Cambridge Grammatical Descriptions

A Grammar of Kham

David E. Watters

CAMBRIDGE

more information - www.cambridge.org/0521812453

This page intentionally left blank

A Grammar of Kham

This is a comprehensive grammatical documentation of Kham, a previously undescribed language from west-central Nepal, belonging to the Tibeto-Burman language family. The language contains a number of grammatical systems that are of immediate relevance to current work on linguistic theory, including a functionally transparent split ergative system, a well developed system of mirativity, restrictive and non-restrictive noun phrases based on word order, a rich class of derived adjectivals, and extensive transitivity alternations in the verb. Its verb morphology has implications for the understanding of the history of the entire Tibeto-Burman family. The book, based on extensive fieldwork, deals with all major aspects of the language including segmental phonology, tone, word classes, noun phrases, nominalizations, transitivity alterations, tense–aspect–modality, non-declarative speech acts, and complex sentence structure. It provides copious examples throughout the exposition and includes three short native texts and a vocabulary of more than 400 words, many of them reconstructed for Proto-Kham and Proto-Tibeto-Burman. This book will be a valuable resource for typologists and general linguists alike.

DAVID WATTERS has been Director of the Oregon SIL linguistic training program in Eugene for the past four years and is currently collaborating with Nepalese scholars on an *Encyclopedia of Nepal's Languages.* He teaches periodically at the University of Oregon as adjunct faculty, and at Tribhuvan University, Kathmandu, as a visiting scholar. Whilst working in Nepal as a field linguist in 1969, Dr. Watters discovered the group of Tibeto-Burman languages known as Kham, and continued numerous successive expeditions into the territory until the early 1990s. He provided the people with written forms of their own language, along with modest amounts of literature, both from their own oral traditions and other outside works.

CAMBRIDGE GRAMMATICAL DESCRIPTIONS

Editors: R. M. W. Dixon, Keren Rice

This series is devoted to the publication of comprehensive descriptive grammars of languages that have not previously been documented, and that have interesting and unusual characteristics which will expand our understanding of human language in all its diversity and challenge the limits of current linguistic theory. Some of these languages are spoken by only a small number of people and can be considered endangered.

Each grammar briefly introduces the society in which the language is spoken, and covers the key areas of phonology, morphology and syntax, together with typological and historical considerations. In each case, a sample text or texts in the language are provided, with full gloss and translation. A glossary of basic vocabulary is also included. The series aims to provide theoretical linguists in the various subdisciplines with reliable data and analysis which will provide a permanent and invaluable set of source materials.

A GRAMMAR OF KHAM

David E. Watters

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published in printed format 2002

ISBN 0-511-03001-0 eBook (Adobe Reader)
ISBN 0-521-81245-3 hardback

To Hasta Ram Budha Kham
my teacher
and lifelong friend
and
to the memory of his son
Sukh Kham
man of faith and vision
whose light was extinguished
in his most promising hour

Contents

page

Tables

Figures

Maps

Preface

The discovery of the Kham group of languages in Nepal in 1969 is one of the remarkable finds in Tibeto-Burman linguistics this century – it happened against the backdrop of nearly two centuries of fairly intense linguistic activity in the whole of the Indian subcontinent. It was in this setting, for example, that Sir William Jones, in 1786, made his now-famous pronouncement before the Royal Asiatic Society in Calcutta that Greek, Latin, and Sanskrit had all 'sprung from some common source'; a source, which, 'perhaps, no longer exists.' His pronouncement profoundly changed the face of linguistics; language origins and language evolution became the new challenge of linguistic inquiry in the nineteenth century.

Sparked by the imagination of a new-found science, the British in India expanded their range of inquiry and began amassing a wealth of linguistic materials from numerous Himalayan languages and dialects – some, like Kusunda, with as few as a dozen speakers. Because the British had no direct access to Nepal, most of the early samples were collected by British military officers from Nepalese tribesmen serving as mercenaries in the British Gurkha army. Colonel Kirkpatrick, for example, collected a short vocabulary of the Magar language, spoken by one of the 'military tribes' of Nepal, as early as 1793, and Francis Hamilton, a British historian and philologist, deposited a more complete specimen of the same language in the Company's library sometime before 1814.

A few years later, Brian Hodgson, the British Minister at the Court of Nepal, beginning as early as 1828, published notes, observations, and essays on the languages and customs of several tribes of Nepal. Grierson's monumental 'Linguistic Survey of India,' published between 1903 and 1909, contains in one of its volumes (contributed by Sten Konow) a broad sampling of Himalayan languages with comprehensive notes on their vocabularies and grammars. Shafer, in an unpublished work of fifteen volumes on Sino-Tibetan linguistics between 1937 and 1941, and later in an edited version of the same work, published between 1966 and 1973, includes works on all the major Himalayan languages from every recognized branch of Tibeto-Burman.

Against this backdrop of linguistic activity, the failure to document Kham in any of its varieties is indeed a curious oversight. Kham, after all, is no small language – it is mother tongue to no less than forty or fifty thousand people living in the remote, upper valleys of mid-western Nepal. I first became aware of the possible existence of such a language from an American anthropologist, John Hitchcock, who had approached the edges of their tribal territory on a month's trek sometime in 1960-1962. He cordially apprised me of their general whereabouts in 1969. It was upon his advice and the

encouragement of Dr. Dor Bahadur Bista and University Vice-Chancellor Dr. T. N. Upraity that I began work on the language the same year under the auspices of The Summer Institute of Linguistics and Tribhuvan University, Kathmandu. It was not until 1971, when I produced my first paper on a mimeograph machine, that the language finally emerged from its long years of obscurity.

A sad fact of our times is the loss of the world's languages at an unprecedented rate. Michael Krauss, in an address at a symposium on language loss (1992), made the startling prediction that 90 percent of the world's languages will be extinct by the end of the twenty-first century. Even if his estimates are off by half, the loss to humankind is staggering. For millennia, the study of language has been viewed as an integral part of scientific inquiry into an adequate understanding of the human mind. The personal loss of the unique cultural nourishment afforded by a particular language to members of a community is even greater. Many have noted that a language, in many respects, is akin to a biological species. It is a uniquely human evolutionary achievement – 'as divine and endless a mystery as a living organism' (Kenneth Hale 1992). The loss of a single language, then, diminishes our world as surely as the loss of a biological species. Language loss is quite naturally a legitimate and critical concern to linguists. But it should be more; it is surely a human concern, one that should be shared by all people.

Reasons for language loss and extinction are not, in most cases, the result of deliberate attempts at 'glottocide,' the destruction of a people's language. The reasons are more subtle and nameless. In fact, there seems to be precious little that most of us can do to stop it. It is no longer economically viable for members of most small linguistic communities to remain isolated from the larger and more powerful majority cultures that surround them. To give them false enticements to continue in their native languages at the expense of economic well-being, however, would be justifiably looked upon as an act of linguistic imperialism and paternalism. Where a minority language does continue to survive in the face of economic and political pressure, it is because its speakers have learned to participate in the majority culture while at the same time receiving benefit, often more communal or spiritual than economic, from the minority culture.

One thing linguists can do, then, to help preserve minority languages at the local level is to help promote community pride in the minority language. Where the subtle pressure of an economically dominate culture encourages people to believe that their future depends on giving up their native language, steps need to be taken to level the playing field. They must be able to view their own language as a valuable heritage worth maintaining. Providing written forms of the language in practical orthographies, along with modest amounts of literature, both from the tribe's oral traditions and other works of high moral value, has proven in many cases to be a good, first step.

A generation ago, Kham began to lose some of its former efficacy. For generations they had lived efficiently in a kind of cultural backwater. In the 1960s, trade links to the north were severed and Kham speakers began to grow more dependent on their Nepali neighbors to the south. It became increasingly impracticable for them to live in isolation

from the mainstream of society. Nowadays, the language is at a crossroads. On the one hand, speakers of Kham have gained a great deal of linguistic and ethnic pride through country-wide nationalistic movements in the wake of a democratic revolution in 1990. On the other hand, some of the nationalistic movements, notably the Maoist movement with its beginnings among Kham speaking peoples, have political ambitions well beyond their traditional tribal territories, and Nepali is the only suitable vehicle. How the situation will play out remains to be seen. It is no longer possible for foreigners to gain safe access to Kham speaking areas, and it is only hoped that Kham speakers, in the midst of their new socio-political situation, will recognize the value of maintaining their language.

It has been a matter of great importance to document Kham in its entirety while it is still a healthy and vigorous language. Language death, where an issue, only makes the need more urgent; its absence does not obviate the still fundamental need for grammatical descriptions of little known languages. Languages need to be documented because they are 'supreme achievements of a uniquely human collective genius' (Kenneth Hale 1992). Language reveals the human mind. Sadly, few grammars, global in coverage, exist for Tibeto-Burman languages; most are short sketches of varying detail. A pressing need for further descriptions is obvious. Bernard Comrie (1991), in an appeal to field linguists everywhere, urged – 'Provide good descriptive grammars and dictionaries: theories come and go; the best descriptive grammars and dictionaries remain as lasting testimonials.'

Clearly, Takale Kham and its relationship to the Kham group of languages is a linguistic phenomenon of important status and deserving of extensive documentation. Because it provides historical links and new insights into a number of intriguing questions relating to the whole of Tibeto-Burman, it is of special interest to Tibeto-Burmanists. But it is more. Since the great diversity of languages in the Tibeto-Burman area is a commentary on the creative genius and diversity of the human mind, the description of another major language with a particular view to its diachronic pathways of creation is of interest to anyone interested in language and mind. Finally, we owe a debt of gratitude to a community of speakers whose language embodies a tradition of intellectual wealth found nowhere else.

Acknowledgments

Studying the grammar of an undescribed language cannot be carried out in the comforts of a study or a library. It must be done in the field, living with the 'keepers' of the language – the people who speak it and pass it on to successive generations. For an outsider to succeed, however dismally, to learn such a language is a testament to the patience and hospitality of the people themselves. The best days of my life have been spent around the fires of the Takales, the Nishels, the Gamales, and the Sheshis – in their villages, in their tents, and in their sheep camps at the foot of the glaciers. It was there that we shared food, swapped stories, laughed, wept, and dreamed. I will always be indebted to them for cheerfully sharing their language and giving me a glimpse of a way of life that is fast disappearing from the face of the earth.

I have many Nepalis to thank too, both in and out of government, without whose help I could never have reached the Kham territories. They provided permits, letters, and study visas. Many became personal friends. They are too numerous to mention. It would be impossible, however, not to mention the special friendship of many at the Central Department of Linguistics, Tribhuvan University: C. M. Bandhu, T. R. Kansakar, B. M. Dahal, Y. P. Yadav, N. M. Tuladhar, K. P. Malla, and M. P. Pokharel.

Numerous foreign scholars in Nepal, too, from many nations and a wide variety of institutions, have through the years provided invaluable insights into complex linguistic issues: Austin Hale, Ross Caughley, Warren Glover, Maria Hari, Sueyoshi Toba, Olavi Vesalainen, Marlene Schultze, Dora Bieri, Esther Strahm, Anita Maibaum, Boyd Michailovsky, Martine Mazaudon, Carol Genetti, George van Driem, Stephen Watters, Erik Andvik, and Balthasar Bickel, just to name a few. The Sino-Tibetan conferences, too, have always been a breeding ground of stimulation and encouragement. Much of my early interest in Tibeto-Burman issues came through Jim Matisoff, Paul Benedict, Graham Thurgood, and Jim Bauman.

My work would have gone nowhere had it not been for the instruction and encouragement of my professors at the University of Oregon. While still travelling to and from Nepal, Scott DeLancey encouraged me to 'Come study with us; write a grammar!' It was a privilege to have him as my advisor. And where would I be without Talmy Givón? He gave me the broad vision of language I needed, the sensible and sane framework around which my thoughts have been organized. More than that, he has been a friend, always available to discuss linguistics, philosophy, or whatever. And there is Doris Payne, who was unafraid to critique my writing where it was not clear. If what I have written is readable, it is in part due to her gentle prodding.

This book is a revision of my 1998 Ph.D. dissertation done at the University of Oregon. I am very grateful to my editors, Bob Dixon and Keren Rice, for taking it on. They were the best. Their list of suggested improvements were genuine improvements. In no way did they stifle me; they provided broad guidelines and gave me the freedom to express myself as I wished. Any errors in this book are truly my own. I am grateful, too, to the outside referee provided by Cambridge University Press, whose comments were both encouraging and constructive.

My deepest debt of gratitude is to my wife, Nancy. She always encouraged me to follow my dreams. In the early days she accompanied me, deprived of the amenities most modern women have come to expect. She did it with cheerfulness. In later years, she experienced days of loneliness, staying at home to raise our sons, while I tramped the remote corners of the earth. Thank you. I am forever in your debt!

Abbreviations

1ST	first (person)
1S	first singular (subject, object, or possessive)
1D	first dual (subject, object, or possessive)
1P	first plural (subject, object, or possessive)
2ND	second (person)
2S	second singular (subject, object, or possessive)
2D	second dual (subject, object, or possessive)
2P	second plural (subject, object, or possessive)
3RD	third (person)
3S	third singular (subject, object, or possessive)
3D	third dual (subject, object, or possessive)
3P	third plural (subject, object, or possessive)
A	the 'A' argument of a transitive clause
ABLE	abilitive
ABLT	ablative
ABS	absolutive
Adj	adjective
ADS	adessive
Adv	adverb
AFT	since, after
ALLT	allative
APPRX	approximative
ASC	associative
BEN	benefactive
CAUS	causative
CEP	counter-expectation particle
CIF	contrary information flow
CIS	cisative
CLSF	classifier
CMPR	comparative
COM	comitative
COME	come purposive
COMP	complement
CON	concessive
CONFIRM	confirmative
CONT	continuous aspect
COORD	coordinator
DAT	dative

DECL	declarative
DEL	delative
Dem	demonstrative
DEON	deontic
DETRANS	detransitivizer
DIM	diminutive
DIR	direct
DIS	distal
DL/dl	dual
DS	different subject
DT	detransitive
DUM	dummy (morphological place holder)
ELAT	elative
EMP	emphatic
ERG	ergative
EXPR	expressive adverb
FEM	feminine
FOC	contrastive focus
FUT	potential mode (future)
GEN	genitive
GO	go purposive
H-D	high followed by downstep pitch
H-H	high followed by high pitch
H-L	high followed by low pitch
H-M	high followed by mid pitch
HO	the ‘ho’ part of a discontinuous ‘probability’ morpheme
HOR	hortative
IA	Indo-Aryan
IF	conditional
IMP	imperative
IMPFV	imperfective
IN	inessive
INCPT	inceptive
INDEF	indefinite
INF	infinitive
INSTR	instrumental
intr	intransitive
INTRG	interrogative
Kh	Kham
LAT	lative
L-M	low followed by mid pitch

LOC/Loc	locative
MASC	masculine
m/f	male/female
MEL	main event line
MIR	mirative
M-L	mid followed by low pitch
MM	middle marking
M-M	mid followed by mid pitch
Mod	modifier
N	noun
NEG	negative
Nep.	Nepali
NF	non-final marker
NML/Nml	nominalizer
NOM	nominative
NP	noun phrase
NUM/Num	number
O	the 'O' argument of a transitive clause
OBJ	object(ive)
ON	superessive
OPT	optative
ORIENT	orientative
PASS	passive
PFV	perfective
PL/pl	plural
POSS	possessive
PRED	predictive
PRIOR	prior past
PROB	probability
PROG	progressive
PROH	prohibitive
PROS	prospective aspect
PROV	provisional
PROX	proximate
PSB	possibility modal
PTB	Proto-Tibeto-Burman
PURP	purposive
QP	question particle
RECIP	reciprocal
REFL	reflexive
Rel	relative clause

REM	remote
RSP	reported speech particle
S	the ‘S’ argument of an intransitive clause
SA	SA (confirmation particle)
SER	serial/concantenated verb
SG/sg	singular
SIM	similative
SUB	subjunctive
SUBJ	subject
T-1	Tone-1
T-2	Tone-2
TAG	tag question
TAM	tense, aspect, modality
UNTIL	until
VBL	verbalizer
V-T	(placement of) verb root and tense–aspect marking in paradigms
WELL	confirmative
WHEN	subjunctive/when

1 *The people and their language*

Kham, in all its varieties, is spoken primarily in the upper valleys of Rukum and Rolpa Districts of the Rapti Zone in Mid-Western Nepal (see map 1). A few thousand of the easternmost speakers spill over into the Nishi and Bhuji Khola regions in the western part of Baglung District, Dhaulagiri Zone. Only Sheshi, the southernmost dialect, is separated from the other dialects by populations of Nepali speakers. All other dialects are contiguous to one another, separated by uninhabited mountain barriers between eleven and thirteen thousand feet in altitude. (For more on geography, see §1.2.)

1.1 Language typology

Kham is a Tibeto-Burman (TB) language of the Bodic branch exhibiting many of the areal features defined for the 'Indospheric' side of the family – a gross generalization for the westernmost TB languages. Such languages have been variously influenced in phonology and grammatical structure by prolonged contact with Indic languages (in contrast to the 'Sinospheric' TB languages of Southeast Asia).

1.1.1 Tonal patterns

Tone in Kham can be described as belonging to a 'four-box' system. Two binary oppositions, 'voice register' (modal and lax) and 'melody' (Tone-1 and Tone-2), intersect to form four contrastive tone patterns. The melody opposition clearly predates the register split and may correlate with Benedict's (1972) tones *A and *B for Proto-Tibeto-Burman (PTB). Voice register was superimposed later and now divides the pitch range of Tones 1 and 2 into an upper and lower range.

In many Bodish languages, register is related (sometimes even synchronically) to the phonation type of onset consonants, and melody is tied to coda consonants and rhyme type. In Kham, the phenomenon is much older, and though the same historic transparency no longer exists, it can be assumed that the same tonogenetic apparatus was at work. Register, the more recent development, can be shown in some reconstructed forms to have derived from a lost *s-* prefix.

1.1.2 Word classes

Of the three major word classes, noun, verb, and adjective, only the first two can be fully supported in Kham on internal structural criteria. Adjective, as an inherent gram-

matical class, is almost non-existent. The entire class is composed of three native words – 'big,' 'small,' and 'short' (plus 'low' in some dialects) – and a handful of loan words from Nepali. All other words that serve in a modifying/adjectival function are derived by nominalization from some other word class, usually verbs or verbalized nouns.

The class of words generally known as adverbs belongs to several heterogeneous word classes – adverbs of time, manner, intensity, and a few others. A class more specific to the region is that of 'expressive adverb.' Expressives modify verbs by designating a specific manner in which the action takes place. A generic verb like 'go,' for example, can be modified to mean 'saunter,' 'amble,' 'stride defiantly,' and so on. Most expressive adverbs are derived from old verbs and occur in reduplicative structures to form rhyming couplets, like *kyasya kisi.*

A special class of 'deictic primitives' occurs in Kham, expressing notions like 'proximate,' 'distal,' 'remote,' 'up,' 'down,' 'front,' 'back,' 'left,' and 'right.' All primitives are bound roots followed obligatorily by locative suffixes or by a special class of landmark locations – 'up-country,' 'down-country,' 'one side of a mountain,' or 'one side of a valley.' Combinations are generative with their own syntax, capable of forming hundreds of complex locative expressions. The demonstratives 'this' and 'that,' in fact, are complex expressions, derived from deictic primitives by nominalization.

i. Morphology

Kham is highly agglutinative with a rich morphology. Nouns are inflected for a single prefixal position (possession or, with a few nouns, a classifier numeral), and several suffix positions, most of which are instantiated by local case markers – locative, adessive, inessive, superessive, cisative, allative, ablative, elative, delative, comparative, lative, orientative, and comitative. In addition, nouns are marked for three numbers, as well as for certain grammatical cases – ergative, instrumental, genitive, primary object, and associative.

Verbs are inflected for five prefix positions and seven suffix ones. Kham exhibits what has been referred to in the literature as 'pronominalizing' morphology. That is, in addition to marking the expected categories of tense, aspect, and modality, verbs also cross-reference the person and number of clause level referents – for Kham, both subject and objects. As far as is known, Kham is the only TB language that consistently marks both. The modern patterns can be shown to have arisen out of simpler material more consistent with the rest of TB, and Kiranti in particular.

Also part of verbal morphology are fairly elaborate derivations related to transitivity and voice – those of causativization and detransitivization.

ii. Case marking alignment

The general alignment of grammatical case markers in Kham is 'split ergative,' based on a person split in which first (1ST) and second (2ND) person rank high on a nominal hierarchy, and third (3RD) person ranks low. The case marking split has radical repercussions in verb morphology as well. Object marking is tied, ultimately, to pragmatic

notions of identifiability.

1.1.3 Constituent order

Kham has a basic constituent order of AOV, SV in both main and dependent clauses. The attendant 'harmonic orders' in phrase level syntax also occur: DemN, NumN, GN, AN, and RelN (where for Kham, A is a type of RelCl). A non-restrictive order for modification in NPs, a kind of appositive, also occurs in which the order of all constituents but DemN and GN are reversed to NNum, NA, and NRel.

1.1.4 Grammatical roles

Unlike what is reported for some TB languages, the grammatical roles 'subject' and 'object' are well motivated in Kham. In unmarked declarative clauses, S and A (usually agents, but not always) are grouped together and form the pivot for syntactic operations like clause chaining and subordination.

Though in normal, running discourse, the NP associated with S and A is usually missing, it is obligatorily indexed in the verb for person and number (as is the O argument). In detransitivizing operations, the clausal agent is deleted and the subject index in the detransitivized verb agrees with the patient.

1.1.5 Nominalization

Nominalization is a major syntactic device in Kham, and operates at all levels of the grammar. Almost all phrasal modifiers are nominalizations (including relative clauses), and all clauses embedded as sentential complements are nominalizations. Even main, independent clauses can be nominalized, and, as such, have special discourse functions.

Nominalizations cut across all speech acts – declaratives, interrogatives, and imperatives. The nominalized versions are 'more discontinuous' and 'less direct' than the regular forms in a specifically defined way. Nominalized imperatives, for example, have the softened force of an optative, and nominalized interrogatives are less intrusive than their regular counterparts.

1.1.6 Clause chains

Clause chains are 'co-subordinate' structures in Kham and differ significantly from the subordinate structure of complements (which are always nominalized). In clause chains, all tense/aspect and person/number information is marked on the chain-final verb, and chain-medial verbs are marked with varying degrees of person inflection depending on whether the subject participant of the following clause has referential continuity with the current clause.

Most chains mark sequential events, but a few specify different aspects of the same

event. The two types are not morphologically distinct. The differences lie, in part, on the level of 'juncture' between the two clauses, and various tests can be devised for teasing them apart.

1.1.7 Evidentials

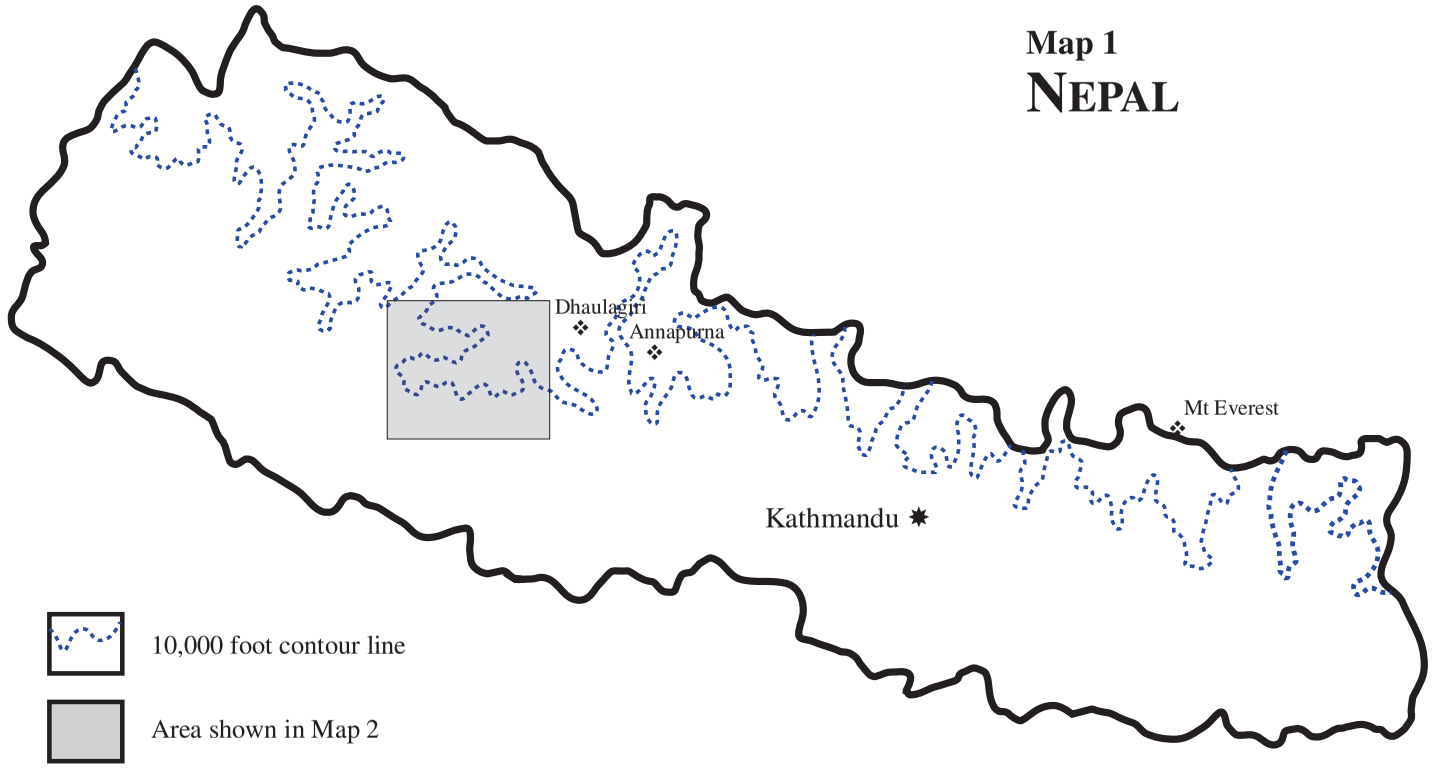
Kham lacks most evidential categories found in many Bodic-type languages. There is, however, a 'mirative' category and a 'reportative' category that covers at least some of the semantic space often associated with the evidential categories of hearsay and inference.

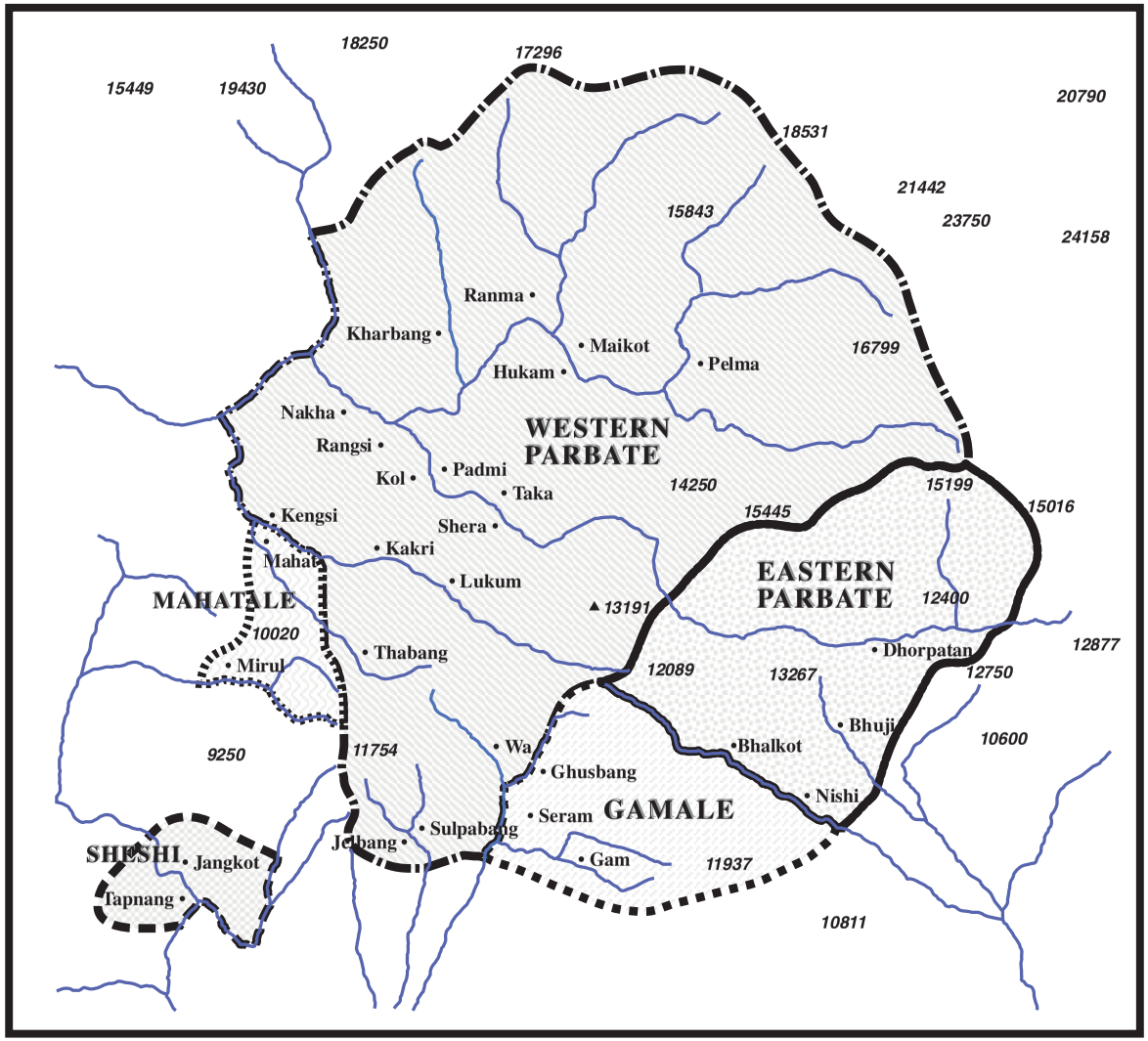
1.2 Villages and grazing lands

Most Kham speakers follow a transhumant pattern of life, maintaining both permanent and summer villages, but travelling extensively throughout much of the year with flocks of sheep and goats. Permanent villages are situated mostly on valley floors between altitudes of six and eight thousand feet, surrounded by fields suited to the cultivation of corn, barley, and millet. Summer settlements are situated at altitudes between nine and ten thousand feet where potatoes can be grown and the people have easy access to timber and subalpine pasture for cattle grazing.

Immediately above the summer settlements at altitudes beginning at about twelve thousand feet, all but the Sheshis have direct access to a large tract of rugged, alpine country running some 50 miles in length off the western end of Mt. Dhaulagiri (see map 2). The northern extreme is blocked by permanent glaciers and high, impassable mountains, except at a few points where passes cross over to the northern side of the Himalaya as low as fifteen or sixteen thousand feet. The area is prime habitat for Blue Sheep and the endangered Snow Leopard. It is in these high regions that the men spend their summers with large flocks of sheep and goats, living in *goths* and goat hair tents. The women and children remain in the summer settlements tending gardens and herding cattle. The large tract of alpine land is carved up into communally owned grazing lands, each controlled by a different village. A village's wealth is determined, in large part, by the size and location of its alpine holdings. Tolls are levied on herds passing through another's communal land, and longer stays are taxed on a per-diem, per-animal basis. Taka, with its cluster of four villages, is relatively wealthy in land holdings, and its occupants own more than 20,000 sheep.

The permanently inhabited villages in the north, especially those of the Maikot, Taka, Lukum, and Thabang valleys, are built in the fortress-like style of northwest Nepal. The villages are complex structures with as many as 300 houses built into a large, interconnecting unit on steep mountain walls. Long rows of flat-roofed houses are built like giant stair-steps, one row on top of another. Log ladders link one level of the near vertical complex to the next, and in some of the more rugged areas, tunnels link adjoining sections. Each house is made of unplastered stone, and heavy timbers support a flat, earthen roof.


Map 1
NEPAL
Dhaulagiri
Annapurna
Mt Everest
Kathmandu
10,000 foot contour line
Area shown in Map 2



Map 2

KHAM SPEAKING TERRITORIES

showing five major dialect groupings:

- Western Parbate
- Eastern Parbate
- Gamale
- Sheshi
- Mahatale

(Altitudes in feet)

The permanent villages of the Nishi and Bhuji valleys, as well as a few others at intermediate altitudes, are not so densely structured as the northern villages. The houses are built of stone with wooden shingle roofs. Further south, the villages occupied by the Gamales and Sheshis are fashioned more according to the style of traditional middle-hill villages found throughout Nepal. They are more spread out, and each house has an open veranda and a small courtyard, sometimes surrounded by small gardens. Houses are plastered with clay and the roofs are usually made of thatch.

1.3 Shamanism and origins

Kham speaking areas are tiny preserves of classical 'Inner Asian Shamanism' (Eliade 1964), a religious complex that once pervaded all of Siberia and Central Asia. Over the centuries, other religious complexes moved into the region either displacing the old shamanism, as in the case of Islam throughout much of Central Asia, or else reinterpreting and institutionalizing many of the old shamanistic motifs, as in the case of Lamaistic Buddhism throughout Tibet and Mongolia. Elements of the tradition have been preserved throughout the Himalayas in varying degrees, the Trans-Himalayan tribes having been heavily influenced by Buddhism in recent generations, and the Sub-Himalayan ones by Hinduism.

Kham communities, especially those in the more remote northern regions, have been only lightly touched by outside influences. Even the major Hindu holidays, those like *Dasain*, have gone unobserved by most communities until very recent years. Instead, they follow ancient shamanistic traditions in a remarkably pure form. All the major elements of Eliade's classical 'Inner Asian' construct are practiced in many Kham communities, some, down to meticulous details (Hitchcock 1967, Watters 1975a). The shaman's call, his initiation, his mythology, his healing technique, his familiar spirits, his costume – all show a common origin with Siberian patterns. In an earlier study (1975a), I reported that the only major element conspicuously missing from the tradition in Taka village was the shaman's escort of the soul of the deceased to the underworld. I have since discovered that this theme, too, has been preserved intact in other valleys.

Shamanistic mythology has it that the speakers of Kham arrived in their present homeland out of the 'North Country.' Though this is likely true, the myths have more to do with *Puran Tsan*, the first mythological shaman, than with the people themselves. Tradition has it that *Puran Tsan* was born in a tree in the north country. Siberian myths speak of the first shaman in the same terms, and the myth is reenacted throughout Siberia as well as in Kham country every time a new shaman is initiated and presented to the community. A pine pole, trimmed of all its branches except the crown, is erected outside the village for the public initiation ceremony. About 12 to 15 feet above ground a small platform is constructed, and it is on this 'nest' that the new shaman is 'born' and accepted into the community of shamans.

Kham myths, as well as Siberian, state that the first shaman was also a blacksmith. He learned his craft from the 'Lord of the Underworld.' Hinduism, of course, puts

blacksmiths into the lower service castes. Siberian shamanism puts a high premium on the technical skills and healing powers of the shaman–blacksmiths. Indeed, the first speakers of Kham may have been metalworkers even in their present homeland[1] – and the name 'Magar,' the tribal name by which Kham speakers identify themselves, may derive from the same etyma as Old Tibetan *mgar-ba* 'smith.'[2] Nowadays, because of modern Hindu influence, other castes, low-caste *Kami* immigrants, do the metal work in Kham villages, and the shamans are smiths only in mythology.

Shamans also have special burial rites not given to laymen. They are buried sitting up and facing north 'to the land of the first shaman.' The upper portion of their bodies are above ground and a stone cairn is erected over them, plastered, and whitewashed. The crown of a pine tree is thrust into the top, giving the shaman access to heaven. This simple burial cairn is the prototype of the Tibetan *stupa* in which the remains of great lamas are said to be housed. The prototype of the Dalai Lama's succession is also found in the shamanistic practices of the Siberian tradition. In Kham villages, after a shaman is chosen by the spirit of his deceased predecessor, he must demonstrate clairvoyant capabilities by finding parts of his predecessor's costume hidden in or around the village. If he fails, his reckless claims to the shamanistic succession are dismissed, and he is declared a charlatan.

Though the shamanism practiced in Kham speaking communities gives no clear indication as to the origin of the people themselves, it does speak to their isolation from the mainstream of commerce and the influx of the major religious philosophies that have been in the region for generations. It is improbable that Siberian shamanism could have reached these southern regions any more successfully than Buddhism has in recent generations. Most probably, the present-day speakers of Kham were practicing Siberian shamanism before they ever entered Nepal, and since their arrival in the Himalayas their language has splintered into three major stocks and arrived at its current state in relative isolation from other TB speaking peoples.

1.4 Language name

Most languages in the Himalayan region are named, at least by outsiders, after the ethnic designation of the people who speak them – i.e. *Sherpa*, for example, is the name of a people, and outsiders refer to their language also as *Sherpa*. Likewise, the Chepang people speak *Chepang*, and Gurungs speak *Gurung*. This is not true for Kham. *Kham*

---

[1] There have long been copper mines in the Char Hajar Parbat and Ath Hajar Parbat regions of Mid-Western Nepal. Casual excavations for new houses and fields in and around most northern Kham villages often reveal beds of copper slag as deep as 10 or 15 feet.

[2] Other words with an original bilabial prefix have been known to survive in some Kham dialects, among them: *pəsi*: < *m-si(y) 'broom,' *psil* < *m-syil 'scrub,' etc. Prefixed *m-* in PTB, however, is usually preserved as a prefixed *p-* in Kham, and one would expect *m-gar > *bəgər*, not *məgər*. The two words may not be related after all.

is the name of a language, or group of languages, spoken by the four northern clans of the Magar tribe: the Budhas, Puns, Ghartis, and Rokhas. The long recognized Magar language is spoken by the southern clans: the Ranas, Thapas, and Ales, i.e. the 'Magars proper.' Several days' walking separates the two groups, and it has been suggested by some that the Northern Magars, the speakers of Kham, are not really Magars at all, but 'originally came of a different stock' (Northey and Morris 1928:189). How both groups came to be called by the same tribal name may never be known. To avoid ambiguity with the Magars who speak Magar, however, I began in 1973 (Watters and Watters) to refer to the people as 'Kham Magars.' Anthropologists have sometimes referred to the same people as 'Northern Magars' (Hitchcock 1967, Fisher 1986, Oppitz 1991).

Kham is known to Nepalis of the region as 'Khamkura,' which, roughly translated, means *Kham-talk* or *Kham-speech.* The word *Kham* itself is of unsure origins and means simply *language* in its broad sense, and *The Language* in its strict sense. In Mid-Western Nepal, where the Kham dialects are spoken, the Nepali use of the word *Kham* or *Khamkura* has the more generalized meaning of a local, non-Nepali dialect. Consequently, at least two other languages in the region, Chantel and Kaike, have received the Nepali appellation *Khamkura*, though neither of them is directly related to the Kham described here. Kham speakers refer to the two languages as Chantel Kham and Tarali Kham, the latter being the same as Kaike, the language spoken in the village of Tarakot.

1.4.1 Early records of Kham as a language name

The earliest direct mention of *Khamkura ('Kamkura')* in the literature is probably that of Professor R. L. Turner (author of the Nepali dictionary) in a book by Northey and Morris entitled *The Gurkhas, their manners, customs and country*, published in 1928. Turner is credited with writing chapter four, *'The people and their languages*,' and makes the following comment in a discussion on the various language families represented in Nepal: 'In addition to the languages which have been provisionally classified as Munda or Tibeto-Burman, there are also many others, of which practically nothing is known but the names, e.g. *Kamkura*, and Rai, with its, at least, ten different dialects' (p. 68, italics added).

Further in the same book, Northey and Morris make the following observation with regard to the division of the Magar tribe into six castes or clans; the Rana, Thapa, Ale, Pun, Burathoki, and Gharti:

> It is probably no exaggeration to state that only the first three named castes are pure Magars, for the latter three do not speak the Magar language and are somewhat different in appearance. The Puns and Burathokis, who live in the high isolated parts of the Magar country, have languages of their own, which differ slightly from valley to valley. These languages have no affinity with Magarkura, and this fact alone is sufficient evidence to prove that they originally came of different stock. (p. 189)

Several years later the same Morris in his book *Handbooks for the Indian Army: Gorkhas* published in 1933, changed his mind on his earlier statement that 'they have languages of their own' and arrived at a new conclusion (based either on hearsay or a whimsical notion of his own) that Kham was only a dialect of Nepali. In a discussion of the major military tribes of western Nepal, the Gurungs and Magars, he writes the following in reference to the three northern clans of the Magar tribe:

> The Burathoki, Gharti, and Pun clans do not speak Magar; nor is it known by any of the Magars living about Argha, Baglung, Dhurkot, Gulmi, Kanchi, Musikot, and Piuthan. Many Puns and Burathokis speak only Nepali, but in some parts they employ a language known as Kamkura. This appears to differ in its pronunciation from valley to valley, and it is said not to be Tibeto-Burman, but a dialect of Nepali. (p. 76)

Unfortunately, it appears that neither Morris nor any of the writers before him managed to record any specimens of Kham, a curious oversight, especially in view of the British interest in the indigenous languages of Nepal. Perhaps more unfortunate was Morris' statement that *Kamkura* is a dialect of Nepali. The error was repeated in numerous later editions of military handbooks, and apparently the matter was never questioned again.

1.4.2 Early reference to Kham speaking peoples

The existence of Kham, at least by negative inference, was apparently known as early as 1819, but owing, perhaps, to the fact that tribes employing Kham were not generally accepted into the British Gurkha regiments in those early years, their language was apparently given little consideration. Hamilton, writing in 1819 in reference to the Bhujel Ghartis, now known to be a Kham speaking clan, states that 'the Ghartis are of two kinds, Khas and Bhujial. The former are admitted to the military dignity; but the latter wallow in all the abominations of the impure Gurungs, and do not speak the [Nepali] language.' He failed, however, to mention what language they did speak.

Three quarters of a century later, records show that the latter class of Gharti was also accepted into the Gurkha regiments after careful screening. Vansittart, writing in 1890, states that:

> by careful selection, excellent Ghartis can be obtained. The Bhujial Gharti lives in the valleys and high mountains to the north of Gulmi, above the Puns. Their tract of country runs along both sides of the Bhuji Khola (river), from which they probably derive their name. The Bhujial Gharti is generally a shepherd. He lives principally on the milk of sheep, and is almost invariably a man of very good physique and heavy limbs. He is remarkably dirty when first enlisted. (pp. 57-58)

In the same publication, Vansittart lists other Magar clans from which recruits had been drawn. Some names are clearly recognizable as Kham village names, while others are the names of village kindreds and lineages. The Gharti clans of *Gamal* and *Walia*, for example, are undoubtedly the Gamales referred to in this volume, along with their

close neighbors, the *Wales* of Wa village. Likewise, the Pun clan of *Takalia* is undoubtedly the Takale of Taka village, the major dialect of this description. Although among present day Takales, only a few belong to the Pun clan, it is generally conceded that Takale Budhas, too, when enlisting in the army, often enlist under the more prestigious name of Pun.

At any rate, no mention was made of Kham in any of the early publications until Turner's observation in 1928 that there existed in Nepal languages like *Kamkura* 'of which practically nothing is known but the names.' Not until John Hitchcock, an American anthropologist, approached the edges of Kham territory on a month's trek sometime in 1960-1962 did it become known that Kham was a TB language. Writing in 1966, he states that the 'Northern Magars … speak a Tibeto-Burman dialect called Khamkura. It resembles Magarkura but the two languages are not mutually intelligible' (Hitchcock 1966:4). If he managed to collect specimens of the language, he apparently never published them.

1.5 Population

In the 1971 census of Nepal, the inhabitants of Rukum, Rolpa, and Baglung Districts speaking 'Magar' as their mother tongue were listed as 27,008. Those speaking 'local district languages' were listed as 13,958. Apart, however, from a small pocket of perhaps 2,000–3,000 Magar speakers in the eastern part of the Baglung District five days removed from Kham speaking territories, Kham is the only non-Nepali language spoken in those districts. It can be assumed, therefore, that apart from about two or three thousand Magar speakers in eastern Baglung District, the total of the two figures 27,008 and 13,958 (40,966) was a fairly accurate representation in 1971 of the number of people whose mother tongue was Kham. Today, their numbers may be closer to 50,000.

1.6 Kham and its dialects

The three major dialects of Kham are Parbate, Gamale, and Sheshi. All three names are exonyms (applied by outsiders), and the speakers themselves, regardless of their dialect, refer to their local language simply as *Kham.* Both Gamale and Sheshi take their names from Gam Khola and Shes Khola (*khola* is a Nepali word for 'river'), though rivers by those names do not actually exist. Rather, the word *khola* in both instances is taken to refer to a region.[3] In fact, both regions are fairly circumscribed with little variation in terms of dialect chains. The name *Parbate*, on the other hand, comes from a Nepali word meaning 'mountain,' or 'belonging to mountainous regions,' and takes in a vast, poorly defined area. Parbate Kham, in other words, is everything that is not Gamale or

[3] In similar manner, *Thak Khola* is an appellation for a particular region along the Kali Gandaki river, the homeland of the Thakhali language.

Sheshi. Because the term is so inclusive and takes in a large dialect continuum, it is useful only as the name of a high level node on a genetic tree. Though all the dialects within Parbate are at least partially intelligible to one another, there are some fairly homogenous groupings corresponding to major population centers in different river valleys. Takale Kham is the prestige dialect of the Parbate group and a *lingua franca* for the whole region. There are other tiny dialects, spoken in single villages, that do not obviously relate to any of the major groups.[4] Figure 1 shows the three major branches of Kham (plus the lower level split in Parbate), along with the names of eleven dialects at the terminal nodes.

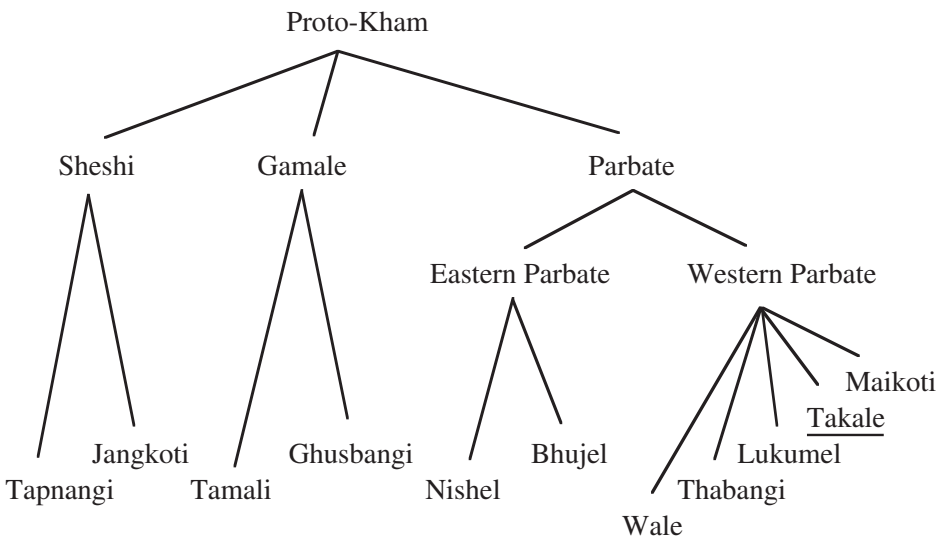


Figure 1. The dialects of Kham

The dialect names at the terminal nodes in figure 1 are, broadly speaking, the names of regional dialects composed of several, more specific, village dialects. In a narrower sense, they are the names of specific village dialects themselves, representative of the larger regional clusters. Thus, for example, Takale is both a village dialect, spoken in Taka village, and a regional dialect composed of several village dialects spoken in the same river valley (see figure 2). Regional dialects dominated by the same mother node are all mutually intelligible, though when situated at opposite ends of a continuum, as are Maikoti and Wale (figure 1), intelligibility may be considerably reduced.

[4] One is Ghusbangi that I have tentatively classified with Gamale, and another Miruli that I have tentatively classified with Mahatale (see map 1).

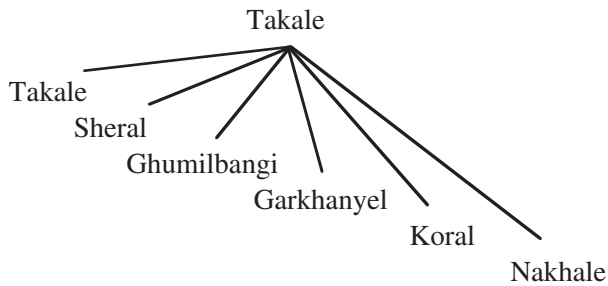


Figure 2. A partial listing of Takale village dialects

1.6.1 Intelligibility levels

At a purely lexical level, based on cognate counts from the Swadesh 100 word list, the similarity between the major branches of Kham is about 85 to 90 percent; enough, it would seem, to insure high levels of intelligibility between them. The count between Takale and Gamale, for example, is 92 percent. Remove loan words and the count goes even higher – up to 96 percent. The striking fact in light of these figures is that the inherent intelligibility levels between the Takale and Gamale dialects should be so low – somewhere in the mid 30 percent range. Intelligibility between Takale and Sheshi is even lower.[5]

High numbers of cognate words between dialects does not always ensure high intelligibility levels. The reasons can be varied and complex. For one, cognate words are not necessarily 'lexically similar.' Any two words in related languages are established as cognate on the basis of diachronic rules of phonological change, not on subjective notions of whether they sound alike or not. Words like Takale *khyo* and Gamale *hlu*, for example, can be established as perfect cognates, both derived from the same proto-form **slo*, but they would not be readily recognized by native speakers as similar. Lexical similarity has to do with recognizability, and so-called 'similarity counts' attempt to make objective the kinds of things a native speaker would resort to unconsciously in deciding if a word were 'similar' or 'different.' Based on lexical similarity, Takale Kham and Gamale Kham have a lexical similarity of just 71 percent, down 21–25 percentage points from pure cognate counts.

1.6.2 Incompatibility of surface morphemes

Other factors contribute even more significantly to reduced intelligibility, the greatest being in the area of morphology. Each of the dialects exhibits a complex morphology with multiple affixes attached to stems of every class. Though the categories marked are much the same from one dialect to another, their surface representations and syntactic

[5] Levels of intelligibility have been measured using the so-called 'Recorded Text Tests.' The technique is described in Casad (1974) and also in Blair (1990).

arrangements are often different. What may be suffixing in one dialect may be prefixing in another, and the morphemes themselves may have derived from entirely different etyma. Worse yet, the morphemes may have derived from etyma once participating in an entirely different system, later reinterpreted to fit the new system. The end result is that the root, which (between Takale and Gamale) has no more than a 71 percent chance of being 'lexically similar,' is surrounded and camouflaged by an array of unfamiliar affixes in unfamiliar syntactic arrangements.

1.7 The place of Kham in Tibeto-Burman

Most major classifications of TB languages in the Himalayan region agree, for the most part, on two major clusters of languages – (1) a Tibetan or Bodish unit that, in addition to Tibetan itself,[6] includes Tamang-Gurung-Thakhali (TGT), and (2) an East Himalayish or Kiranti unit that takes in the so-called 'Rai-Limbu' languages. Everything else gets lumped rather differently. Shafer (1966), for example, sets up a 'West Himalayish' and a 'West-Central Himalayish' between Bodish and East Himalayish, essentially placing all four groups on a par with one another. Nishida's classification (1970), apart from the placement of Tibetan, follows closely. Most other classifications – Benedict (1972), Egerod (1974), and DeLancey (1987), for example – put the languages of Shafer's West Himalayish unit closer to Bodish, and his West-Central Himalayish languages closer to Kiranti.

Kham, unclassified in any of the major classifications until DeLancey (1987), falls into a western division of the Kiranti group along with Vayu, Chepang, and Magar. Benedict was unsure of Magar, and regarded it as a Bodish-Bahing link. With the addition of Kham to Western Kiranti, the whole picture becomes much more clear. As I will show in chapter 17, there is little doubt that Kham and the Eastern Himalayan languages share a common, high level node. At the same time, Kham shares a considerable amount of innovative vocabulary with Magar – the kinds of items that do not show up on basic vocabulary lists. As such, Kham turns out to be an important link between the East Himalayish/Kiranti languages and the loose grouping of languages in West-Central Nepal.[7]

Using two different comparative lists – the Swadesh 100 word list and Matisoff's CALMSEA list[8] – both of which utilize basic vocabulary fairly resistant to change, I arrive at comparable results for Kham's relationship to other languages. Table 1 compares

---

[6] Nishida (1970) is an exception, who has a Tibetan group in opposition to everything else that he calls 'Himalayan.'

[7] Shafer (1966:142) lumped these languages together into West-Central Himalayish because, he says, 'they have more in common with each other than with any other language or group of languages.'

[8] An acronym for 'Culturally Appropriate Lexicostatistical Model for SouthEast Asia' (Matisoff 1978). See Vocabulary, chapter 19.

the two lists (both are slightly expanded to include, in some cases, synonymous Kham terms for a single original item).

Table 1. Cognate percentages between Kham and other TB languages

	Swadesh list 116 items (%)	Matisoff list 230 items (%)
Modern Kham to:		
PTB	59	56
Chepang	44	33
Magar	38	33
Kaike	26	17
Thakhali	25	16
Tamang	18	10
Tibetan	12	15
Khaling	9	5
Sunwar	5	3

The two columns in table 1 yield similar results: 59 percent of the Kham words on the Swadesh list are relatable to PTB, and 56 percent on Matisoff's list are. Chepang and Magar show a considerably higher number of cognates with Kham than any other languages in Nepal, about 40 percent on the Swadesh list. The Tamang-Gurung-Thakhali (TGT) group, which includes Kaike, also has a fairly high number of cognates, from 18 to 26 percent. The Kiranti languages, represented by Khaling and Sunwar, are the lowest at less than 10 percent. (See Vocabulary in chapter 19.)

Although Kham exhibits low lexical similarity with the East Himalayish/Kiranti languages, it does show remarkable correspondences in the morphology and organization of verbal paradigms, including the retention of an archaic set of person agreement forms (see chapter 17). This suggests that Kham, along with Magar and Chepang, can be regarded as distant cousins of Kiranti belonging to a western branch.[9] Slightly revising some of the earlier TB classifications, then, I would add Kham, along with Magar, to Benedict's Vayu-Chepang division (as DeLancey's Kham-Magar has already done). In addition, Kham likely forms a link to Bodish (Tibetan-Kanauri) along the lines suggested by Benedict for Magar. My revision is in figure 3.

[9] Basing genetic relationships on reconstructed morphological patterns is a debatable issue. My proposal is strengthened considerably by the presence of archaic cognate forms within those patterns.

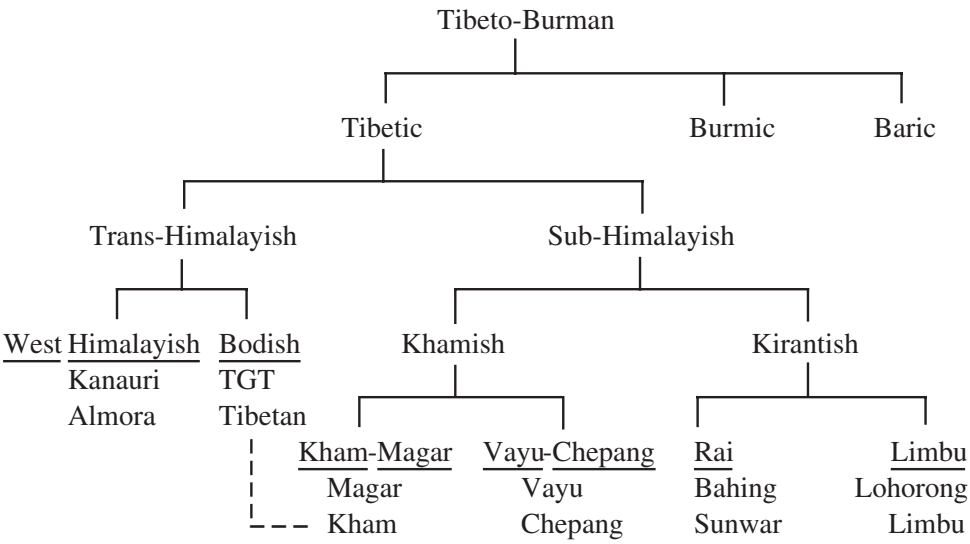


Figure 3. The relationship of the Tibetic languages within TB

Kham, then, lends cohesiveness to the loosely knit Magar-Chepang-Vayu group (which I will call Khamish because of Kham's key position) and binds it to Kiranti (Rai-Limbu). At the same time it forms a link between the new Khamish group and Bodish through Tibetan. In essence, then, Kham links the two major branches of Tibetic – what I will call Trans-Himalayish (Egerod's Bodish-Himalayish and DeLancey's Bodish) and Sub-Himalayish (Egerod's Kirantish and DeLancey's East Himalayish).

2 *Segmental phonology*

In chapters 2 and 3, I will address two major topics in Kham phonology: (1) the inventory of segmental phonemes with their major allophonic variants, and (2) tonal phenomena. In the first case I will address segmental phonology with only enough detail to settle on a reliable and working orthography for the remainder of the chapters. This will necessarily include some discussion of syllable structure as it relates to allophony and morphophonemics. In chapter 3, I will discuss Kham tonal phenomena within the larger framework of Himalayan and Southeast Asian tone systems. Kham tones appear to be unique in a number of areas, which, I suspect, will turn out to be more common than our current knowledge of Himalayan tone systems permits us to recognize. It is hoped that a fairly detailed treatment of Kham tone will be another step in helping define the features of a Himalayan tone typology.

2.1 Consonants

The consonantal inventory in Kham is simple, with twenty-two consonants occurring at only three points of articulation – bilabial, alveolar, and velar (ignoring the consonant /h/). Palatal consonants, including the affricates [tʃ] and [dʒ], occur only as allophonic variants of the alveolar series. All consonants occur in the onset of a syllable, while a more restricted inventory occurs in the coda – little more than half the total, twelve consonants. Table 2 gives the full inventory of consonant phonemes in Takale Kham, using orthographic symbols that will be used throughout the rest of the book.

In terms of distribution, none of the voiced obstruents and none of the aspirate series occur in syllable-final position. Furthermore, of the approximants, only /l/ and /r/ occur syllable finally.[1] The affricate /c/ in final position is rare, being limited to a few loan words from Nepali.

Syllable-initial consonant clusters, apart from clusters with /y/ or /w/, are disallowed. All but /ŋ/ cluster with /y/, and all but /p/ cluster with /w/. Some of the clusters are limited to onomatopoeic words, in keeping with a general cross-linguistic observation that onomatopoeic words often overstep the norms.

[1] Orthographically, *h* occurs syllable finally, but in all cases it indicates a preceding vowel with lax phonation. Phonation types will be discussed in detail in chapter 3 on Tone.

Table 2. Consonant phonemes in Takale Kham

	labial		alveolar		velar		glottal
	–voice	+voice	–voice	+voice	–voice	+voice	
stops	p	b	t	d	k	g	
+ aspirated	ph		th		kh		
fricatives			s	z			
affricates			c	j			
+ aspirated			ch				
nasals		m		n		ŋ	
approximants:							h
liquids				l r			
glides				y		w	

2.1.1 *Stop consonants*

Voiceless stop consonants occur in syllable onset and coda positions with virtually any vowel combination. Syllable-final /t/ and /k/, however, are rare in Takale Kham. Where they do occur, they occur primarily in new Nepali loans[2] or onomatopoeic words. Final /t/ and /k/ occur extensively in Proto-Kham, and they are preserved in varying degrees in dialects other than Takale. Nishel is a good source for reconstructing final /t/ and Lukumel is a good source for final /k/. Also, both are preserved in Sheshi verbs, given the right morphological contexts.

In Takale, the loss of either one results in compensatory lengthening (plus other vocalic reflexes which will be discussed in §2.2.5 and §2.2.6):

(1)

Sheshi	Nishel	Takale	gloss
-sit-	sit-	si:-	'to sweep'
-sik-	si:-	'sí:-	'to step on'

The voiced stops /b/, /d/, and /g/ are restricted to syllable onsets.[3]

[2] What are presumably older loans have adjusted to Kham phonology by adding a vocalic release, as in *cetə* < *cet* 'memory,' and *buhtə* < *bhut* 'ghost.'

[3] I am aware of one case of /b/ occurring in coda position – *nehblo* 'two.' Evidence suggests that *blo* was an early classifier (see §9.4.1). Occurring between voiced elements with lax phonation on the first, the voicing on /b/ has remained. Also, in some Nepali loans with geminate consonants a voiced coda is possible, as in *giddə* 'vulture.'

2.1.2 Stop variants

Allophonic variation in stops is relatively minor. An allophone of /k/, for example, that occurs preceding the front vowels /i/ and /e/ is more fronted than one that occurs preceding back vowels. The differences, however, present no problems for interpretation and I will discuss it no further.[4] Also, stop consonants in the coda of a syllable are unreleased.

Voiceless stops occurring in the coda of VC and CVC syllables become voiced when followed by the nominalizing suffix *-o*. In such cases, the resulting sequence of CVC-V is resyllabified as CV-CV, as in:

(2)

a.	təp-	'to drum'	/tə.bo/	'drummer'
b.	cep-	'to ride'	/ce.bo/	'rider'
c.	dup-	'to gather'	/du.bo/	'gatherer'

2.1.3 Fricatives and affricates (sibilants)

There are two fricatives – /s/ and /z/ – and three affricates – /c/, /j/, and an aspirate counterpart for /c/, namely /ch/. Only /s/ and /c/ occur syllable finally. Takale is one of the few dialects that has preserved final /s/.

With fricatives and affricates there is a distinct allophonic variation anywhere that a front vowel follows – the consonant becomes palatalized, as in the following contrastive sets.

(3) INITIAL /s/:

a. (followed by front vowels)

/si-/	[ʃi]	'to die'
/sül-/	[ʃyl]	'to be slippery'
/syul/	[ʃjul]	'footprint'
/sya:/	[ʃjæ]	'meat'

b. (followed by non-front vowels)

/sa:/	[sa:]	'strength'
/sɨ:/	[sɯ:]	'breath'
/su:/	[su:]	'who?'
/so-/	[so]	'to rise'

The consonant /s/ is unaffected by a preceding front vowel:

---

[4] In Gamale, on the other hand, allophones of /t/, /th/, and /d/ following front vowels are pronounced with the flat of the tongue, and some are beginning to collapse with /k/, /kh/, and /g/ in the same environment, as in: *ge:* 'we (pl)' and *ge:* 'you (pl)' < **dʒe:*.

(4) FINAL /s/:

/pis-/	[pis]	'to be firm fitting'
/pos/	[pos]	'beer malt'

Recall that syllable codas with voiceless stops become resyllabified and voiced when followed by the nominalizer *-o*, as in *təp-* 'to drum' > *təb-o* /tə.bo/ 'drummer.' The generalization, however, does not hold for voiceless sibilants, as in *kəs-* 'to be tight' > /kə.so/ 'tight' – but not */kə.zo/.

Like the palatal variants of /s/ to [ʃ] shown in (3a), so also /z/, /c/, /ch/, and /j/ have palatalized variants [ʒ], [tʃ], [tʃʰ], and [dʒ], respectively, in identical environments.

2.1.4 Nasals

There are three nasal consonants – /m/, /n/, and /ŋ/. Though all three occur in syllable-initial and syllable-final positions, there is a general tendency in Takale Kham for final **-n* and **-ŋ* to erode, being replaced by nasalization on the preceding vowel, and in some cases, modification of the vowel itself. This is shown in table 3.

Table 3. Final **-ŋ* and **-n* in three Kham dialects

	Takale	Mahatale	Sheshi	gloss
a. proto **-ŋ*	sĩ:	siŋ	siŋ	'tree, wood'
	cõ:-	cuŋ-	cuŋ-	'alight, perch'
	hã:	haŋ	haŋ	'cliff'
b. proto **-n*	chĩ:-	chen-	chen-	'to lift'
	jẽ:h-	jehn-	jehn-	'high'
	phĩ:-	pin-	pin-	'boil food'

2.1.5 Approximants

The approximant class is comprised of what have sometimes been called 'frictionless continuants.' It includes the liquids /l/ and /r/, and the glides /y/ and /w/, as well as /h/. The consonant /h/ has no particular vocalic shape. It is the voiceless equivalent of the vowel that follows. I classify it as a consonant because it never occurs in the nucleus of a syllable. Following are examples:

(5) INITIAL /h/:

/he:/	'parched grain'	/ha:/	'tooth'
/ho:/	'that'		

Final orthographic *h* does not occur as a consonant, but only as a marker of lax vowels, as in *ka:h* 'dog' (see footnote 1).

i. Liquids

The liquids /r/ and /l/ both occur in Kham. The rhotic liquid /r/ is a flapped consonant [ɾ] word initially, and trilled [r] word finally. The lateral /l/ approximates an alveolar laterally released affricate [dl] ~ [dɮ] word finally. There are no consonant clusters with /r/ or /l/ as the second member. Where clusters in /r/ are borrowed from Nepali, they are resyllabified in Kham with an epenthetic /ə/, as in /jã:hkə̲ri/ < Nepali *jhākri* 'shaman.'

ii. Glides

The glides /y/ and /w/ occur both syllable initially and as the second member in complex onsets like /py/, /kw/, etc. The glide /y/ is phonetically equivalent to [i] or [e], and /w/ is equivalent to [u] or [o]. Here I present the two glides /y/ and /w/ as consonantal onsets. I will treat them as clusters in §2.4 on the syllable.

(6) INITIAL /y/:

/yem/	'trail, road'
/ya-/	'to give'
/yo:-/	'to sell'

(7) INITIAL /w/:

/wi: phwi:/	'worry'
/wa/	'about'
/wohr/	'right hand'

2.1.6 Aspiration

Aspiration occurs with stops and affricates, but only with the voiceless series /ph/, /th/, /kh/, and /ch/. Aspiration is further restricted to syllable onsets. The absence of a voiced aspirate series /bh/, /dh/, /gh/, and /jh/ is partly a matter of interpretation. This will become more clear in chapter 3 on Tone. For now, it is enough to know that 'lax phonation,' a breathy laryngeal quality of the vowel, occurs with voiceless onsets as well as voiced ones. In syllables with *voiceless* onsets, the difference between aspiration and lax phonation on the vowel is qualitative and easy to perceive, as in the following three-way contrast:

(8) a. NORMAL (modal) PHONATION:
/pi:-/ 'to milk'

b. LAX PHONATION:
/pe:h/ 'bamboo box'

c. ASPIRATION:
/phe:/ 'ball of string'

In syllables with *voiced* onsets, on the other hand, only a two-way contrast exists – a

contrast between normal phonation, as in /be:/ 'skirt,' and *something else* that could, without further insights, be interpreted either as lax phonation *or* aspiration:

(9) a. ASPIRATION?:
/bhe:/ 'basket'

b. LAX PHONATION?:
/be:h/ 'basket'

To the uninitiated ear, the syllable represented in (9) sounds very much like the voiced aspirates prevalent in Nepali and other Indic languages.[5] Given the fact, however, that lax phonation occurs in syllables with almost any onset (voiced, voiceless, or no consonantal onset at all), and that all such syllables display identical tonal properties (as opposed to aspirated syllables which have different tonal properties), it becomes clear that such syllables should be interpreted as lax phonation types. Phonologically, then, voiced aspirates like /bhe:/ or /jha:/ do not occur in Kham.

Aspiration on voiceless obstruents is heavy in Kham, with a considerably stronger burst of air than that which occurs in English.

2.2 Vowels

There are nine vowels in Kham, six of which are basic, and three of which can be traced to secondary developments resulting from the loss of consonants. The six basic vowels are two front, /i/ and /e/, two central, /ə/ and /a/, and two back, /u/ and /o/. The vowels /ü/, /ö/, and /ɨ/ are recent innovations and do not occur in all dialects (see §2.2.4 and §2.2.5).

Table 4. Vowel phonemes in Takale Kham

	front		central	back	
	–round	+round		–round	+round
high	i	ü		ɨ	u
mid	e	ö	ə		o
low			a		

The six basic vowels are close to the cardinal values associated with those symbols and require no further explanation apart from allophonic variants which I will present shortly. The two front vowels that I write as /ü/ and /ö/ are both rounded and close to the IPA representation [y] and [ø], respectively. The back vowel that I represent as /ɨ/ is phonetically a high-back unrounded vowel and close to the IPA representation [ɯ].

[5] Indeed, in the Devanagri spelling system I devised for Kham, I found strong reader support in writing these three stops as भ 'bh,' ध 'dh,' and घ 'gh.'

Table 4 gives the full inventory in the Takale dialect using orthographic symbols that will be used throughout the book.

2.2.1 *Contrastive length*

All vowels are contrastive for length. Orthographically, long vowels are marked by a colon immediately following the vowel symbol – /i:/, /e:/, /ə:/, /a:/, /u:/, /o:/, /ü:/, /ö:/, and /ɨ:/. Following are contrastive sets:

(10)		SHORT:		LONG:	
	a.	pi-	'to suck'	pi:-	'to milk'
		pa-	'to fall'	pa:-	'to break'
	b.	te-	'to drop'	te:-	'to press down'
		tu-	'to be poisoned'	tu:-	'to be spicy'
	c.	ki-	'to pluck'	ki:-	'to plow'
		ko-	'to crow'	ko:-	'to peel, skin'

The short/long contrast is found primarily on the five vowels i, e, a, u, and o. The vowel /ə/ is almost always short, though there are a few examples of long /ə:/, as in:

(11)			
	a.	kə:	'a lid'
	b.	mələ:	'weevil'
	c.	thə:rəi-	'to determine'

Likewise, the vowel /ɨ/ occurs almost exclusively as a long vowel. Recall that the vowel comes from the loss of final *-k*, and that the loss of a final consonant normally results in compensatory lengthening.

The secondary vowels /ü/ and /ö/ are primarily short since they occur in closed syllables like *sür-* 'sour,' *sül-* 'slippery,' and *süs* 'disciple,' etc.

i. Length and stress in polysyllabic morphemes

In polysyllabic morphemes, all syllables can be short, but not all can be long. If there is a long syllable somewhere in the string it will usually be the final syllable, though there are some exceptions. Furthermore, stress will fall on that final, long syllable (though this is not to be equated with high pitch). In a bisyllabic morpheme in which both syllables are open and short there is no discernible stress placement, as in:

(12)			
	a.	kata	'what?'
	b.	baza	'bird'
	c.	gohga	'corn'

In trisyllabic morphemes, on the other hand, in which all syllables are open and short, there is discernible stress on the final syllable (in spite of its shortness),[6] as in: