verb, *tukma* 'to hurt'. Other person-part terms are reanalysed as part of experience nouns. An example is *remsumik* 'envy' which derives from *remsu* 'envy' and *mik* 'eye'. The first part is still attested in the word for the 'other wife of one's father' (in polygamy), *remsuma*, where *-ma* is related to the common teknonymic marker for women, and in a collocation with *hitma* 'to look':

(28)   *Remsu ka-ŋ-hiʔ-yu.*
envy   INC.U-3NS.A-look-NPT
They envy us.

A look at related neighbouring languages shows that such fluctuation is not uncommon also in linguistic history. In closely related Limbu, for instance, the idiomaticity of 'feeling sleepy' is distributed differently from Belhare. In Belhare the construction is non-compositional. The subject noun (*mik*) is a term eidemically related to a body part (*mik* 'eye') and the verb is a fully specialised experience term *yuma* 'sleepy', as in example (29a). In the compositional Limbu expression (29b), there is a general verb *yuːmaʔ* 'to experience' and an experience noun (*miʔ*) denoting 'sleepiness' (van Driem 1987; s.v. *yuːmaʔ*). Limbu *miʔ* is still related to the body part *mik* 'eye' by way of paronymy.

(29) a. Belhare: *A-mik   yus-e.*
1SG.POSS-eye sleepy-PT
I am sleepy.

b. Limbu: *A-miʔ   yuːs-ɛ.*
1SG.POSS-sleepiness activated-PT
I am sleepy.

The Limbu body-part term *mik* itself appears again with polysemous *yaːmaʔ* 'tickled, feel tickled, ticklish; horrified' (van Driem 1987, see *mik yaːmaʔ*) in example (30).

(30)   *Ku-mik   yaːs-ɛ.*
3SG.POSS-eye horrified-PT
She was horrified to see it.

Here, the noun is assigned an experience function both by the verbal mate and the constructional meaning of the possessive of experience.

Whereas in Limbu it is the noun that has become specialised, in Belhare it is the verb. This is corroborated by the fact that in Belhare *yuma* still has another meaning than 'sleepy', which is slightly closer to its more general Limbu counterpart *yu:ma?* 'to experience'. In collocations like (31a) and generally in negation (31b), the verb means 'to experience sleep', which can mean to be asleep (31a) or to be able to sleep (31b).

(31) a. *A-mik   yu-ma   yur-he.*
        1SG.POSS-eye experience.sleep-CIT enough-PT
        I have slept enough.

     b. *A-mik-to   n-yu-at-ni.*
        1SG.POSS-eye-ID   NEG-experience.sleep-PT-NEG
        I couldn't sleep.

4. POSSESSORS AS SUBJECTS

The preceding examples suggest that the possessed noun phrase as a whole is the subject of the clause. This is evidenced by third person singular agreement (marked by zero) in all instances. In other respects, however, it is the possessor in the experiential noun phrase that functions as subject. I use the term subject here in a standard definition, which assumes a clustering of subject properties. A noun phrase has subject properties if it is syntactically treated like the single argument of an intransitive clause and the actor of a transitive clause. In the following I shall explore the subject properties of the possessive of experience.

4.1 SAME-SUBJECT CONSTRUCTIONS

The clause linkage suffix *-sa* requires that the subject of the subsequent clause is the same as the current one (cf. Bickel 1993 and 1995). In example (32) it is not possible to have the storyteller as the one who is freed from lice.

(32)   *Kʌtha mas-sa un-na un sik   set-pi-yakt-he.*
        story tell-SS 3-ERG 3 louse kill-BEN-IPFV-PT-3U
        She was delousing her when telling a story.

In clauses linked by *-sa,* the actor cannot be resumed by a subsequent undergoer, as in (33a), but only by an actor, as in (32), or the single argument of an intransitive verb, as in (33b, c).

(33) a. **Un-na   chis-sa   yeti lui-t-u-ga?*
        3-ERG   meet-SS   what tell-NPT-3U-2
        What will you say to him when he finds you?

     b. *Min-cek-pa   piī-sa   ap-khai?-ŋe.*
        NEG-say-LOC   run-SS   come.across-TEL-RES
        He has came over here running and without telling anybody.

     c. *Hap-sa   hap-sa   a-nucha   ta-he.*
        weep-SS   weep-SS 1SG.POSS-younger.sibling come-PT
        My younger brother came crying.

This notion of co-reference is also satisfied by a possessor of experience as in example (34).

(34) *Hap-sa   hap-sa   a-niūa   tug-he.*
      weep-SS weep-SS 1SG.POSS-mind hurt-PT
      I was sad and cried.

Inverting the sequence does not alter this finding:

(35) a. *U-hakchiūa   kas-sa   khim rak-lamma   hond-he.*
        3SG.POSS-cough come.up-SS house interior-MED appear-PT
        She came out of the house coughing.

   b. *A-sak   lu-sa   imm-har-e-ŋa.*
      1SG.POSS-hunger feel-SS sleep-TEL-PT-EXC
      Though hungry, I fell asleep.

   c. *A-ppa   la-har-e,   uŋ, u-ris   kas-sa?*
      1SG.POSS-father return-TEL-PT Q   3SG.POSS-anger come.up-SS
      My father went back angrily, didn't he?

Also in Camling a possessive of experience can be monitored by same-subject marking.

(36) *I-homa   i-homa   pa-dhit-aci-nʌ*
      one-mana   one-mana INV-find-DL-SEQ

      Kic-sikha   la-sa   pa-tat-aci-ko   raicha.
      3NS.POSS-joy AUX-SS INV-bring-DL-NOML DISC
      They found one mana here and one mana there and happily brought them
      (home).

The same goes for the Papuan language Amele (Roberts 1987:300):

(37) *Ija   tataw-ig   ija   am-i   wal-do-i-a.*
      1SG SIM.stand-1SG.SS 1SG eye-1SG.POSS spin-3SG-3SG-TODAY'S.PT
      As I stood I became dizzy.

4.2 TRANSITIVE VERB AGREEMENT

It goes without saying that, in transitive clauses, a possessor does usually not trigger verb agreement. A transitive verb agrees with the macro-roles (cf. Foley & Van Valin 1984) of actor ('A') and undergoer ('U'), which are defined as the most agentive and the least agentive arguments, respectively. In transitive verb agreement the possessive of experience shows interesting behaviour. Some experience constructions, such as *niūa tima* 'happy', *kitma katma* 'afraid' and *cipma katma* 'to detest', allow transitivisation. The semantic effect is that the source of the experience is referred to by specific determination:

(38) a. *Cia a-niūa   ti-yu.*
        tea 1SG.POSS-mind pleased-NPT
        I like tea.

   b. *Cia a-niūa   tiu-t-u-ŋ.*
      tea 1SG.POSS-mind pleased-NPT-3U-1SG.A
      I like this tea.

The syntactic effect is that it is the possessor that agrees with the verbal actor affix. This is in parallel with a regular transitive paraphrase *kiiʔtuŋ*:

(39)   *A-kipma*   *kai?-t-u-ŋ.*   cf. *kii?-t-u-ŋ*
        1SG.POSS-fear bring.up-NPT-3U-1SG.A   fear-NPT-3U-1SG.A
        I am afraid of him.   I fear him

The undergoer is a regular undergoer and not the possessed noun. This is evidenced by number agreement in examples like (40a). The possessed noun is also not a secondary object. In example (40b) the slot for the secondary object, abbreviated as 'SO', is filled by another argument, *ekchuma* 'sash'. It follows that the possessor, realised by the possessive prefix *a-* in (40a) and the proclitic *hani* in (40b), functions as actor.

```
            U             A
            |             |
```

(40) a. *Na   ma?i-chi   saro   a-niūa*
        DEM human-NS   very   1SG.POSS-mind

```
            U  A
            /\/\
```

        n-tiu- ?-ni-ŋ-chi-ŋ.
        NEG-pleased-NPT-NEG.3U-1SG.A-NS.U-1SG.A
        I don't like these people very much.

```
        b.  U              SO        A
            |              |         |
```

        Ŋka   hale   ekchuma   hani-niūa
        1SG   before   sash   2PL.POSS-mind

```
            U         A
            |         /\
```

        ka-tiu-s-i-k-kha.
        1SG.U-pleased-PERF-2PL-2-PERF
        Before, you have liked me for my sash.

Possessor agreement in transitives seems to occur also in Limbu, as the following example (van Driem 1987, s.v. *luŋma*) suggests.

(41)   *An-dzum-in*   *sa:?rik a-luŋma*   *hi:pt-u-ŋ.*
        1SG.POSS-friend-ABS   very   1SG.POSS-liver yearn-3U-1SG.A
        I miss my friend very much.

The intransitive counterpart of (41) is (42) (s.v. *luŋma himma?*).

(42)   *Sa:?rik a-luŋma*   *him.*
        very   1SG.POSS-liver yearn
        I am extremely overwhelmed by the grief of separation, by nostalgia.

There are, however, also other examples suggesting that in Limbu this subject property is not as pervasive as in Belhare (van Driem 1987, s.v. *niŋwa tama?*).

(43)   *A-niŋwa*   *kɛ-das-u.*
        1SG.POSS-mind 2-reach.upward-3U
        You please me.

Constructions like (43), where the experience noun phrase functions as undergoer, do not seem to occur in Belhare.

4.3 RECIPROCAL FORMATION AND NOMINALISATION

Two other constructions in which the possessive of experience has subject properties are reciprocal formation and nominalisation.

Reciprocal formation involves two actors who are in an identical relation to the predicate. An example is (44a). The same construction can be formed from possessive of experience expressions. For this to be possible the possessor must be treated like an actor, as in (44b).

(44) a. *Mit-ka-mit   ca-i!*
          think-REC-think AUX-1PL
          Let's think of each other!

     b. *Un-chik niūa   ti-ga-ti   n-cai-chi.*
          3-NS   mind pleased-REC-pleased 3NS-AUX-DU
          They love each other.

Like most Kiranti languages, Belhare has a means to derive nominals referring to the actor argument ('A') of transitive (example (45a)) or to the single argument ('S') of intransitive verbs (example (45b)).

(45) a. *cama ka-thuk-pa*
          food NOML.S/A-cook-M
          a cook

     b. *Ka-lik-pa.*
          NOML.S/A-enter-M
          The one who went into (the house).

The derivative *ka-* is not sensitive to the specific semantic role that the subject of intransitives plays. Unlike a *nōmen agentis* form, *ka-* is not restricted to agentive subjects but derives nominals referring to any type of subject, including also themes (46).

(46) a. *Hene   ka-yuŋ-ma-ga?*
          where NOML.S/A-be-F-2
          Where are you (woman) from? (lit. Where (are) you a female dweller from?)

     c. *Ka-pikg-a-ba.*
          NOML.S/A-fall-downwards-M
          The one who fell down.

Thus, the form involves the notion of subject in the strict sense, which combines the semantically general single argument of intransitives with the actor of transitives. It is only undergoer arguments of transitive verbs that are excluded: *kathukpa* in (45a) cannot refer to the food being cooked. This notion of subject is also satisified by possessors of experience. As example (47) illustrates, they can be made the referent of the *ka*-form:

(47) *Unchi-ris   ka-kat-pa-chi-ŋa   n-seŋs-e.*
       3NS.POSS-anger NOML.S/A-come.up-M-NS-ERG 3NS.A-tear-PT.3U
       Those who got angry (about the results) tore down (the announcement).

Also in Limbu verbal nouns can be formed from possessive of experience constructions. The copula *co:kma?* in the following example (van Driem 1987:66; s.v. *tukma?*) indicates a subjective or temporary ascription:

(48)   *A-laŋ   kɛ-duk-pa   co:k.*
        1SG.POSS-leg NOML-hurt-M COP
        There is something wrong with my leg.

5. ALTERNATIVES: UNDERGOERS OF EXPERIENCE AND UNACCUSATIVES

The diagnostic used to identify the possessive of experience also finds resonance in some expressions that look like dative subject constructions. The small set of such constructions in example (49), which I refer to as 'undergoer of experience' constructions, takes the generalised impersonal agreement marker *-i* just as a possessive of experience does (example (50)). The constructions are for the most part in syntactic parallelism with Nepali, since dative noun phrases in Nepali regularly correspond to undergoer affixes in Belhare. The only difference is the ergative case marking on the stimulus. The verb *setma*, which appears in example (49d) in the forms *kasei?ni* 'it doesn't make us drunk' and *kase?yu* 'it makes us drunk', derives from *setma* 'to kill' but behaves syntactically differently from its etymon.

(49) a. *Khalamba-ŋa mai-lab-he.*   cf. Nep. *Ma-lāī   rughā lāg-yo.*
        cold-ERG   1SG.U-catch-PT   1SG-DAT cold   strike-3SG.PT
        I have got a cold.

     b. *Cuŋ-ŋa   mai-tar-he.*   cf. Nep. *Ma-lāī   jvaro ā-yo.*
        fever-ERG 1SG.U-bring-PT   1SG-DAT fever come-3SG.PT
        I have got fever.

     c. *A-chepma-a   mai-lett-he.*
        1SG.POSS-urine-ERG 1SG.U-urge-PT
        I had to urinate.

     d. *Iŋa-ŋa   ka-sei-?-ni,*   *rʌksi-a*
        beer-ERG INC.U-make.drunk-NPT-NEG liquor-ERG

        ka-se?-yu.
        INC.U-make.drunk-NPT
        You don't get drunk from beer but from liquor.

(50)   *Cuŋ-ŋa   (mai-)tar-i-ha-e?wa   lu-het-na.*
        fever-ERG (1SG.U-)bring-IMPERS-NOML-like feel-IPFV-EXC
        I feel as if I have got fever.

In the construction type (50) it is possible, although not very common, to inflect the verb for the undergoer. This confirms the observation in §2 that *-i* is indeed reanalysed as an impersonal marker.

Like possessives of experience, the undergoers of experience in example (49) have subject properties to a certain extent. Verbal nouns with subject reference are derived without problems (see (51)). As was shown in the preceding section, regular undergoers cannot be taken as the referent of *ka*-derivations.

(51) a. *cuŋ-ŋa   ka-tat-pa*
fever-ERG   NOML.S/A-bring-M
one who has fever

   b. *chepma-a   ka-let-pa*
urine-ERG   NOML.S/A-urge-M
one who has to urinate

   c. *Ka-set-pa-chi   he-lleŋ   ŋ-khar-e?*
NOML.S/A-make.drunk-M-NS   where-DIR   3NS-go-PT
Where did the drunkards go?

Same-subject constructions are more restricted. They are grammatical (52a) but the same-subject form in *-sa* optionally inflects for the undergoer (52b, c). This results in a kind of anticipatory reference marking. Such marking is not possible if the undergoer is not an experiencer but in another semantic role. A regular transitive verb like *luma* 'to tell' has either a same-subject affix in *-sa* or it is inflected for person, number and role. The two paradigms are in complementary distribution.

(52) a. *Cuŋ-ŋa   tas-sa-ro   Dhankuta tas-e-ŋ.*
fever-ERG   bring-SS-ID   Dhankutā reach-PT.3U-1SG.A
I reached Dhankutā although I indeed had this fever.

   b. *Cuŋ-ŋa   mai-tas-sa   Dhankuta tas-e-ŋ.*
fever-ERG   1SG.U-bring-SS   Dhankutā reach-PT.3U-1SG.A

   cf. **mai-lu-sa*
1SG.U-tell-SS
telling me
Although I had fever I reached Dhankutā.

   c. *A-chepma-a   mai-les-sa   yuŋ-he-ŋa.*
1SG.POSS-urine-ERG   1SG.U-urge-SS   sit-PT-EXC
I sat having to urinate.

The anticipatory inflection on same-subject forms suggests that Belhare undergoers have less subject properties than possessors of experience. Obviously the latter have enough referential prominence to be unambiguously monitored as subjects in clause linkage. They do not invite anticipatory inflection as a clarifying device.

Another alternative to possessive of experience constructions is unaccusative verbs. These verbs denote experiences such as *luma* 'perceptible, felt', *tama* 'activated, enabled', *khanma* 'nice, beautiful, good', *suma* 'sour', *khikma* 'bitter', *limma* 'delicious, tasty' or *kuma* 'warm, hot'. Although intransitively inflected, they are subcategorised both for the experiencer and the experienced thing (the 'stimulus'). Both arguments qualify as absolutives (marked by zero). Therefore, clauses may be ambiguous. Topicalisation, although more often associated with the experiencer than with the stimulus, does not preclude ambiguity:

(53)   *Na(-na)   khikt-he-bu.*
DEM(-TOP)   bitter-PT-REP
This one is bitter. *or*: To this one it tastes bitter.

From the point of view of verb agreement, it is the stimulus noun phrase that counts as subject:

(54)   Ŋka   hakliũa   lus-e.
        1SG   sweat   perceptible-PT
        I am hot.

Yet in other respects the experiencer argument has subject properties. In cross-clausal same-subject marking it is the experiencer that is treated as subject:

(55)   Hakliũa   lu-sa   thaŋŋ-har-e-ŋa.
        sweat   perceptible-SS   go.up-TEL-PT-EXC
        I started to climb up in the heat.

Unlike undergoers of experience the experiencer argument of unaccusatives is referentially prominent enough to be treated as subject. Anticipatory inflection does not appear.

In verbal noun derivation, the experiencer argument is treated like actors and the single argument of ordinary intransitives (cf. examples in (45) and (46) above), that is like a subject:

(56)   Na   (caleppa) ka-khik-pa   m-pi-n-an-u-m!
        DEM bread   NOML.S/A-bitter-M   NEG-give-NEG-IMP.PL-3U-2PL.A
        Don't give any more to this one to whom (the bread) tastes bitter!

This syntactic association of experiencers with the subject is in line with the claim in Role and Reference Grammar that experiencers are likely to group with the actor macrorole, whereas stimulus or theme nominals cluster with undergoer roles (Van Valin 1993:44).

6. SUMMARY AND CONCLUSIONS

The possessive of experience construction in Belhare involves nouns denoting parts of a person, body products and experiences. They are ordered along a continuum of decreasing subcategorisation for an 'experience' function. If the collocation is lexically compositional, the experience meaning results from five different patterns: (type i.a) general nouns with specialised monosemous experience verbs or with (type i.b) polysemous experience verbs, (type ii.a) specialised experience nouns with general or with (type ii.b) experience verbs, and (type iii) idiom generation. Idiom generation is remarkably rare in the Kirant when compared to similar constructions in Papuan and Austronesian languages, where 'body-image expressions' prevail (see McElhanon 1977, Bugenhagen 1990). In Belhare, and probably in the Kirant in general, it is more common to express experiences by specific experiential terms. This is even the more interesting because syntactically, experiential terms like remsumik 'envy' or lamma 'appetite' are treated like ordinary person-part terms such as taŋghek 'head' or niũa 'mind' and body-product expressions such as hakliũa 'sweat' or hi 'shit'. From this one could have expected experiences to be coded more in the figurative manner of idiom generation. Historically, however, it seems that constructional metaphors and metonymies were more common. Evidence for this are experiential verbs like tama 'activated in a positive way' and honma 'perceptible' that are still closely related to general verbs for coming and appearing.

Except for intransitive agreement, the possessive of experience functions as a subject. Experiencers belong to what Bally (1926) called "la sphère personelle", which gives them a high degree of referential prominence or empathy. The linguistic problem of experiencers is to code them as such but to grant them at the same time the grammatical prominence a true person deserves. This often gives rise to subject properties clustering with oblique

experiencers. This is well testified by dative subject constructions. It can also be observed with what is called possessor promotion or raising in the literature (see, for instance, Bally 1926 and Seiler 1983). Two examples of this often discussed phenomenon may suffice here. In both cases, one (example (57)) from Vedic (Seiler 1983:42) and one (example (58)) from Yimas (Lower Sepik; Foley 1991:301), the possessor of an experiencing body part, that is the ultimate experiencer, surfaces as a dative argument.

(57)   *yó   mē   kukṣí   sutásomaḥ   pṛṇáti*
    REL 1SG.DAT belly.ACC having.pressed.the.soma fills
    who has pressed the soma and fills (with it) my belly

(58)   Ŋarwa   *wa-ŋa-kwalca-t.*
    penis.IX.SG IX.SG.S-1SG.DAT-arise-PERF
    I have an erection.

In apparently rare cases, experiencing possessors are also promoted to subject. This is known from the Bantu language Haya (Hyman 1977, Seiler 1983:46). To this gallery of promoted experiencing possessors, we may now add the subject properties of the Belhare possessive of experience. It is yet another instance of the widespread and well-known propensity of human language to give a privileged status to the 'sphère personelle'.

REFERENCES

Allen, Nicholas J., 1975, *Sketch of Thulung grammar*. Cornell University East Asia Papers No.6. Ithaca, N.Y.: Cornell University.

Bally, Charles, 1926, L'expression des idées de sphère personelle et de solidarité dans les langues indo-européennes. In F. Frankhauser and J. Jud, eds *Festschrift Louis Gauchat*. Aarau: Sauerländer.

Bickel, Balthasar, 1993, Belhare subordination and the theory of topic. In Karen H. Ebert, ed. *Studies in clause linkage*. Arbeiten des Seminars für Allgemeine Sprachwissenschaft No.12. Zürich: ASAS-Verlag.

    1995, In the vestibule of meaning: transitivity inversion as a morphological phenomenon. *Studies in Language* 19:73-127.

    1996, Spatial operations in deixis, cognition, and culture: where to orient oneself in Belhare. In Jan Nuyts and Eric Pederson, eds *Language and conceptualization*, 46-83. Cambridge: Cambridge University Press.

Bugenhagen, Robert D., 1990, Experiential constructions in Mangap-Mbula. *Australian Journal of Linguistics* 10/2:183-215.

Driem, George van, 1987, *A grammar of Limbu*. Berlin: Mouton.

Foley, William A., 1986, *The Papuan languages of New Guinea*. Cambridge: Cambridge University Press.

    1991, *The Yimas language of New Guinea*. Stanford: Stanford University Press.

Foley, William A. and Robert D. Van Valin, Jr., 1984, *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Gupta, Sagar Mal and Jyoti Tuladhar, 1979, Dative subject constructions in Hindi, Nepali and Marathi and Relational Grammar. *Contributions to Nepalese Studies* 7:119-153.

Hardman, Charlotte E., n.d., Conformity and self-expression: a study of the Lohorung Rai of East Nepal. PhD dissertation, School of Oriental and African Studies, London.

Hook, Peter Edwin, 1990, Experiencers in South Asian languages: a gallery. In Verma and Mohanan, eds.

Hyman, Larry M., 1977, The syntax of body parts. In Ernest Rugwa Byarushengo, Alessandro Duranti and Larry M. Hyman, eds *Haya grammatical structure: phonology, grammar, discourse,* 99-117. Southern California Occasional Papers in Linguistics No.6. Los Angeles: University of Southern California.

Masica, Colin P., 1976, *Defining a linguistic area: South Asia.* Chicago: Chicago University Press.

Matisoff, James A., 1986, Hearts and minds in South-East Asian languages and English: an essay in the comparative lexical semantics of psycho-collocations. *Cahiers de linguistique asie-orientale* 15:5-57.

McElhanon, K.A., 1975, Idiomaticity in a Papuan (non-Austronesian) language. *Kivung* 8:103-44.

1977, Body image expressions in Irianese and Papua New Guinean languages. *Irian* 6:3-27.

1992, On the concept of person. In Tom Dutton, Malcolm Ross and Darrell Tryon, eds *The language game: papers in memory of Donald C. Laycock. PL,* C-110.

Reesink, Ger P., 1983, Switch reference and topicality hierarchies. *Studies in Language* 7:215-246.

Reh, Mechthild, 1993, Zur Versprachlichung mentaler Erfahrung in Lwo-Sprachen. Paper presented at the workshop 'From body to emotion and cognition', Universität zu Köln, 2–3 July 1993.

Roberts, John R., 1987, *Amele.* London: Croom Helm.

Seiler, Hansjakob, 1983, *Possession as an operational dimension of language.* Tübingen: Narr.

Van Valin, Robert D., Jr., 1993, A synopsis of Role and Reference Grammar. In: Robert D. Van Valin, Jr., ed. *Advances in role and reference grammar.* Amsterdam: Benjamins.

Verma, Manindra K. and Karavannur Puthanvettil Mohanan, eds, 1990, *Experiencer subjects in South Asian languages.* Stanford: CSLI.

Wallace, William D., 1985, Subjects and subjecthood in Nepali: an analysis of Nepali clause structure and its challenges to Relational Grammar and Government & Binding. PhD dissertation, University of Illinois, Urbana-Champaign.

Wilkins, David, 1996, Natural tendencies of semantic change and the search for cognates. In Mark Durie and Malcolm Ross, eds *The comparative method reviewed.* New York: Oxford University Press.

A NEW ANALYSIS OF THE LIMBU VERB

GEORGE VAN DRIEM

1. THE LIMBU VERB

Limbu is a Kiranti language native to eastern Nepal and the western fringe of Sikkim. The author devoted the fourth chapter of his 1987 grammar of the Phedãppe dialect of Limbu to a morphological analysis of the Limbu simplex verb (van Driem 1987). Simplicia are non-periphrastic finite indicative verb forms, from which the various Limbu periphrastic tense forms, gerunds, participles, adhortative and optative forms are derived. Limbu distinguishes eleven pronominal categories: first, second and third person, singular, dual and plural number, and an inclusive–exclusive distinction in the first person dual and plural. The Limbu transitive verb shows agreement with both *agent* (transitive subject) and *patient* (transitive object or beneficiary), and the transitive paradigm distinguishes 44 different forms, as shown in Tables 2 and 3. Limbu intransitive and reflexive verbs agree with the *subject* (intransitive or reflexive subject), and the intransitive and reflexive paradigm distinguishes eleven different forms, as shown in Table 1.

After I had completed the manuscript for the Limbu grammar in the summer of 1986, Professor Emeritus Carl Ebeling of the University of Amsterdam and Caucasologist Rieks Smeets of Leiden University suggested that alternative analyses of Limbu conjugational morphology were possible, some of which might require positing fewer slots or functional positions. Since I was eager to start my work on the Dumi language in the late summer of 1986, I did not give the matter high priority at the time. In subsequent comparative work on conjugational morphology in Kiranti and other Tibeto-Burman languages, I proposed morphological analyses for the conjugations of other languages, none of which, as it turned out, presumed as many suffixal slots as my Limbu analysis. Even the morphological analysis of the older Bahing paradigm, the transitive conjugation of which distinguished as many as 64 of the 75 theoretically possible forms, presumed fewer slots than the Limbu analysis. As time went on, my dissatisfaction with the 1987 analysis grew, and I began to assign the students of my Limbu course at Leiden University the task of revising my 1987 morphological analysis, an exercise which proved both instructive and entertaining. Recently, Ebert (1991, 1992) presented papers, which included diagrams of morphemic analyses of Limbu, Chamling and Athpare simplicia. The complete Chamling and Athpare data have not been made available, but her Limbu diagram represents an insufficient analysis. Because of the relevance of Limbu conjugational morphology to the comparative study of Tibeto-Burman verbal flexion, therefore, I felt it was high time to present a new morphological analysis of verbal agreement in the Limbu simplex.

157

TABLE 1: LIMBU AFFIRMATIVE AND NEGATIVE INTRANSITIVE
AND REFLEXIVE SIMPLEX CONJUGATIONS IN THE
NON-PRETERITE AND PRETERITE TENSES

(In each box the negative form is listed below the corresponding affirmative form.)

		INTRANSITIVE			REFLEXIVE	
		NPT	PT		NPT	PT
	1SG	Σ-$^?\varepsilon$ mɛ-Σ-$^?$ɛn	Σ-aŋ mɛ-Σ-aŋnɛn	1SG	Σ-siŋ$^?\varepsilon$ mɛ-Σ-siŋ$^?$ɛn	Σ-siŋaŋ mɛ-Σ-siŋaŋnɛn
	1DU.INC	a-Σ-si an-Σ-sin	a-Σ-ɛtchi an-Σ-ɛtchin	1DL.INC	a-Σ-nɛtchi an-Σ-nɛtchin	a-Σ-nɛtchi an-Σ-nɛtchin
s	1DU.EXC	Σ-sige mɛ-Σ-sigɛn	Σ-ɛtchige mɛ-Σ-ɛtchigɛn	1DU.EXC	Σ-nɛtchige mɛ-Σ-nɛtchigɛn	Σ-nɛtchige mɛ-Σ-nɛtchigɛn
u	1PL.INC	a-Σ an-Σ-nɛn	a-Σ-ɛ an-Σ-ɛn	1PL.INC	a-Σ-siŋ an-Σ-siŋnɛn	a-Σ-siŋɛ an-Σ-siŋɛn
b	1PL.EXC	Σ-ige mɛ-Σ-igɛn	Σ-m$^?$na mɛn-Σ-m$^?$na	1PL.EXC	Σ-siŋige mɛ-Σ-siŋigɛn	Σ-siŋŋ$^?$na mɛn-Σ-siŋŋ$^?$na
j	2SG	kɛ-Σ kɛn-Σ-nɛn	kɛ-Σ-ɛ kɛn-Σ-ɛn	2SG	kɛ-Σ-siŋ kɛn-Σ-siŋnɛn	kɛ-Σ-siŋɛ kɛn-Σ-siŋɛn
e	2DU	kɛ-Σ-si kɛn-Σ-sin	kɛ-Σ-ɛtchi kɛn-Σ-ɛtchin	2DU	kɛ-Σ-nɛtchi kɛn-Σ-nɛtchin	kɛ-Σ-nɛtchi kɛn-Σ-nɛtchin
c	2PL	kɛ-Σ-i kɛn-Σ-in	kɛ-Σ-i kɛn-Σ-in	2PL	kɛ-Σ-siŋi kɛn-Σ-siŋin	kɛ-Σ-siŋi kɛn-Σ-siŋin
t	3SG	Σ mɛ-Σ-nɛn	Σ-ɛ mɛ-Σ-ɛn	3SG	Σ-siŋ mɛ-Σ-siŋnɛn	Σ-siŋɛ mɛ-Σ-siŋɛn
	3DU	Σ-si mɛ-Σ-sin	Σ-ɛtchi mɛ-Σ-ɛtchin	3DU	Σ-nɛtchi mɛ-Σ-nɛtchin	Σ-nɛtchi mɛ-Σ-nɛtchin
	3PL	mɛ-Σ mɛn-Σ-nɛn	mɛ-Σ-ɛ mɛn-Σ-ɛn	3PL	mɛ-Σ-siŋ mɛn-Σ-siŋnɛn	mɛ-Σ-siŋɛ mɛn-Σ-siŋɛn

In the following, I shall use the words 'old' and 'new' to refer respectively to the morphemes, morpheme labels and slots of the 1987 analysis and of the new analysis proposed here. New slots are proposed, and some old slots have been abolished. Certain morpheme labels have been made more precise. Zero morphemes have been reassessed, and the problem of negation is discussed. Implications of the new analysis for the diachronic view of conjugational morphology in Kiranti and in Tibeto-Burman are discussed.

Abbreviations used in this paper are:

1, 2, 3	first, second, third person	PL	plural
AG	agent of a transitive verb	PT	preterite
DU	dual	REF	reflexive
EXC	exclusive	S	subject of an intransitive or
INC	inclusive		reflexive verb
NDU	non-dual	SF	suffixal slot
NEG	negative	SG	singular
NPT	non-preterite	→	indicates the direction of a
NSG	non-singular or		transitive relationship
PAT	patient of a transitive verb	Σ	verb stem
PF	prefixal slot		

2. THE PREFIXAL CHAIN EXPANDED

A flaw in the old analysis is that the old prefixal slot PF1 could be occupied by two morphemes, viz. any combination of the first person prefix *a-* , the second person prefix *kɛ-* and an old third person zero morpheme Ø. Slots are functional positions in the affixal string of a verb, each of which can be occupied by a definable set of morphemes. The morphemes sharing a position in a string define the function of that position. There appears to be a general tendency for semantically related morphemes to occupy the same slot. Slots are language-specific and analysis-dependent and represent the non-random sequential ordering of morphemes in conjugated verb forms. To have more than one morpheme occupying a slot defeats the purpose of having slots in the first place and necessitates making statements about the relative position of two morphemes within a single slot, something which is not ascertainable in those cases in which one of these is a zero morph. Conversely, attempts at slot reduction may not be allowed to supersede the goal of formulating a maximally explanatory analysis.

The first person prefix *a-* always precedes the second person prefix *kɛ-* and therefore must be analysed as occupying an anterior slot. This prefix indicates first person in forms in which first person involvement is not indicated by some portmanteau. It now seems a bit overwrought to have posited a zero allomorph of this morpheme in forms containing the exclusive suffix *-ge ~ -be* (see Sprigg 1989). It is reasonable to say that the meaning of the exclusive suffix comprises the sense of first person involvement in addition to exclusion of the person or persons addressed. Michailovsky (1989:472) proposes to analyse the prefix *a-* as a first person non-singular inclusive morpheme despite its occurrence in 2→1 forms in which Michailovsky maintains that the opposition between inclusive and exclusive is effectively 'neutralised'. I cannot concur with this view, as the prefix clearly functions as a marker of first person, not only in 2→1 forms, but also in non-finite forms such as the supine.

The second person prefix *kɛ-* indexes second person in forms in which second person involvement is not indexed by some portmanteau. The old third person zero morpheme may be abolished by a more precise labelling of two other morphemes in old prefixal slot PF2: The old non-singular agent/subject morpheme *mɛ- ~ m-* specifically marks the involvement of a *third person* non-singular agent or subject, and should therefore be relabelled as (3NSG.AG.S). Its abbreviated allomorph *m-* occurs between a preceding prefix and the root

TABLE 2: LIMBU AFFIRMATIVE TRANSITIVE SIMPLEX CONJUGATION IN THE NON-PRETERITE AND PRETERITE TENSES

(In each box the preterite form is listed below the corresponding non-preterite form.)

p a t i e n t

agent	1SG	1DU.INC	1DU.EXC	1PL.INC	1PL.EXC	2SG	2DU	2PL	3SG	3D	3PL
1SG						Σ-nɛ / *Σ-nɛ*	Σ-nɛtchiŋ / *Σ-nɛtchiŋ*	Σ-niŋ / *Σ-niŋ*	Σ-uŋ / *Σ-uŋ*	Σ-uŋsiŋ / *Σ-uŋsiŋ*	
1DU.INC									a-Σ-su / *a-Σ-ɛtchu*	a-Σ-susi / *a-Σ-ɛtchusi*	
1DU.EXC						Σ-nɛtchige / *Σ-nɛtchige*			Σ-suge / *Σ-ɛtchuge*	Σ-susige / *Σ-ɛtchusige*	
1PL.INC									a-Σ-um / *a-Σ-um*	a-Σ-umsim / *a-Σ-umsim*	
1PL.EXC						Σ-nɛtchige / *Σ-nɛtchige*			Σ-umbe / *Σ-mʔna*	Σ-umsimbe / *Σ-mʔnasi*	
2SG	kɛ-Σ-ʔɛ / *kɛ-Σ-aŋ*								kɛ-Σ-u / *kɛ-Σ-u*	kɛ-Σ-usi / *kɛ-Σ-usi*	
2DU	agɛ-Σ		agɛ-Σ		agɛ-Σ				kɛ-Σ-su / *kɛ-Σ-ɛtchu*	kɛ-Σ-susi / *kɛ-Σ-ɛtchusi*	
2PL	*agɛ-Σ-ɛ*		*agɛ-Σ-ɛ*		*agɛ-Σ-ɛ*				kɛ-Σ-um / *kɛ-Σ-um*	kɛ-Σ-umsim / *kɛ-Σ-umsim*	
3SG	Σ-ʔɛ / *Σ-aŋ*	a-Σ-si / *a-Σ-ɛtchi*	Σ-sige / *Σ-ɛtchige*	a-Σ / *a-Σ-ɛ*	Σ-ige / *Σ-ige*	kɛ-Σ / *kɛ-Σ-ɛ*	kɛ-Σ-si / *kɛ-Σ-ɛtchi*	kɛ-Σ-i / *kɛ-Σ-i*	Σ-u / *Σ-u*	Σ-usi / *Σ-usi*	
3DU									Σ-su / *Σ-ɛtchu*	Σ-susi / *Σ-ɛtchusi*	
3PL	mɛ-Σ-ʔɛ / *mɛ-Σ-aŋ*	am-Σ-si / *am-Σ-ɛtchi*	mɛ-Σ-sige / *mɛ-Σ-ɛtchige*	am-Σ / *am-Σ-ɛ*	mɛ-Σ-igɛ / *mɛ-Σ-igɛ*	kɛm-Σ / *kɛm-Σ-ɛ*	kɛm-Σ-si / *kɛm-Σ-ɛtchi*	kɛm-Σ-i / *kɛm-Σ-i*	mɛ-Σ-u / *mɛ-Σ-u*	mɛ-Σ-usi / *mɛ-Σ-usi*	

of the verb, but the full form occurs between a preceding prefix and a following negative morpheme. The old singular agent/subject zero morph must be retained but more precisely labelled as the specific marker of a *third person* singular agent or subject (3SG.AG.S). These considerations necessitate positing more prefixal slots for person in the Limbu verb than in the old analysis, that is a first person slot (PF1), a second person slot (PF2) and a third person slot (PF3), replacing older slots PF1 and PF2.

LIMBU PERSON AND NUMBER AGREEMENT PREFIXES

PF1	PF2	PF3
		mɛ- ~ m-
		3NSG.AG.S
a-	*kɛ-*	3NSG.AG.S
1	2	
		∅
		3SG.AG.S

3. THE PROBLEM OF NEGATION

An analysis of negation in Limbu simplicia would have to account for the fact that the prefix *mɛn-* occurs in negative preterite forms with a first person singular or first person plural exclusive agent or subject and the lack of an overt negative suffix in such forms. In all other negative simplicia, there are at least two negative and no more than three overt negative affixes, of which one is a prefix.

ANALYSIS A		ANALYSIS B	ANALYSIS C
PF4	PF5	PF4	PF4
		mɛ- ~ n-	
		NEG	
mɛ-	*n-*		*mɛ- ~ n- ~ mɛn-*
NEG	NEG		NEG
		mɛn-	
		1SG.AG.S/1PL.EXC.AG.S/PT.NEG	

Analysis B posits a distinct negative morpheme *mɛn-* in forms with a first person singular or first person plural exclusive agent or subject, bearing the clumsy morpheme label (1SG.AG.S/1PL.EXC.AG.S/PT.NEG). The sequence *mɛn-* in such forms is homophonous with the morpheme sequence *mɛ- + n-* (3NSG.AG.S + NEG) in negative forms with a non-singular third person agent or subject, although this does not lead to homophony anywhere in the paradigm. Analysis A, on the other hand, would explain the sequence *mɛn-* in negative preterite forms with a first person singular or first person plural exclusive agent or subject as the co-occurrence of two negative morphemes in adjacent slots, viz. *mɛ- + n-* (NEG + NEG). An advantage of analysis A is that this co-occurrence could be seen as a motivated phenomenon, in keeping with the regularity consistently observed elsewhere throughout the simplex paradigm that all negative forms are marked by at least two negative morphemes. In all other forms, one negative prefix is consistently seen to occur with one or two negative suffixes. The occurrence of two prefixed negative morphemes in negative preterite forms with a first person singular or first person plural exclusive agent or subject can be interpreted as being attributable to the fact that the negative suffix is 'blocked' by the negative preterite

TABLE 3: LIMBU NEGATIVE TRANSITIVE SIMPLEX CONJUGATION IN THE NON-PRETERITE AND PRETERITE TENSES

(In each box the preterite form is listed below the corresponding non-preterite form.)

p a t i e n t

agent	1SG	1DU.INC	1DU.EXC	1PL.INC	1PL.EXC	2SG	2DU	2PL	3SG	3DU / 3PL
1SG						me-∑-nen me-∑-nen	me-∑-netchiŋnen me-∑-netchiŋnen	me-∑-niŋnen me-∑-niŋnen	me-∑-ʔen men-∑-baŋ	me-∑-enchin men-∑-baŋsiŋ
1DU.INC									an-∑-sun an-∑-etchun	an-∑-sunchin an-∑-etchunchin
1DU.EXC				me-∑-netchigen me-∑-netchigen					me-∑-sugen me-∑-etchugen	me-∑-susigen me-∑-etchusigen
1PL.INC									an-∑-umnen an-∑-umnen	an-∑-umsimnen an-∑-umsimnen
1PL.EXC				me-∑-netchigen me-∑-netchigen					me-∑-umben men-∑-mʔna	me-∑-umsimben men-∑-mʔnasi
2SG	ken-∑-ʔen ken-∑-aŋnen								ken-∑-un ken-∑-un	ken-∑-unchin ken-∑-unchin
2DU	agen-∑-nen agen-∑-en		agen-∑-nen agen-∑-en		agen-∑-nen agen-∑-en				ken-∑-sun ken-∑-etchun	ken-∑-sunchin ken-∑-etchunchin
2PL									ken-∑-umnen ken-∑-umnen	ken-∑-umsimnen ken-∑-umsimnen
3SG	me-∑-ʔen me-∑-aŋnen	an-∑-sin an-∑-etchin	me-∑-sigen me-∑-etchigen	an-∑-nen an-∑-en	me-∑-igen me-∑-igen	ken-∑-nen ken-∑-en	ken-∑-sin ken-∑-etchin	ken-∑-in ken-∑-in	me-∑-un me-∑-un	me-∑-unchin me-∑-unchin
3DU									me-∑-sun me-∑-etchun	me-∑-sunchin me-∑-etchunchin
3PL	men-∑-ʔen men-∑-aŋnen	amen-∑-sin amen-∑-etchin	men-∑-sigen men-∑-etchigen	amen-∑-nen amen-∑-en	men-∑-igen men-∑-igen	kemen-∑-nen kemen-∑-en	kemen-∑-sin kemen-∑-etchin	kemen-∑-in kemen-∑-in	men-∑-un men-∑-un	men-∑-unchin men-∑-unchin

first person singular morpheme *-paŋ* or the preterite first person plural exclusive agent/subject morpheme *-mʔna*, although it is not clear what factor could be responsible for blocking the slot of the negative suffix.

A question arising from analysis A is that of allomorphy. In the original analysis, the negative prefix is analysed as *mɛ-* ~ *n-*, whereby the allomorph *mɛ-* occurs whenever the negative prefix is the first overt morpheme in a verb, and the allomorph *n-* occurs when it is not. In analysis A, allomorphy could be dispensed with, but there seems to be no other reason for analysing the negative prefix in *mɛ-hip-ʔɛ-n* 'I shan't hit him' as being an altogether different morpheme from that in *kɛ-n-hipt-u-n* 'you won't hit him'. Likewise, it seems neither motivated nor plausible to assign the negative morpheme *mɛ-* of prefixal slot PF4, assumed in analysis A, to some slot anterior to our current PF1 which would be 'blocked' by the occurrence of one of the overt prefixes *a-* (1), *kɛ-* (2) or *mɛ-* ~ *m-* (3NSG.AG.S). Rather, it is simpler to assume two formally identical negative prefixes NEG$_1$ and NEG$_2$, exhibiting the same pattern of allomorphy *mɛ-* ~ *n-* described in the original analysis, although the second negative morpheme only actually occurs in the form of its allomorph *n-* because it invariably co-occurs with the first negative morpheme. The occurrence of the second negative prefix in prefixal slot PF5 is taken to be conditioned by the lack of a second negative morpheme in the suffixal string of the verb. This situation occurs only in negative preterite forms with a first person singular or first person plural exclusive agent or subject where the negative suffix is 'blocked' by the negative preterite first person singular suffix *-paŋ* or the preterite first person plural exclusive agent/subject suffix *-mʔna*.

These competing analyses cannot be meaningfully assessed without consideration of the suffixal negative morphemes. With the exception of preterite forms with a first person singular or first person plural exclusive agent or subject, simplicia are negated by simultaneous prefixation and suffixation of negative morphemes. The obligatory prefixed morpheme was termed the first negative morpheme, and the suffix the second negative morpheme. The second negative morpheme *-nɛn* ~ *-n* (NEG$_2$) is located in the last position in the suffixal string, which, in the new analysis, is suffixal slot SF10. Negated non-preterite 1SG→3NSG forms also obligatorily take a third negative morpheme, which is a suffix *-n* (NEG$_3$) located in new suffixal slot SF6, for example *mɛ-ni-ʔɛ-n-chi-n* 'I don't see them'. The third negative morpheme also occurs facultatively in 1DU.INC→3NSG, 2SG→3NSG, 2DU→3NSG, 3SG→3NSG and 3DU→3NSG forms. Because the negative prefixal sequence *mɛn-* contains two nasal segments and because this sequence occurs only in those negative simplicia in which no negative suffix occurs, viz. preterite forms with a first person singular or first person plural exclusive agent or subject, it appeared intuitively satisfying to assume, as I did in the original analysis, that a Limbu simplex is negated by at least two negative morphemes. However, an analysis of the sequence *mɛn-* as both allomorphs of the negative prefix *mɛ-* ~ *n-* co-occurring in succession within a single slot is unsatisfactory because two morphemes ought not to occur in a single slot. Alternatively, the sequence *mɛn-* may be analysed as a special allomorph of this negative prefix occurring in negative preterite forms with a first person singular or first person plural exclusive agent or subject. This is analysis C.

Finally, it would appear to be more expedient to abandon the unsophisticated idea of several negative morphemes in Limbu simplicia and to adopt the concept of a single 'discontinuous morpheme' for the negative in simplicia, which is how I analysed the Limbu active participle *kɛ-...-pa*, Limbu negative active participle *mɛn-...-mna* and Limbu negative

perfect gerund *mɛn-...-ʔeː*, although it somehow did not occur to me to analyse the negative in simplex verbs in these terms. I have recently adopted Hagège's term 'simulfix' for this phenomenon, in particular with respect to the non-singular second person actant number simulfix *-a...-ni* in Lohorung (van Driem, forthcoming). The older Kiranti active participial ending *khə-...-kɾa* reflected in the petrified Dumi forms *khənɨkpa* (*nɨ-nɨ* 'be good') 'good, comely, beautiful' and *khəyiːkpa* (*iː-nɨ* 'be bad') 'horrid, ugly, foul, bad' is also such a discontinuous morpheme or simulfix (van Driem 1993b). A simulfix consists of two or more elements which are affixed simultaneously at different locations in the affixal string. Hagège (1986:26) also considered the neologisms *diffixe, ambifixe* and *transfixe* for this phenomenon, which he described for the Austronesian language Palau spoken in the archipelago of the same name in the Pacific Ocean between 134°05'E and 134°45'E and between 6°55'N and 8°15'N, but he prefers and adopts the term *simulfixe*. The case for analysing negation in Limbu simplicia as a case of simulfixation is compelling. This analysis would entail positing a single polylocational negative morpheme consisting of minimally one and maximally three elements in a well-defined pattern of paradigmatic distribution. The labels (NEG$_1$), (NEG$_2$) and (NEG$_3$) will be retained to indicate the elements of the Limbu negative simulfix. The analysis proposed therefore superficially resembles the old analysis and analysis C above, although it is substantially different from either of these. The new analysis posits a single negative simulfix with one to three affixal manifestations, the occurrence of which is paradigmatically determined and one of which is always a prefix. The new model for the prefixal string of the Limbu simplex should therefore look like this:

THE LIMBU PREFIXAL STRING

PF1	PF2	PF3	PF4
		mɛ- ~ *m-*	
a-	*kɛ-*	3NSG.AG.S	*mɛ-* ~ *n-* ~ *mɛn-*
1	2		NEG$_1$
		Ø	
		3SG.AG.S	

4. THE REFLEXIVE AND ALLOMORPHY

The old reflexive morpheme *-siŋ* ~ *-nɛ* ~ *-n* was interpreted as having an allomorph *-nɛ* ~ *-n* in dual forms and an allomorph *-siŋ* in singular and plural forms. In reflexive forms, the dual reflexive allomorphs *-nɛ* ~ *-n* always occur as part of the sequence *-nɛtchi*, whereby the element *-tchi* is identified as the allomorph *-tchi* of the dual patient/subject morpheme *-si* ~ *-tchi* in suffixal slot SF3. The allomorph *-tchi* occurs regularly following the preterite morpheme *-ɛ*, the dual reflexive allomorph *-nɛ* ~ *-n* and the 1→2 portmanteau *-nɛ* ~ *-n*. The dual reflexive preterite sequence *-n-ɛ-tchi* is homophonous with the corresponding non-preterite sequence *-nɛ-Ø-tchi*. In fact, consistently throughout the paradigm, cases of preterite/non-preterite homophony are satisfactorily accounted for by adjacency of a vocalic morpheme to the preterite morpheme *-ɛ*, which elides. The dual reflexive allomorphs *-nɛ* ~ *-n* are homophonous with the 1→2 portmanteau *-nɛ* ~ *-n*, which occurs in the 1SG→2SG ending *-nɛ*, the 1SG→2DU ending *-nɛtchiŋ*, the 1SG→2PL ending *-niŋ* and the 1NSG→2 ending *-nɛtchige*. These endings also exhibit preterite/non-preterite homophony, and attempts

to segment sequences like the dual reflexive ending -*nɛtchi* as -*nɛt* + -*chi* have not proved to be illuminating.

There seems to be no semantic common denominator on the basis of which the dual reflexive allomorphs -*nɛ* ~ -*n* and the 1→2 portmanteau -*nɛ* ~ -*n* could be analysed as a single morpheme, and so they were analysed as distinct entities in the 1987 analysis. The reflexive allomorph -*nɛ* ~ -*n* is restricted to dual reflexive forms, and the portmanteau -*nɛ* ~ -*n* indexes the transitive relationship between a first person agent and a second person patient. Subsequent comparative study seems to have vindicated this synchronic analysis. Whereas the dual reflexive allomorphs -*nɛ* ~ -*n* appear to be cognate with the Tibeto-Burman root for 'two' *g-ni-s, also reflected in the Limbu regular numeral *nɛtchi* 'two' (cf. *sumsi* 'three', *lisi* 'four', *nasi* 'five', all with the non-singular or 'generalised dual' suffix -*si*) and the collective numeral *nɛpphu* 'two' (cf. *sumbhu* 'three', *libhu* 'four', with the collective suffix -*phu*), the 1→2 portmanteau -*nɛ* ~ -*n* has cognates in the Dumi 1SG→2 suffix -*n*, the Hayu 1SG→2 suffix -*no*, the Kulung 1SG→2 suffix -*an*, the Thulung 1SG→2 suffix -*nini*, the Bahing 1SG→2 suffix -*na* and the Jinghpaw 1→2 suffix -*nteʔ1* ~ -*teʔ1*, all pointing to the existence of some 1SG→2 proto-morpheme at the Proto Tibeto-Burman level, which has tentatively been reconstructed as *-*nya*.

<div align="center">

ANALYSIS D   ANALYSIS E

SF1   SF1

-*siŋ* ~ -*nɛ* ~ -*n*   -*siŋ*

REF   NDU.REF

-*nɛ* ~ -*n*   -*nɛ* ~ -*n*

1→2   DU.REF

                     o

               -*nɛ* ~ -*n*

                 1→2

</div>

The reflexive morph -*nɛ* ~ -*n* may be analysed as a specific dual reflexive morpheme contrasting with a reflexive morpheme with non-dual meaning -*siŋ*, as in analysis E. Alternatively, the reflexive morph -*nɛ* ~ -*n* may be analysed as an allomorph of the reflexive morpheme in dual forms alongside an allomorph -*siŋ* indexing the reflexive in non-dual forms restricted to dual forms, as in analysis D. The latter analysis in terms of allomorphy is more satisfactory because it assumes a single morpheme -*siŋ* ~ -*nɛ* ~ -*n* with a single meaning and a well-defined pattern of allomorphy. This reflexive morpheme does not index actant number because this meaning is indexed by other morphemes in the affixal string. Analysis E may reflect the distinct provenance of the two morphs, although such a diachronic criterion, even if correct, need not be of decisive synchronic importance.

5. THE DUAL MORPHEME

The second suffixal slot SF2 is the tense slot. Non-preterite time is marked by zero, and preterite time is marked by the suffix -*ɛ* ~ *Ø* with a zero allomorph before another vowel.

The non-preterite zero suffix does not occur in forms in which tense is indexed by the non-preterite first person singular patient/subject morpheme *-ʔε*.

The old dual agent morpheme *-s ~ -tch*, the sole occupant of old suffixal slot SF3, immediately precedes and invariably co-occurs with the third person patient morpheme *-u*. Old suffixal slot SF3 could be eliminated by analysing the sequence *-s + -u*, or *-tch ~ -u* after the preterite morpheme *-ε*, as a DU→3 portmanteau morpheme *-su ~ -tchu*, indexing the transitive relationship between a dual agent and a third person patient and occupying the same position in the suffixal string as the third person patient morpheme *-u*. The sole advantage to approach G, however, is that it would eliminate a slot from the analysis. The drawback to analysis G is that the new portmanteau morpheme is too easily segmented into its two component parts identified in the original analysis F. The element /-u/ in the proposed DU→3 portmanteau *-su ~ -tchu* is clearly the third person patient morpheme *-u*, and the element *-s ~ -tch* clearly indexes dual number of agent. Analysis F, therefore, does not just represent a diachronic dissection of some fused morpheme.

	ANANLYSIS F			ANALYSIS G	
SF2	SF3	SF4		SF2	SF3
					-u
					3PAT
-ε~ Ø		*-u*		*-ε~ Ø*	
PT		3PAT		PT	
	-s ~ -tch				*-su ~ -tchu*
	DU.AG				DU→3
Ø		*-si ~ -tchi*		*Ø*	
NPT		DU.PAT.S		NPT	*-si ~ -tchi*
					DU.PAT.S

The old dual agent morpheme *-s ~ -tch* indexes dual number of the agent in 1DU.EXC→3, 1DU.EXC→3, 2DU→3 and 3DU→3 forms. The old dual patient/subject morpheme *-si ~ -tchi* indexes dual number of patient in 1SG→2DU, 3→2DU, 3→1DU.INC and 3→1DU.EXC forms, and dual number of subject in all intransitive dual forms. The old dual agent morpheme *-s ~ -tch* and the old dual patient/subject morpheme *-si ~ -tchi* can be analysed as a single entity, viz. a dual morpheme *-si* in suffixal slot SF3, with the regular allomorphs *-s ~ -tch* before the third person patient morpheme *-u*, and the regular allomorphs *-tch ~ -tchi* after the preterite morpheme *-ε ~ Ø*, the 1→2 portmanteau *-nε ~ -n* and the reflexive suffix *-siŋ ~ -nε ~ -n*. This is analysis H.

	ANALYSIS H	
SF2	SF3	SF4
-ε~ Ø		
PT		
	-si ~ -s ~	
	-tchi ~ -tch	*-u*
		3PAT
Ø		
NPT	DU	

6. FIRST PERSON SINGULAR MORPHEMES

The morpheme *-paŋ* remains puzzling, as was explained in my Limbu grammar. In modern Phedāppe Limbu, the suffix *-paŋ* is obligatory in negative preterite 1SG→3 forms, for example *mɛ-biː-baŋ-si-ŋ* 'I didn't give it to them' (1SG→3NSG.PT) – cf. *mɛ-biː-ʔɛ-n-chi-n* 'I shan't give it to them' (1SG→3NSG.NPT), but the suffix *-paŋ* is optional in intransitive negative preterite forms with a first person singular subject. This leads to doublets such as *yaʔ-mɛn-laːk-paŋ* 'I did not partake in the rice harvest dance' and *yaʔ-mɛ-raːkt-aŋ-nɛn* with the same basic meaning. Where such doublets exist, difference in form expresses difference in meaning, and I have recently come to understand the semantic distinction between such negative preterite forms with a first singular subject. Moreover, the fact that the suffix *-paŋ* is optional in intransitive forms but obligatory in transitive forms leads to the conjecture that this morpheme was originally limited to the transitive paradigm and was absorbed into the intransitive paradigm secondarily. The difference in meaning between negative preterite doublets with a first singular subject strongly supports this conjecture.

Negative preterite first singular intransitive forms in *-paŋ* co-exist with the 'regular' intransitive forms. The intransitive forms *kɛraŋ* 'I have arrived' and *khiːʔraŋ* 'I am (i.e. have become) tired'[1] are negated either as *mɛgɛraŋnɛn* 'I haven't arrived' and *mɛkhiːʔraŋnɛn* 'I am not tired' or as *mɛŋgɛppaŋ* and *mɛŋkhiːppaŋ*, respectively. The difference in meaning, as explained by my old friend Nārāyaṇ Prasād Paṅyāṅgu of Tamphulā village (pers. comm., Kathmandu, June 1993) is that the 'regular' first singular intransitive affixation in *mɛ-Σ-aŋ-nɛn* is simply the negative alternative to its affirmative preterite counterpart in *-aŋ*, for example *Kɛraŋŋiː mɛgɛraŋnɛnniː?* 'Have I arrived or haven't I?'. The use of the suffix *-paŋ* in negative intransitive forms, however, adds a semantic dimension to the utterance. The added implication is that something has not happened that was supposed to happen or was expected to happen. For example, as an answer to a question one can either respond with *mɛdhaŋaŋnɛn* 'I didn't come up', the regular intransitive negative of *thaŋaŋ* 'I have come up', or with the more abrupt *mɛndhaŋbaŋ*. The latter form with the originally transitive suffix *-paŋ* can be construed as rude because the implication is: 'I was supposed to have come up, but I just didn't'. The regular intransitive form *mɛdhaŋaŋnɛn* on the other hand is a simple negation and is therefore more polite, that is 'I didn't come up' or, by implication, 'I couldn't come up'.[2] It appears that the meaning of the forms in *-paŋ* has something to do with the greater volitionality or agency of this originally transitive suffix. In combination with the lexical meaning of certain verbs the form in *-paŋ* is the more obvious choice, for example *mɛnchibaŋ* 'I'm not dead, I didn't die', suitable when the speaker has just survived a life-threatening situation or when the speaker, presumed dead, turns out not to have died. This is why the form *mɛsyaŋnɛn* 'I'm not dead' is an odd, albeit acceptable, utterance in modern Limbu, for it is as if the speaker is making the remark out of the blue, and, in fact, there are reportedly few conceivable situations in which *mɛsyaŋnɛn* would be a relevant utterance.

---

[1]   The stems of the Limbu verb *khiːpmaʔ* 'to get tired' are given incorrectly in the glossary of the Limbu grammar. They should be *khiːʔr ~ khiːʔ*.

[2]   Nep. *'āunu sakinā' jasto alikati bhāv huncha.*

Historically, then, the label of the morpheme -*paŋ* would have to have been (1SG→3.PT), whereas it has acquired the modern meaning (1SG.AG.S.PT) whilst retaining some of its original sense of agency or deliberacy associated with transitivity. This enables us to identify the element /p/ in -*paŋ*, originally (1SG→3.PT), with the Bahing anticipatory 2/3SG→3 copy morpheme -*p*, which occurs in the preterite forms of open-stem verbs before the Bahing preterite tense morpheme -*ta*. This is the first cognate element to be identified in a Kiranti language for any of the Bahing anticipatory copy phenomena. The original meaning of the suffix -*paŋ* (1SG→3.PT) explains the special sense in which it entered the intransitive paradigm, for example *mɛ-dhaŋ-aŋ-nɛn* 'I didn't come up' versus *mɛn-dhaŋ-baŋ* 'I didn't do it, viz. come up', yet this explanation is diachronic. Synchronically, the preterite first person singular agent/subject portmanteau -*paŋ* could then best be reassigned to new suffixal slot SF4, which groups this originally agentive affix together with other agent morphemes in a single slot and may more faithfully reflect historical reality.

Analysis I, however, assumes an alternative approach to the four first person singular morphemes identified in the old analysis, viz. the non-preterite first person singular patient/subject morpheme -*ʔɛ*, the preterite first person singular patient/subject suffix -*aŋ*, the first person singular agent morpheme -*ŋ* and the preterite first person singular agent/subject portmanteau -*paŋ*. A tenseless first person singular morpheme -*aŋ* ~ -*ŋ* is posited, which exhibits the regular allophone -*ŋ* after the third person patient suffix -*u*. Conversely, the regular zero allomorph of the preterite suffix -*ɛ*~ Ø occurs before the vowel-initial suffix -*aŋ* ~ -*ŋ* (1SG). This analysis unites the old suffixes -*aŋ* (1SG.PAT.S.PT) and -*ŋ* (1SG.AG). Similarly, the portmanteau -*paŋ* can be segmented into an element -*p* and the newly posited first person singular morpheme -*aŋ* ~ -*ŋ*. The element -*p* may be identified as a preterite third person patient morpheme, which would tally with its semantically motivated occurrence in negated first person intransitive forms, whereby third person patient agreement is with the logical proposition denoted by the verb and not enacted by the first person singular agent, as in the example *mɛn-dhaŋ-baŋ* 'I didn't do it, viz. come up'.

Historically, the preterite third person patient suffix -*p* reflects the labial character of the Proto Kiranti third person patient morpheme *-*u* and the plosive character of the initial /t/ of the Proto Kiranti preterite suffix *-*tɛ*. It is unclear what historical relationship exists between the Limbu preterite third person patient morpheme -*p* and the Bahing anticipatory 2SG/3SG→3 morpheme -*p*, but the occurrence of the several Bahing anticipatory copy morphemes is likewise connected with preterite tense, indexed by the suffix -*ta* in Bahing.

The old non-preterite first person singular patient/subject suffix -*ʔɛ* is relabelled as a non-preterite first person singular morpheme, rather than as a patient/subject morpheme, in accordance with its occurrence in negative 1SG→3 forms. In the old analysis, the occurrence of the suffix in these forms was attributed to the reduced transitivity of a negated situation. Now, a zero morph of the third person patient suffix may be assumed to occur in non-preterite negative 1SG→3 forms before the non-preterite first person singular suffix -*ʔɛ* (1SG.NPT) and in preterite 1PL.EXC→3 forms before the preterite first person plural exclusive agent/subject portmanteau -*mʔna*.

ANALYSIS I

SF3	SF4	SF5
	$-u \sim \emptyset$	$-\text{?}\varepsilon$
$-si \sim -s \sim$	3PAT	1SG.NPT
$-tchi \sim -tch$		
DU	$-p$	$-a\eta \sim -\eta$
	3PAT.PT	1SG

7. THE SUFFIXAL STRING COMPACTED

The old non-singular agent number morpheme *-tchi* in old suffixal slot SF7 occurs only in the 1NSG→2 form and was posited to account for the element *-tchi* in the 1NSG→2 ending *-nɛtchige*. The fact that the 1→2 portmanteau *-nɛ ~ -n* can be semantically and formally identified within the ending argues against defining the sequence *-nɛtchi* as a 1NSG→2 portmanteau co-occurring with the exclusive suffix *-ge ~ -be*. Moreover, positing such a portmanteau would necessitate adopting some independent *ad hoc* way of accounting for the otherwise regular preterite/non-preterite homophony in the 1NSG→2 form. The transitive paradigm does not distinguish a dual from a plural third person patient, and non-singular number of a third person patient is indexed by the third person non-singular patient morpheme *-si* in old suffixal slot SF8. However, there are no formal or positional reasons not to identify the old non-singular agent morpheme *-tchi* in 1NSG.EXC→2 forms with the old third person non-singular patient morpheme *-si*. A single non-singular number morpheme *-si ~ -tchi* can therefore be posited, with the regular allomorph *-tchi* after either the preterite tense suffix *-ɛ* or the 1→2 portmanteau *-nɛ*. The old singular patient number zero morph $\emptyset$ should likewise be analysed as a singular number morpheme. This redefined zero morph not only indexes singular number of a third person patient but also of a second person patient or subject.

The plural agent suffix *-m* must be assigned to new suffixal slot SF5 because the suffix follows the third person patient morpheme *-u* and precedes the non-singular morpheme *-si ~ -tchi* in suffixal slot SF7. The portmanteau *-mʔna*, which indexes a first person plural exclusive subject or agent in preterite time, must also be assigned to a position in the suffixal string before the non-singular morpheme *-si ~ -tchi* in suffixal slot SF7. The portmanteau *-mʔna* can best be assigned to the same suffixal position as the plural agent suffix *-m* to which it appears to be semantically and formally related.

The plural patient/subject morpheme *-i ~ $\emptyset$* indexes plural number of a first or second person patient in 3→1PL.INC, 3→1PL.EXC, 1SG→2PL and 3→2PL forms, and plural number of first or second person subject in intransitive forms. The zero allomorph of this morpheme occurs in indicative forms with a first person plural inclusive patient (viz. 3→1PL.INC forms) and in intransitive forms with a first person plural inclusive subject. However, the full allomorph *-i* of the morpheme occurs in intransitive adhortative forms with a first person plural inclusive subject. Attempts to identify the patient/subject morpheme *-i ~ $\emptyset$* with the element /-i/ in the dual morpheme *-si ~ -s ~ -tchi ~ -tch* are not productive. The suffix *-i ~ $\emptyset$* must be assigned to a position following the tense slot SF2 and before the copy morpheme slot, new suffixal slot SF8. Semantic arguments could be

advanced for assigning the morpheme -*i* ~ Ø to either suffixal slot SF3 or to suffixal slot SF5, but in view of the heterogeneous functions of suffixal slot SF5, it seems reasonable to assign the plural patient/subject suffix -*i* ~ Ø to this position.

A zero morph Ø was posited in the old analysis to index singular number of a first person agent in 1SG→2SG forms and singular number of a second person agent. This zero morph is retained in the new analysis and assigned to new suffixal slot SF5. Elsewhere in the paradigm, singular number of a first person actant is indexed independently by the first person singular morpheme -*aŋ* ~ -*ŋ* or the non-preterite first person singular morpheme -*ʔɛ*. Singular number of a third person agent or subject is indicated by a zero morph Ø in new prefixal slot PF3. Singular number of a third person patient, second person subject and second person patient is indicated by the singular zero morph Ø in new suffixal slot SF7. The new analysis therefore posits three zero morphs indexing singular number, whereas only two singular zero morphs would, in principle, be necessary to account for actant number marking in, for example, 1SG→2SG and 2SG→3SG forms. However, the distribution of the three zero morphs (viz. Ø (3SG.Ag.S) in PF3, Ø (SG.AG) in SF5 and Ø (SG) in SF7) more accurately reflects the split ergative structure of the Limbu paradigm, whereby patient and subject roles are opposed to agent role in the first and second person, and agent and subject roles are opposed to patient role in the third person and, curiously, in the preterite of first person plural exclusive forms. The posited pattern of singular number zero morphs relies less upon what Michailovsky (1989:472) calls "choix arbitraires" in the analysis of the Limbu verb than would positing two zero morphs to account for all forms, but follows instead from symmetries inherent to the Limbu affixal agreement pattern.

INITIAL PORTION OF THE LIMBU SUFFIXAL STRING
(suffixal slots SF1 to SF5)

SF1	SF2	SF3	SF4	SF5
				-*ʔɛ*
				1SG.NPT
				-*aŋ* ~ -*ŋ*
				1SG
-*siŋ* ~ -*nɛ* ~ -*n*	-*ɛ* ~ Ø		-*u* ~ Ø	
REF	PT	-*si* ~ -*s* ~	3PAT	-*m*
		-*tchi* ~ -*tch*		PL.AG
-*nɛ* ~ -*n*	Ø	DU	-*p*	
1→2	NPT		3PAT.PT	-*mʔna*
				1PL.EXC.AG.S.PT
				Ø
				SG.AG

Because, in the new analysis, the plural agent morpheme -*m* and the preterite first person plural agent/subject portmanteau -*mʔna* have been assigned to suffixal slot SF5, and because the old non-singular agent morpheme -*tchi* has been eliminated from the analysis as an

independent morpheme, old suffixal slot SF7 has been abolished. The copy slot, new suffixal slot SF8, now contains copies only of morphemes occurring in new suffixal slot SF5 rather than copies of morphemes occurring in either of two suffixal slots, as in the old analysis.

The remaining suffixal slots are unproblematic. The position of the third element of the negative simulfix -*n* is posited as immediately preceding the slot containing the non-singular number morpheme -*si*, with which it invariably co-occurs. The first person singular copy morpheme -*ŋ* and the plural agent copy morpheme -*m* occur together in suffixal slot SF8. Whereas the plural agent copy morpheme -*m* echoes the homophonous agentive morpheme in suffixal slot SF8, the first person singular copy morpheme -*ŋ* reiterates the notion of first singular agent indexed by either the first singular morpheme -*ŋ* in suffixal slot SF5 or the 1→2 portmanteau -*nɛ* ~ -*n* in suffixal slot SF1. In suffixal slot SF9, the exclusive suffix -*ge* ~ -*be* indexes first person involvement and the exclusion of second person, and the inclusion of second person in first person forms is indexed by a zero morph ∅.

FINAL PORTION OF THE LIMBU SUFFIXAL STRING
(suffixal slots SF6 to SF10)

SF6	SF7	SF8	SF9	SF10
	-*si*	-*ŋ*	-*ge* ~ -*be*	
	NSG	1SG	EXC	
-*n*				-*nɛn* ~ -*n*
NEG₃	∅	-*m*	∅	NEG₂
	SG	PL.AG	INC	

8. CONCLUSIONS

The model of Limbu conjugational morphology has been refined. Both the number of suffixes and the number of suffixal slots have been reduced. Conversely, the number of prefixal positions identified in the analysis has increased. Morphemes have been redefined or more accurately formulated. One zero morpheme has been eliminated, and the distribution of zero morphs indexing singular actant number appears to more satisfactorily reflect the psychological reality of Limbu conjugational morphology. Negation in the Limbu simplex has been analysed as a discontinuous morpheme, for which we adopt Hagège's term 'simulfix'.

The choice between the two alternative negative preterite first person singular simplicia in the intransitive paradigm has been shown to be semantically motivated. The difference in meaning enables us to retrace the diachronic development which led to the existence of alternative negative preterite first person singular forms in the intransitive paradigm.

The number of dual morphemes has been reduced from two to one, and the number of generalised dual morphemes, that is non-singular morphemes derived from older dual **-si*, has likewise been reduced from two to one. The dual morpheme occurs amidst the initial suffixes, and the generalised dual morpheme occurs in the final portion of the string, so that the expansion of the original dual meaning to encompass the notion of non-singularity is a

phenomenon associated with the final portion of the suffixal string. This corroborates the idea that this 'generalisation' of the meaning was a secondary development observed in more recent accretions.

There are two diachronic implications of this new analysis of the Limbu verb. The first of these is that it is more plausible to posit a single dual morpheme in the initial portion of the suffixal string of the Proto Kiranti verb, as in the first and second models developed for the common ancestral verbal agreement system (viz. van Driem 1990, 1992). However, this protomorpheme *-ci* would have to be glossed simply as a dual morpheme, rather than a morpheme-indexing dual number of specifically first and second person actants. The two homophonous dual morphemes posited in the third model of the Proto Kiranti verb (viz. van Driem 1991) do not reflect distinct entities in Proto Kiranti, but subsequent developments in Bahing and Limbu resulting from the split ergative nature underlying Kiranti conjugational morphology, whereby first and second person actants are encoded in the verb according to an ergative pattern, and third person actants are encoded according to an accusative pattern.

The second diachronic implication of this new analysis is the greater verisimilitude of the simpler, first model of the Proto Kiranti verb in terms of the way it posits a single first person singular morpheme *-ŋ* in the same suffixal position as the third person patient protomorpheme *-u*, that is immediately after the position of the dual protomorpheme *-ci*. Whereas the Hayu and Limbu data reflect an old sequence *-u-ŋ*, Thulung and Dumi reflect an older element order *-ŋ-u*, and the conjugations of some Kiranti languages provide no indication for the relative position of these protomorphemes. In the second and third models of the Proto Kiranti verb, two additional tensed first person singular protomorphemes were posited in a position immediately anterior to the dual protomorpheme *-ci*, viz. the preterite first person singular protomorpheme *-aŋ* and its non-preterite counterpart *-ŋa*. These too seem to be artificial constructs not corresponding to distinct entities in the Proto Kiranti verbal agreement system, but reflecting subsequent developments in some Kiranti languages whereby a first person singular protomorpheme, perhaps best reconstructed as *-ŋa ~ -ŋ*, when immediately preceded by the preterite tense protomorpheme *-tε*, gave rise in some languages to a reanalysis of adjacent segments in terms of a preterite tense first person singular morpheme *-aŋ*.

The third model of the Proto Kiranti verb (van Driem 1991:354) should therefore be modified as follows: (1) The tensed first person singular morphemes *-aŋ* (1SG.PT) and *-ŋa* (1SG.NPT) may be eliminated, and (2) instead of the two dual morphemes preceding the third person patient protomorpheme *-u*, a single dual protomorpheme *-ci* should be posited at this location. Data will have to be continuously reassessed if we are to trace the evolution of Tibeto-Burman conjugational systems and to discern just which phenomena are regional developments. In Tibeto-Burman languages beyond the Kirant, there is little evidence for an older element order *-u-ŋ* (3PAT-1SG), which is reflected in the Limbu and the Hayu verb, and the many reflexes of the Proto Tibeto-Burman first person singular morpheme *-ŋ ~ -aŋ* outside of the Himalayas suggest that this morpheme occupied an anterior position following the verb in the Proto Tibeto-Burman verbal agreement system. Yet it is still too early to say just which first person singular agreement phenomena in Kiranti are the results of secondary developments. A Tibeto-Burman dual protomorpheme *-si* in a position anterior to the third person patient protomorpheme *-u* is well reflected outside the Himalayas, yet the hypothetical Tibeto-Burman dual patient protomorpheme *-si* in the final portion of the suffixal string was posited primarily on the basis of Kiranti data (van Driem

1993a). This, together with the fact that the Kiranti evidence itself suggests that the posterior reflex is a later accretion, casts doubt on the existence of the latter suffix at the Tibeto-Burman level and implies that this detail in the model for Proto Tibeto-Burman verbal agreement should be revised accordingly.

REFERENCES

Driem, George van, 1987, *A grammar of Limbu.* Berlin: Mouton.

    1990, An exploration of Proto-Kiranti verbal morphology. *Acta Linguistica Hafniensia* 22:27-48.

    1991, Bahing and the Proto-Kiranti verb. *Bulletin of the School of Oriental and African Studies* 54/2:336-356.

    1992, Le proto-kiranti revisité, morphologie verbale du lohorung. *Acta Linguistica Hafniensia* 24:33-75.

    1993a, The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies* 56/2:292-334.

    1993b, *A grammar of Dumi.* Berlin: Mouton de Gruyter.

    forthcoming, *A grammar of Lohorung.*

Ebert, Karen, 1991, Inverse and pseudoinverse prefixes in Kiranti languages: evidence from Belhare, Athpare and Dungmali. *Linguistics of the Tibeto-Burnan Area* 14/1:73-92.

    1992, Kiranti suffix copying (and possible origins of tense-aspect markers). Paper presented at the XXVth International Conference on Sino-Tibetan Languages and Linguistics, University of California at Berkeley, 14 October 1992.

Hagège, Claude, 1986, *La langue palau, une curiosité typologique.* München: Wilhelm Fink Verlag.

Michailovsky, Boyd, 1989, Compte-rendu de *A grammar of Limbu* par George van Driem. *Bulletin de la Société de Linguistique de Paris* 84/2:470-473.

Sprigg R. Keith, 1989, Review of *A grammar of Limbu* by George van Driem. *Bulletin of the School of Oriental and African Studies* 52/1:163-165.

LEPCHA ORTHOGRAPHY: AN EARLIER AND A LATER STAGE

RICHARD KEITH SPRIGG

It is remarkable that the language of the Lepchas, with only 34,894 speakers recorded for it in 1909, "roughly estimated as follows: – Sikkim...25,000; Darjeeling...9,894" (*Linguistic survey of India* III/1 p.233), when Nepali was just beginning to oust the numerous tribal languages of Sikkim and eastern Nepal, should be so prominent in Sino-Tibetan reconstruction; but both Bodman (1980, 1988) and Chang and Chang (1975) have made quite wide use of Lepcha cognates. Indeed, Chang and Chang (1975:398) declare that "the frequent usefulness of comparisons with Lepcha will be apparent in our discussions below". One of Bodman's (1980:73) examples, "*ta-rók* 'six', Tibetan *drug*, Chinese *C-rùk*", appears in §2.2.1 below. Where romanised, all my Lepcha examples are in accordance with Mainwaring (1876), though with slight modifications, including those cited from Grünwedel's editing of Mainwaring (1898).

It is because of this fairly wide use of Lepcha in Tibeto-Burman and Sino-Tibetan reconstruction that I consider variation in Lepcha spelling to be important. If Lepcha cognates are to be reliable for this purpose, inconsistencies in spelling need to be accounted for; otherwise they are liable to cause confusion to specialists in this field.

The Lepcha language is also remarkable for having a script of its own, a distinction that it shares with only two other languages of the Himalayas, Newari and Limbu (see Sprigg 1983:306-308 and Sprigg 1986:27-29). The Lepcha script, it has been claimed, was devised by the third Rajah of Sikkim, Chador Namgyal (1700–1717) (Risley 1894:13). If this account of its origin is correct, the script had been in existence for at least a hundred years by the time that the two earliest datable texts known to me were written; they are eyewitness accounts of the murder of the Lepcha Prime Minister of Sikkim, in 1826 (Risley 1894:19), and together form item 190 in the Foreign Department's 'Persian' section of the National Archive, Delhi, dated 14 April 1828. The one I have cited in this article, as item (i), is headed *gyá-mú-nun shu* 'submitted by Gyamu'. Twenty-one years later, in 1849, came a short statement of accounts, which I have published in Sprigg (1983); this is my source (ii). The next sources that I have used are printed: Colonel (later General) G.B. Mainwaring's grammar (Mainwaring 1876) and his dictionary (Mainwaring 1898). These two sources, (iii)

---

and (iv), are approximately one hundred years old. Finally, as a contemporary source, my source (v), I have drawn on my typescript copy of Tamsang (1994).

The interesting thing for me about these five sources, jointly extending over a period of 165 years, is that the later sources show fairly systematic differences in spelling from the earlier sources. I interpret these differences as evidence of a move away from a Tibetan model at an earlier stage of the language's orthography towards a later stage in which the model is closer to the Devanagari script, used for Hindustani and Nepali.

Since space is limited, it is necessary to keep comments on the fifteen types of variation to the minimum, and let the chronological order of the change, (i) 1828, (ii) 1849, (iii) 1876, (iv) 1898, and (v) 1994, speak for itself.

1. FINAL VOWEL ([ə:]); CHANGE FROM -*a* TO -*â*

I have used the circumflex to romanise "a sort of circumflex...called *rân*" (Mainwaring 1876:5), which Grünwedel has accounted for as a loan word from Tibetan, *sgron* 'cover, lay over, adorn':

	issue	cause to issue	know	date	pool	warm	certain	hand
i.	*pla*	*plya*	*ya*					
ii.				*tsha*	*da*			
iii.	*pla*		*ya/yâ*		*da*	*'â*	*nga*	(*'á-*) *ka/kâ*
iv.	*pla*	*plya*	*ya/yâ*	*tsha*	*da*	*'al'â*	*nga/ngâ*	(*'á-*)*ka/kâ*
v.	*plâ*	*plyâ*	*yâ*	*tshâ*	*dâ*	*'a*		*ka*

1.1 GRAMMATOLOGY AND THE SYMBOLS FOR -*a* AND -*â*

The above examples show a close relationship between the -*a* spelling and the -*â* spelling for these vowel-final lexical items, so close that one spelling can easily develop into the other; but from the point of view of grammatological categories the two types of spelling are very different: the -*â* type of spelling, as in *yâ* ᴇ᷈ 'know', for example, is alphabetic; and the two sounds have their own separate symbols, the syllabic vowel sound ([ə:]) being symbolised by the circumflex-like superscript symbol *rân*. The -*a* spellings, on the other hand, are diphonic; and the two sounds of *ya* ᴇ 'know' have a single symbol; so both the initial sound and its following syllabic vowel sound ([ə:]) share this symbol.

In Lepcha it is only that single syllabic-vowel unit, pronounced [ə:], that is (or used to be in 1828 and 1849) symbolised diphonically, the other syllabic vowel units having each its own alphabetic symbol; Japanese, on the other hand, makes much wider use of the diphonic category, for all five syllabic vowels, in forty out of the fifty members of the *gojuu on*; for example *ka*, *ki*, *ku*, *ke*, *ko*, *ya*, *yu*, *yo*, *wa*, *wo*, leaving only *a* and *o*, *u*, and the various forms of *i* and *e*, and the final consonant -*n* to monophonic symbolisation.

The example *ya* or *yâ* ɛ/ɛ̃ 'know' in the above table, like a number of other verbal items that have a vowel-final root, has an inflected form in *-m*, for example *yâm-bo* ɛ̃ ꠰ꠐ 'one acquainted with'; cf. also *á-'âm* 'warmth' (Mainwaring 1876:100, 111), *plâm-bo* 'going forth', *plyâm kón* 'let it issue' (Mainwaring 1898:228). These inflected forms are consonant-final ([ə:m]) and, as such, are invariably spelt with the *rân* vowel symbol. Possibly this obligatory spelling *-âm*, with the *rân* symbol, for these consonant-final inflected forms made it easier for this same alphabetic spelling to spread to their corresponding root forms, helping to promote the change from *-a* to *-â*.

2. FINAL CONSONANT ([-am/p/t/n/r/l; -ɒk/ŋ; -ɛm/p/t/n/r/l])

2.1 FINAL [-am/p/t/n/r/l]; CHANGE FROM *-am/p/t/n/r/l* TO *ám/p/t/n/r/l*

2.1.1 [-am]; CHANGE FROM *-am* TO *-ám*

	three	-ing	altogether	Lingdam (?slope crown)
i.	*sam/sám*	*-bam/bám*		
ii.	*sam/sám*		*gun-jam*	*líng-dam*
iii.	*sám*	*-bám*		
iv.	*sam/sám*	*-bam/bám*	*gun-jám*	*dám*
v.	*sám*	*-bám*	*gun-jám*	*dám*

2.1.2 [-ap]; CHANGE FROM *-ap* TO *-áp*

	-teen	bury	noose-trap	short	knock together
ii.	*thap*				
iii.	*tháp*	*lap*			
iv.	*tháp*	*lap/láp*	*tap/táp*	*map*	*bap/báp*
v.	*tháp*	*láp*	*táp*		*báp*

2.1.3 [-at]; CHANGE FROM *-at* TO *-át* (BUT IN SOME WORDS *-ât*)

	prime minister	blow	earth	mouldy	copulate
i.	*sháng-zat*				
iii.		*mat*	*fat*		
iv.	*chóng/chang-zât*	*mat/mut*	*fat/fát*	*mat/mát*	*'at/'át*
v.	*cháng-zât*	*mát*	*fát*	*mát*	*'át*

2.1.4 [-an]; CHANGE FROM -*an* TO -*án*

	having	burn	drink (n.)	watcher	warm
i.	-*bán*				
iii.	-*bán/ban*	*fan*	*'á-than*	*ran-bo*	
iv.	-*ban/bán*	*fan/fán*	*'á-than*	*ran-bo*	*'an*
v.	-*bán*	*fán*	(*'á-thóng*)	(*róng-bú*)	*'án*

2.1.5 [-ar]; CHANGE FROM -*ar* TO -*ár* (BUT IN SOME WORDS -*âr*)

	therefore	price	rust	rotten	prosper
i.	*'ar-nun*				
ii.		*far/fár*			
iii.	*'âr-nun*	*fár*			
iv.	*'âr-nun*	*fár*	*far/fár*	*sar/sár/ser*	*tar/tár*
v.	*'âr-nun*	*fár*	*fár*		*tár*

2.1.6 [-al]; CHANGE FROM -*al* TO -*ál*

	new	dibble	tomorrow	fall off
iii.	*'al*	*mal*	*lúk-kal*	
iv.	*'al/'ál*	*mál*	*lúk-kal*	*fal/fál*
v.	*'ál*	*mál*	*lúk-kál*	

2.2 FINAL [-ɒk/ŋ]; CHANGE FROM -*ak/ng* TO -*ók/ng*

2.2.1 [-ɒk]; CHANGE FROM -*ak* TO -*ók*

	seven	six	pour	target	stomach	hurt	befall
ii.	*ka-kyak*	*ta-rak*					
iii.	*ka-kyak*	*ta-rak*	*lak*	*mak*	*ta-bak*	*dak*	*zak*
iv.	*ka-kyak*	*ta-rak*	*lak*	*mak*	*ta-bak*	*dak*	*zak*
v.	*ka-kyók*	*ta-rók*	*lók*	*mók*	*ta-bók*	*dók*	*zók*

2.2.2 [-ɒŋ]; CHANGE FROM -*ang* (*nyín-dó*) TO -*óng*

	now	upon	stone	Lepcha	tiger
i.	*'á-lang*				
iii.	*'á-lang*	*'á-plang*	*lang*	*róng*	*sa-thang*
iv.	*'á-lang*	*'á-plang*	*lang*	*róng/rang*	*sa-thang/thóng*
v.	*'á-lóng*	*'á-plóng*	*lang*	*róng*	*sa-thóng*

2.3 FINAL [-ɛm/p/t/n/r/l];

change from -yam/p/t/n/r/l to -(y)em/p/t/n/r/l;

change from c/j/nyap, etc. to c/j/nyep, etc.

2.3.1 [-ɛm]; CHANGE FROM -yam TO -(y)em

	play	put on	wrist/upper arm	ball of thread
iv.	*lyam*	*dyam*	*ka-tyam/('á-)kâ-tyám*	*ki-tyam*
v.	*lem*	*dem*	*ka-tyem*	*ki-tem*

2.3.2 [-ɛp]; CHANGE FROM -yap TO -ep, AND FROM cap TO cep

	smeared	bore	along with
iv.	*lyap*	*cap*	*dyap/dyep*
v.	*lep*	*cep*	*dep*

2.3.3 [-ɛt]; CHANGE FROM -yat TO -(y)et, AND FROM nyat TO nyet

	two	about to	father-in-law
i.	*nyet*		
ii.	*nyat*		
iii.	*nyat*	*det*	*'á-fyat*
iv.	*nyat*	*dyat/d(y)et*	*'á-fyat*
v.	*nyet*	*det*	*'á-fyet*

2.3.4 [-ɛn]; CHANGE FROM -yan TO -(y)en, AND FROM j/nyan TO j/nyen

	faith	bad	listen	enemy/warfare
ii.	*yuk-dyan*			
iii.	*den*	*'á-jan*	*nyan*	*fyan*
iv.	*dyan/den*	*'á-jan/jen*	*nyan*	*fyan/fyen*
v.	*den*	*jen*	*nyen*	*fyen*

2.3.5 [-ɛr]; CHANGE FROM -yar TO -yer, AND FROM nyar TO nyer

	thunderbolt	wane	mud	sunset
i.	*sa-dyar*			
ii.		*nyar*		
iii.	*sa-dyar*		*suk-vyar*	
iv.	*sa-dyar/dyer*	*nyer*	*suk-vyar/vyer*	*tsuk-kyar/kyer*
v.	*sa-dyer*	*nyer*		*tsuk-kyer*

2.3.6 [-ɛl]; CHANGE FROM -*yal* TO -*(y)el*, AND FROM *j/nyal* TO *j/nyel*

	pronounce	weary	roll down	feathers	gums
iii.	*jal*	*pyal*			
iv.	*jal/jel*	*pyal/pyel*	*tyal/tyel*	*myal/myel*	*nyal/nyel*
v.	*jel*	*pyel*	*tel*	*mel*	*nyel*

In this section (§2.3) the vowel is [ɛ] throughout, that is front, spread, and half-open; this vowel contrasts with a half-close vowel [e], in the case of lexical items having a final plosive or nasal consonant (for example -*em/p/t/n* versus -*em/p/t/n*), but not in the case of those having a final rolled or lateral consonant, [-ɛr/l]. This means that such lexical items as *sa-dyar* 'thunderbolt' and *nyar* 'wane', and *jal* 'pronounce' and *nyal* 'gums', can be written with either an -*ar* or an -*er* spelling, or an -*al* or an -*el* spelling, without confusing one lexical item with another; but Tamsang's spellings with -*em/p/t/n* under-symbolise Lepcha phonology; they fail to distinguish the [ɛ] vowel unit from the [e] vowel unit (except where the initial consonant is nasal, in which case there is no such distinction; see Sprigg 1989:226).

2.4 GRAMMATOLOGY AND THE SYMBOLS FOR -*am* AND -*ám*, ETC.

From the grammatological point of view the -*am/p/t/n/r/l* spellings of §2.1 above are just as different from the -*ám/p/t/n/r/l* spellings as the -*a* spellings are from the -*â* spellings of §1. and §1.1; and the same is true of the -*ak/ng* spellings of §2.2 versus the -*ók/ng* spellings, and the -*yam/p/t/n/r/l* spellings of §2.3 versus the -*em/p/t/n/r/l*; only the pairs of vowel symbols are different. The lexical items in which -*a*- is used all make use of the diphonic category as regards the initial consonant and the vowel, with the result that *sam* 'three', for example, and *dak* 'hurt' have a single diphonic symbol for [sa-] and [dɒ-] (ﺳﻦ , ﭪ), while the later spellings, with a separate symbol in each case for the vowel (ﺷﻦ , ﭪ), have alphabetic spellings. The same principle is at work for forms such as *lyam* 'play' (ﻟﯘ), for [lɛ-], but with a contribution towards symbolising the vowel from -*y*-, a postscript symbol in Lepcha writing.

3. CHANGE OF SPELLING OR CHANGE OF PRONUNCIATION

Lepchas themselves are sometimes disconcerted by these variations. While reading with me through the text of Hooker's accounts (1849, my source (ii), A.R. Foning, of the Kalimpong Lepcha Association, and author of *Lepcha, my vanishing tribe* (Foning 1987), was astonished to find the word for 'three' spelt not *sám*, the spelling that he was used to, but *sam*. He concluded that since the spelling -*a* represents a long half-close central spread vowel ([ə:]) when it occurs in final position, as in the word *da* 'pool', it must be the same for -*a*- when it occurs in consonant-final words as well, as in the word *sam* in this text; therefore the change in spelling from *sam* to *sám* must also, following his view, represent a change of

pronunciation from [sə:m] to [sam]. I, on the other hand, concluded that only the spelling had changed, from *sam* to *sám*, between 1849 and the present time, while the pronunciation had remained the same during that time, as [sam]; and I was able to support my conclusion from the fact that the writer of these accounts had been inconsistent: though he had used the spelling *sam* on lines 1, 17, 29, 33 and 34, he had used *sám* on line 10 (Sprigg 1983:315). A somewhat similar variation occurs in the word for 'price', which he has spelt *far* on line 12 but *fár* on lines 4, 8, 9, 11, 25 and 29; and this I see as further support for my view that, as regards *a* and *á*, the writer's orthographic practice was not fixed.

I further believe that the -*am* spelling in *sam* 'three' (‮ّﺶ‬) is based on such models from Tibetan, spoken and written by Sikkim's neighbours to the north and to the east, as *sram* 'otter' (སྲམ་) and *bsam* 'thought' (བསམ་), both of which have the final [-am] in their pronunciation. In vowel-final Lepcha lexical items this same vowel sound, [a], is invariably symbolised by -*á*, as in *khá* 'score' and *dá* 'rest', from my source (i) (1828), and *rá* (name of the second month) and -*ká* 'to', from my source (ii) (1849) (‮ﭬﺭ‬ , ‮ﺥﺭ‬ , ‮ﻉﺭ‬), whence the final sequence of sounds [-am] can equally well be symbolised by -*ám*. After a period of alternation, when the writers of my sources (i) and (ii), in 1828 and 1849, hesitated between the -*am* spelling and the -*ám*, and in fact used both spellings in *sam* and *sám* and in -*bam* and -*bám* (‮ﺶ‬ , ‮ﺮّﺶ‬ , ‮ﻉ‬ , ‮ﺮّﻉ‬), an alternation that Mainwaring's dictionary recorded fifty years later, in 1898, the -*ám* type of spelling seems to have prevailed, and is the only spelling to occur in [am]-final lexical items in Tamsang's (1994) dictionary .

The predominance of the -*ám* over the -*am* spelling means that the Tibetan-like spelling for [-am], using -*am*, has been superseded by a spelling after the fashion of the Devanagari script, used for the nearby languages Nepali and Hindi, in which words with final [am] are symbolised by -*ám* (alternatively -*ām*), with a letter for the vowel [-a-], for example *nám* 'name', *ghám* 'sunshine', and *kám* 'work' (नाम , घाम, काम), and also after the fashion of the Bengali and Maithili scripts, all of which scripts are in use immediately to the west and to the south of the Lepcha-speaking area. This observation can stand equally well for the other consonant-final lexical items, in [-ap/t/n/r/l], considered in §2.1 above; and, with a change of vowel symbol to *ó*, it can stand for the [-ɒk/ŋ] of §2.2 and, with a change of vowel symbol to *e*, for the [-ɛm/p/t/n/r/l] of §2.3.

REFERENCES

Bodman, Nicholas C., 1980, Proto-Chinese and Sino-Tibetan: data towards establishing the nature of the relationship. In Franz van Coetsem and Linda R. Waugh, eds *Contributions to historical linguistics: issues and materials*, 33-199. Leiden: Brill.

1988, On the place of Lepcha in Sino-Tibetan: a lexical comparison. *Linguistics of the Tibeto-Burman Area* 11/1:1-26.

Chang, Betty Shefts and Kun Chang, 1975. Gyarong historical phonology. *Bulletin of the Institute of History and Philology*, Academia Sinica 46/3:391-524.

Foning, Arthur, 1987, *Lepcha, my vanishing tribe.* New Delhi: Sterling.

Grierson, George Abraham, ed., 1909, *Linguistic Survey of India,* III. (reprinted 1967, Delhi: Motilal Banarsidass)

gyá-mú-nun shu, 1828, National Archive, Foreign Department, Persian Section, item 90 (Delhi).

Mainwaring, George Byres, 1876, *A grammar of the Róng (Lepcha) language.* Calcutta: Baptist Mission Press.

1898, *Dictionary of the Lepcha language.* Revised and compiled by Albert Grünwedel. Berlin: Unger.

Risley, H.H. ed., 1894, *Gazetteer of Sikhim.* Calcutta: Bengal Government Secretariat.

Sprigg, Richard Keith, 1983, Hooker's expenses in Sikkim: an early Lepcha text. *Bulletin of the School of Oriental and African Studies* 46/2:305-325.

1986, The syllable finals of Tibetan loan words in Lepcha orthography. *Linguistics of the Tibeto-Burman Area* 9/1:27-46.

1989, Oral vowels and nasalized vowels in Lepcha (Rong) as the key to a puzzling variation in spelling. In Jeremy H.C.S. Davidson, ed. *South-East Asian linguistics: essays in honour of Eugénie J.A. Henderson*, 219-235. London: School of Oriental and African Studies.

Tamsang, Kharpoo, 1994, *Lepcha-English encyclopedic dictionary.* Kalimpong: Mani Press.