THE CHEPANG LANGUAGE: PHONOLOGY, NOMINAL AND VERBAL
MORPHOLOGY – SYNCHRONY AND DIACHRONY OF THE VARIETIES OF THE
LOTHAR AND MANAHARI RIVERS

by

MARIE-CAROLINE LYDA PONS

A DISSERTATION

Presented to the Department of Linguistics
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

September 2022

DISSERTATION APPROVAL PAGE

Student: Marie-Caroline Lyda Pons

Title: The Chepang Language: Phonology, Nominal and Verbal Morphology – Synchrony and Diachrony of the Varieties of the Lothar and Manahari Rivers

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Linguistics by:

Scott DeLancey	Chairperson
Spike Gildea	Core Member
Doris Payne	Core Member
Zhuo Jing-Schmidt	Institutional Representative

and

Krista Chronister	Vice Provost for Graduate Studies

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded September 2022

DISSERTATION ABSTRACT

Marie-Caroline Lyda Pons
Doctor of Philosophy
Department of Linguistics
September 2022

Title: The Chepang Language: Phonology, Nominal and Verbal Morphology –
    Synchrony and Diachrony of the Varieties of the Lothar and Manahari Rivers

Chepang is a Trans-Himalayan (a.k.a. Tibeto-Burman or Sino-Tibetan) language mainly spoken in 4 districts of Nepal: Makawanpur, Chitwan, Dhading and Gorkha. Around 48,000 speakers declared speaking the language in 2011 (Nepal CBS). Chepang is an oral language, which means that the community has not yet developed a written tradition.

This dissertation offers in-depth and comprehensive descriptions of many aspects of the language structures and valuable insights on their variation and historical developments. It includes an analysis of phonetics and phonology, drawing a new understanding of the positions and realizations of the laryngeal features at the source of the development of tonal distinctions. Chepang's nominal morphology shows typologically well attested word formation processes combining nouns, nouns and verbs, compounds and nouns, and nouns and derivational morphology. A detailed description of kinship terms is provided, specifying how people use them. The description of verbal morphology provides a clear understanding of the type of morphology and constructions associated with verbs, covering all typical contextual settings where such constructions

can occur. The complex non-canonical direct-inverse system of Chepang shows further pragmatic or epistemic motivations behind the use of a construction over another. When possible, historical explanations for the origin of certain forms or constructions are provided, allowing a better understanding of the verbal system in its entirety. Throughout the dissertation, the examples provided in Chepang were carefully chosen to illustrate the daily life, culture, spiritual beliefs, history, and other knowledge that pertain to the Chepang people.

Such description contributes to a better understanding of Trans-Himalayan typology and more broadly linguistics typology. It provides a solid foundation for further historical linguistics research on the reconstruction of Proto-Chepang (ancestor language) and on the nature of its relationship to other Trans-Himalayan languages at lower and higher-level clades of the family.

It offers two comprehensive orthographies based on Devanāgarī and Roman alphabet to the Chepang community. Part of the recordings is being made accessible to the public in an online accessible format. This project aimed at participating in raising interest and awareness to promote the preservation of the language and the culture it conveys.

CURRICULUM VITAE

NAME OF AUTHOR: Marie-Caroline Lyda Pons

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

University of Paris IV Sorbonne, Paris

DEGREES AWARDED:

Doctor of Philosophy, Linguistics, 2022, University of Oregon

Master of Arts, Linguistics, 2013, University of Paris IV Sorbonne

Bachelor, 2010, Linguistics, University of Paris IV Sorbonne

AREAS OF SPECIAL INTEREST:

Typology, Historical and Comparative Linguistics, Descriptive Linguistics, Morphosyntax, Discourse, Pragmatics, Semantics, Trans-Himalayan Languages

PROFESSIONAL EXPERIENCE:

Graduate Teaching Fellow, Department of Linguistics, University of Oregon,]
    Fall 2014 - Spring 2021

GRANTS, AWARDS, AND HONORS:

Graduate Student Research Award, Department of Linguistics, University of
    Oregon, 2021

Dissertation Improvement Grant, National Science Foundation NSF-DEL
    (Documenting Endangered Languages), 2019

Oregon Humanities Center Graduate Research Support Fellowship, University of
    Oregon, 2019

Special "Opps" Travel and Research Awards, University of Oregon, 2019

College of Art and Science Continuing Student Scholarships and Fellowships, University of Oregon, 2019

Global Oregon Graduate Research Award, University of Oregon, 2018, 2015

Small Professional Grant, Center for Asian and Pacific Studies, University of Oregon, 2017

Firebird Foundation for Anthropological Research Fellowship in Oral Literature, 2016

Center for the Study of Women in Society Research Grant, University of Oregon, 2016

PUBLICATIONS:

Pons, Marie-Caroline. 2021. On the origin of 2nd person prefix *#tV-* in Trans-Himalayan languages. *Linguistics of the Tibeto-Burman Area* 44:2, 226-263. https://doi.org/10.1075/ltba.20004.pon

Pons, Marie-Caroline. 2021. Review: The Dura language: Grammar and phylogeny. *Himalayan Linguistics* 20:1, 116-132. http://dx.doi.org/10.5070/H920155279

Pons, Marie-Caroline. 2014. Valence and voice strategies in Magar. *Nepalese Linguistics* 29, 101-108.

Chepang, Buddi Raj Chepang, Bir Bahadur Chepang, Pratap Singh Chepang, Deu Bahadur Chepang, Bhim Bahadur Chepang, Jit Bahadur Chepang, Kalisman Chepang, Yani Maya Chepang, Bhim Bahadur Chepang, Shanta Ram Praja, Samjana Chepang, Rabi Lal Chepang, Sani Maya Chepang, Jal Maya Chepang, and Rasila Chepang.

Amongst the writers of the story of this journey, I am forever grateful to my advisor Scott DeLancey who gave me the confidence to grow as a linguist. Scott, thank you so much for your support, your vision and scholarship, and for always reminding me to chill out. I have been so lucky to be surrounded by some of the best people in this world. Thank you to all of you. To my dissertation committee members, Spike Gildea, Doris Payne, and Jing Zhuo-Schmidt, and the UO Linguistics Department professors who helped me sharpen my linguistic analytical skills, Melissa Baese-Berk, Vsevolod Kapatsinski, Tyler Kendall, Eric Pederson, Tom Payne, and Lisa Redford; to my UO fellow linguists for their friendship and support: Krishna Boro, Charlie Farrington, Richard Griscom, Xuan Guan, Kaylynn Gunter, Jaeci Hall, Misaki Kato, Ellen Kress, Dae-yong Lee, Paul Olejarczuk, Manuel Otero, Jaime Peña, Sara Pacchiarotti, Becky Paterson, Shahar Shirtz, Matt Stave, Allison Taylor-Adams, Amos Teo, Zoe Tribur, and Julia Trippe; to the many Trans-Himalayan linguists who have inspired me: Oliver Bond, David Bradley, George van Driem, Carole Genetti, David Hargreaves, Kristine Hildebrandt, Nathan Hill, Gwendolyn Hyslop, Guillaume Jacques, Linda Konnerth, Aimée Lahaussois, Martine Mazaudon, Boyd Michailovski, Pavel Ozerov, Narayan Sharma, Camille Simon, Mark Turin, Manuel Widmer and Marius Zemp.

You are my home everywhere I go, my friends and family who give me incredible love and support, who put up with me, and make me a better person, I love you.

To the Chepang people

TABLE OF CONTENTS

Chapter   Page

13

LIST OF TABLES

23

LIST OF FIGURES

LIST OF MAPS

CHAPTER I

INTRODUCTION

This dissertation is the first part in the development of a grammar of Chepang, as spoken in the basins of the Lothar and Manahari Rivers, in the hills of the districts of Chitwan and Makawanpur, in Central Nepal.

The present work builds on four years of research in close collaboration with Chepang community members. While primarily focused on providing a good understanding of the linguistic structures of Chepang as spoken in the Lothar and Manahari River basins, I seek to give a representation of language variation and change; it is hence based on a digital language corpus that comprises 25 hours of audio and video recordings (APPENDIX II) collected with around 100 people from 28 villages of different Municipalities and Rural Municipalities: Rapti (wards 3, 5, 6, 11, 13) in the district of Chitwan; Raksirang (wards 6, 7, 8), and Manahari (wards 4, 7) in the district of Makawanpur (APPENDIX I).

This descriptive project aims at expanding our knowledge of Chepang from earlier studies mainly conducted by Ross Caughley who provided a description of the phonology (1969), verbal morphology (1982), and a dictionary (2000; 2016) while mainly focusing his research on the variety spoken in the village of Maisirang, in the Rural Municipality of Raksirang (ward 8) in the district of Makawanpur.

In this dissertation, I tried to offer thorough accounts of the described linguistic structures through both synchronic and diachronic analyses, in addition to providing two accurate and practical writing systems to the Chepang community, one based on Devanāgarī, and the other on the Roman alphabet (§ 2.11).

The present chapter is organized as follows. I introduce the Chepang language, and its speakers in § 1.1. I present the first two Nepalese official documents that mention the Chepang people in § 1.2. A review of earlier research on the Chepang language and people is given in § 1.3. In § 1.4, I present the different terms used to refer to the Chepang language and people and discuss linguistically misconceived etymologies which unfortunately became very strongly attached to these, bringing their load of negative

social consequences for the Chepang community. In § 1.5, I briefly present the origin of the use of the name Praja which, like the name Chepang, is used to talk about the people and the language. In § 1.6, I show that earlier ethnographical accounts of the Chepang people by Hodgson (1848; 1857a; 1874a) in particular, which depict the community as primitive hunter-gatherer nomadic people who live outside a system in the first half of the 19th century, are highly contested by today's Chepang community members. I discuss this narrative considering community members' accounts of their personal and community history at the local level and Chepang vocabulary, which indeed comprises native terminology related for instance to agriculture. In 1.7, I give an overview of the endangerment status of the Chepang language, and the attitude lately adopted by community members, especially those that are politically engaged, towards the preservation and revitalization of their language. I briefly introduce in § 1.8, the current views on the internal branching of the Trans-Himalayan (TH) language family and discuss earlier phylogenetic classifications of Chepang, suggesting three proposals that should be further investigated through a thorough comparison of phonological and morphosyntactical structures at a lower level. In § 1.9, I describe the language variation observed and propose three main language varieties: Lothar, Manahari, Rapti, and Handikhola. I clarify the present approach and terminology used to describing language variation throughout the grammar. In § 1.10, I give a brief historical account of language policy in Nepal and an overview of the historical and current language contact situation between Chepang and other Trans-Himalayan (TH) and Indo-Aryan (IA) languages, in particular Nepali (IA), the lingua franca of Nepal, and Tamang (TH). In § 1.11, I describe the collaborative dynamic of this project, the language data that have built the digital corpus on which the present analysis of the language is based, and our efforts towards providing the Chepang community members with an access to the collected documentation of their language. Finally, I describe the organization of the dissertation in § 1.12.

1.1. The Chepang language and its speakers

Chepang is a Trans-Himalayan (TH) language (a.k.a. Tibeto-Burman or Sino-Tibetan), spoken by the Chepang people. The phylogenetic position of Chepang within the family of TH languages is discussed in § 1.8.

The number of speakers was estimated to around 47,000 in 2011, the last Nepal Census to date (Central Bureau of Statistics, National Planning Commission Secretariat 2012). This corresponds to 72% of the Chepang population which by then, counted around 65,000 people. In 2011, the Chepang population represented 0.24% of the total population of Nepal.

Based on contemporary and historical accounts of the Chepang community's familial or local history, passed on through generations, it is likely that the Chepangs have inhabited the hills of the Mahabharat Range since the 17th century, and in particular an area that covers Northeast Chitwan, Northwest Makawanpur and Southwest Dhading. This is discussed in § 1.2 and § 1.6.

Previous literature attempted to situate the original location of the Chepangs, and the hills of the Mahabharat Range was first suggested by Hodgson (1848) and later by Jest (1966). In his 5th edition of *People of Nepal*, Bista (1987: 98) mentions the area of Sunathali in the district of Dolakha in the East of Nepal, inhabited by the Limbu, as an original place of some Chepang people: "Some of the Chepangs believe that their community is an offshoot of the Kiranti (Rai-Limbu) group that inhabits Sunathali, in Dolkha in the east." Bista does not mention further detail on this matter, nor does he mention the place where he collected such information. This information may relate to a myth of origin that is sometimes told by Chepang community members along with another one that similarly situates the origin of the Chepangs in another area of Dolakha district, i.e., Pukunthali (§ 1.4). Both areas correspond to places considered spiritually sacred and dedicated to the divinity of Bhumi, deity of Earth, often invoked in shamanic chants. It is also possible that some Chepang communities of some villages in the area migrated from the East in the second half of the 19th century, as is attested in other places. For instance, the settlement of Polkim in Rapti-13, consists of late migrations of Thakuri community members who arrived from the district of Morang in Eastern Nepal. This

migration goes back to six generations according to Polkim community members, that is, 150 years ago (25 years per generation), i.e., 1870. In the case of Polkim, Thakuri men speaking Nepali married Chepang women, and Chepang was preserved as the language spoken by their descendants.

In Table 1, I present the number of Chepang people who declared speaking their native language in 2011 (Central Bureau of Statistics, National Planning Commission Secretariat 2012) in the four districts where it is mainly spoken (Chitwan, Makawanpur, Dhading, and Gorkha). Chitwan shows the highest number of speakers, with more than 20,000, but the district of Makawanpur is where the language is the most preserved, or where language transmission is the most sustained, with 84% of the Chepang population speaking it.

Table 1. Chepang speakers in Chitwan, Makawanpur, Dhading, and Gorkha

Nepal CBS 2011	Chepang speakers	Chepang ethnicity	speakers' percentage
Chitwan	21,469	28,655	74.92 %
Makawanpur	16,004	19,038	84.06 %
Dhading	8,437	14,476	58.28 %
Gorkha	1,625	3,446	47.16 %
Total	**47,535**	**65,615**	**72.45 %**

Map 1 represents the number of Chepang speakers per district where the Chepang population was reported living in the 2011 Nepal Census (Central Bureau of Statistics, National Planning Commission Secretariat 2012). The highest number of speakers is concentrated in the districts of Chitwan and Makawanpur, the principal areas where Chepang language varieties have been investigated for this project (§ 1.9.1).

Map 1. Chepang speakers in Nepal
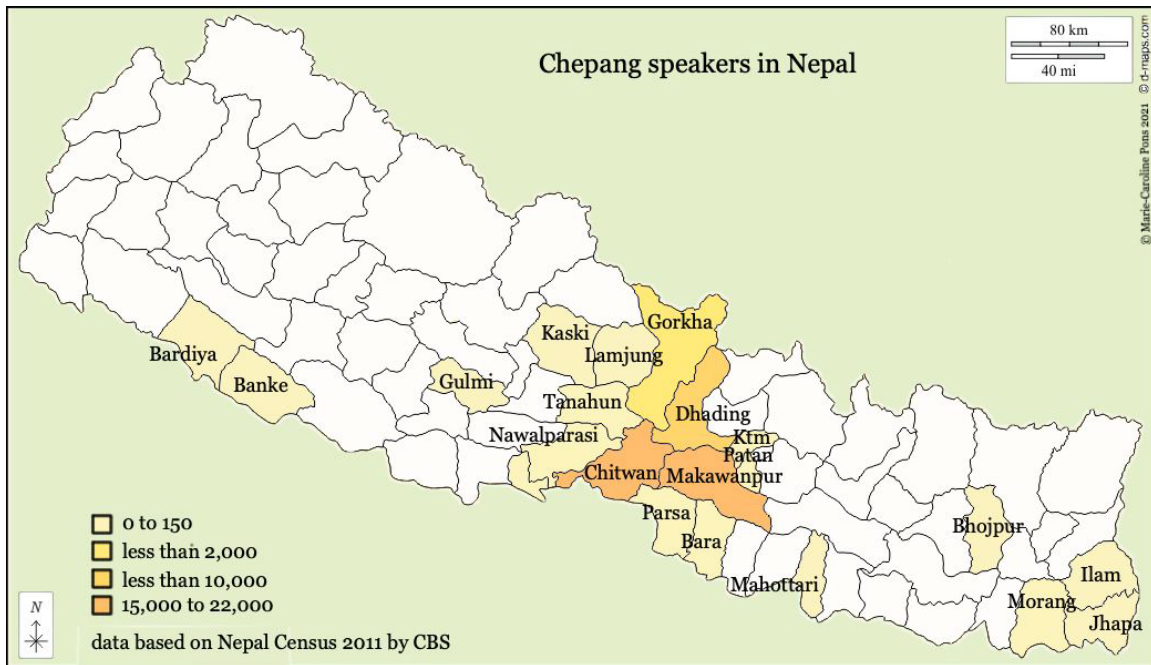
The coming Nepal Census was held from November 11th to 25th 2021 and likely will be released in 2022-23. It will tell us how different the socio-linguistic situation amongst the Chepang community is today by contrast to ten years ago.

1.2.   First mentions of the Chepangs in Nepalese official documents

Jest (1966: 171) noted that no official historical Nepalese document had ever been found mentioning the existence of the Chepang people. In fact, two documents mention the Chepangs, one as early as 1776 by King Pratap Singh Shah, and the other in 1847 by Prime Minister Jung Bahadur Rana.

The first document, dated from 1833 vs[1] (1776 CE[2]) is an order from King Pratap Singh Shah, son of King Prithivi Narayan Shah, who reigned for two years after the death of his father, from 1775 to 1777. This order serves his interests and that of the Kingdom

---

[1] Vikram Samvat

[2] Common Era

of Nepal through those of the Gosai Saints who were important Indian Kashmiri traders travelling between India, Nepal, and Tibet since the first half of the 18th century (Dahal & Bhattarai 2019). At the time of King Prithivi Narayan Shah who ruled the Kingdom of Nepal from 1768 to 1775, the Gosai Saints had already been given facility of customs for the King to gain their favor and maintain the control of the India-Nepal-Tibet trade (Dahal & Bhattarai 2019). In this document, King Pratap Singh Shah orders tax collectors (called Umrao, Dware and Amalidar, who are government officers or representatives) to take 1 ana (6.5 penny Nepali rupee[3]) from each household that belongs to the following 14 communities: Manjhi, Kumal, Darai, Danuwar, Newar-Kumal, Kusale, Tharu, Jwalaha, Pahari, Kusahari, Thami, Hayu, Sunuwar and Chepang. This tax would then be reversed to the Gosai Saints, and if it happened that they would reach the villages of such communities, they had to be given food and shelter. In this order, King Pratap Singh Shah further warns these communities that actions would be taken in case they obstruct the tax collection process. Through this offer, at the death of a Gosai Saint, in case of absence of descendant, their land would become the property of the Kingdom and could be used by these communities as well for cultivation. An image of this order written in 18th century Nepali (not translated) is given in Figure 1, as published by Dinesh Raj Patna (1968: 24–25).

---

[3]25 penny = 4 ana; 1 ana = 6.25 penny; 100 penny = 1 rupee

Figure 1. First 1776 Nepalese official document mentioning the Chepangs

श्री ५ प्रतापसिंह शाहले भगवन्तनाथलाई
वि. सं. १८३३ मा लेखेको पत्र

(८ संख्या)

स्वस्तिश्री.गिरिराजचक्रचूडामणिनरनारायणेत्यादिविविधविरुदावलिविराजमानमानोन्नतश्रीमन्महा-
राजाधिराजश्रीश्रीश्रीमहाराजेप्रतापसिंहसाहबहादूरसमसेजङ्गदेवानां सदा समरविजयिनाम्

आगे श्रीगोसाइ भगवंतनाथजिके हाम्रो मुलुकभरिको जोगिको मंडलाइ चहाँइउं. हाम्रा
मुलुकभरिका उमरा¹ द्वारा अमालिदार सबैले माझिकुह्याल दरबं दनुवार नेवारकुह्याल
कुसल्या थाप ज्वलाहा पहरि कुसहरि थामि हायु सुनुवार चेपंग एति जातका घरहि² एक एक्
आना दस्तुर लिनु. साज बिहान षान दिनु. जोगिका षतहित दंडकुंड मोरोअपुतालि³ महाषत-
छित जोगिको टिको अम्बलिको एस हिसाबले अमालिदारले तिराइदिनु. बितलपक⁴ कुरियामाहा⁵
पनि तिराइदिनु. जो छेक्थुन् गर्ला. सो अप्सरिया⁶ होला. इति सम्बत १८३३ साल अगहन

१. उमरा=भारादार । २. घरहि=घरपीछे । ३. मोरोअपुतालि=अपुतालि खाने
हकवाला नहुने सम्पति । ४. बितलप=बिर्ता । ५. कुरिया=सोही जग्गामा बसेको मोही ।
६. अप्सरिया=विद्रोही ।

३५

शुदि ८ रोज ४ मुकाम कांतिपुर शुभम्

वि. सं. १८३३ पौष ७ गते श्री ५ प्रतापसिंहले प्रसिद्ध जोगी भगवन्तनाथलाई आफ्नो
राज्यभरिका जोगीहरूका नाइके तथा बाहुन, क्षत्री, मगर, गुरुङ, तामाङ आदि बाहेकका अरू
केही जातका दुनियाँका प्रत्येक घरबाट १।१ आना वार्षिक दस्तूर उठाउने, जोगीहरू मरे
तिनको अपुताली खान पाउने आदि अधिकार दिई यो पत्र लेखिदिएका हुन् । यो रकम द्वारे
उमराउमार्फत मिलाइदिने उल्लेख पनि यस पत्रमा भएको छ ।

पश्चिमतिरको षाड्गुण्यप्रयोगमा यी भगवन्तनाथले पृथ्वीनारायण शाहको निकै मदत गरेका
थिए । यसो हुँदा पृथ्वीनारायण शाहले यिनको निकै कदर गरेका थिए² ।

The second document is an order from Prime Minister Jung Bahadur Rana, which dates from 1904 VS (1847 CE), addressed to the Chepang local authorities, referred to as मिझार <mijhār>[4] who live in the villages of Pinda or Pida situated South of the district of

---

[4] The Nepali term मिझार <mijhār> specifically refers to the local authority in charge of the local organization of *kipāt* lands; the *kipāt* land system was put in place by King Prithivi Narayan Shah in 1773. *kipāt* lands correspond to not state-owned lands, property of the local community, which entails that no tax

Dhading. Prime Minister Jung Bahadur Rana warns them to return to their land and continue their cultivation, after having been reported to have fled their villages "(for fear of) troops sent there to search for people engaged in illicit movements from Nepal (i.e., Kathmandu valley) and the Tarai." An image of this order is given in Figure 2, as translated and published by Regmi (1970: 46).

Figure 2. Second 1847 Nepalese official document mentioning the Chepangs

Order To Kipat-Owing Chepangs in Pinda (West No.1), 1847

From Prime Minister Jung Bahadur,

To Gore Mijhar, Dhane Mijhar and all other Kipat-owing Chepangs of Pinda (West No. 1 district).

It appears that you have fled into forests (for fear of) troops sent there to search for persons engaged in illicit movements from Nepal (i.e. Kathmandu Valley) and the Tarai. You have represented to the respectable people of Pinda that you have not committed any offense, that you feel afraid of coming back to your homesteads and engaging yourselves in cultivation; and that you will do so if an order assuring you of security is granted. The matter was then reported here. As you are subjects, you should not commit any wrong action. If you do anything against the interests of His Majesty, if any person reports the matter here and if you cannot face interrogation, your Kipat lands may be taken away from you and you may be enslaved along with the other members of your family. Understand this well, and come back to and live in your homesteads. Perform the customary function of attending to the Maula (places where sacrifices are offered during religious festivals) and other functions assigned to you and pay the prescribed taxes and levies.

Jestha Sudi 15, 1904
(May 1847)
(Regmi Research Collections 33/20)

****

Contd...

---

can be collected on such lands (Regmi 1974; 1976; 1978; Armbrecht Forbes 1996). The term मिझार <mijhār> has since been found used as a clan name within the Chepang community.

1.3. Earlier research

In this section, I start by situating the contexts and methodologies attached to the earliest literature published on the Chepang language and culture since the first half of the 19th century to the end of the 1960's (§ 1.3.1). I discuss the results of such research in addition to the succeeding literature on the Chepang language (§ 1.3.2) and on the Chepang people and culture (§ 1.3.3), by scholars from Nepal and abroad. Any literature that has ever been written about the Chepangs comes from oral accounts given by Chepang community members to outsiders.

1.3.1. Contexts and methodologies of 19th century to 1960's literature

The first ever accounts of the Chepang language, people, and culture, are found in the work of Brian Houghton Hodgson, British Resident, ethnographer, and botanist, who lived in Nepal from 1820 to 1847 (1848; 1857a; 1874a). In none of these accounts there is mention of any metadata that would locate the Chepang people he met.

At the end of the Anglo-Nepalese war (1814-1816), the reign of King Rajendra Bikram Shah had started (1816-1847), under an absolute monarchy. King Rajendra Bikram Shah was forced to abdicate by Prime Minister Jung Bahadur Rana who established a totalitarian regime and the hereditary succession of the Rana prime ministers up to 1951 holding figurehead positions. Prime Minister Jung Bahadur Rana then imposed King Surendra Bikram Shah (King Rajendra Bikram Shah's nephew) as King figure (1847-1881) and the monarchy became constitutional.

Within this political setting, Brian Houghton Hodgson occupied the position of Assistant British Resident in Kumaon (1819-1820) until the region was annexed from Nepal; he then lived in Calcutta, before he came back to Nepal in 1824 to become the British Resident at Kathmandu in 1833, a position he occupied until 1847. During this period, doubting the British diplomatic agenda, the government did not allow Hodgson to leave the Kathmandu valley (Waterhouse 2004: 4; Whelpton 2004: 25–26). The only place outside the valley Hodgson could reach was the Royal Palace in Nuwakot (Inskipp

& Datta 2004: 135). Hodgson used to employ assistants to go outside the Kathmandu valley to bring him back information, plants and animal specimens for his studies.

What Hodgson wrote has often been cited in the succeeding literature without questioning the sources of his claims or the methodology he employed to collect information, while Hodgson's work may be based on data he did not collect himself, i.e., second-hand data, as noted by Shneiderman & Turin (2006: 99): "Since the Ranas did not permit Hodgson to travel out of Kathmandu, he never conducted ethnographic research per se, working rather with several high-caste assistants who collected data throughout the country and shared their notes with him back in Kathmandu." In addition to relying on information collected before him (Arnold 2004: 197), by geologist Herbert (1842) or zoologist and botanist Hooker (1824; 1848) for instance, Gaenzle (2004: 226) mentions additional methods that Hodgson had to rely on to collect data, such as at least once bringing a native informant in a cage to study their language. Only after he left his position of British Resident of Nepal in 1847, was he able to travel a few times from Darjeeling to the Tarai region in the South of Nepal, gathering botanical and zoological data (Inskipp & Datta 2004: 136; Datta 2004: 156).

Hodgson's assistants were not solely of high caste; as noted by Allen (2016: 5, 194), one of them was Chebu Lama, a hunter, who Allen suggests was of Sherpa, Tamang, or Gurung origin:

> Even before his restoration as assistant to Edward Gardner in April 1825 Hodgson had begun to pay a hunter to bring him animal trophies from outside the Valley - a shikari who, from the evidence of the first specimens brought in, was a Bhotia or trans-Himalayan trader rather than a Valley dweller. Only the name of one of these hunters employed by Hodgson over the course of many years is known: Chebu Lama, which suggests a Sherpa, Tamang or Gurung origin. However, the dramatic improvement in Hodgson's fortunes in April 1825 meant that he could now afford to employ more hunters and trappers as well as a local artist to draw some of these same specimens to his specifications. That first artist was initially put to work doing what he did best: drawing and

painting miniatures for religious works such as illustrated manuscripts, wall hangings and temple decorations according to the set traditions of his ancestors. (Allen 2016: 5)

Thanks to the industry and artistry of Raj Man Singh and at least two of his fellow chitrakars, Hodgson was now accumulating an ever-expanding portfolio of zoological drawings and watercolours that, in his opinion, could be matched against the best of the European artists. (Allen 2015:194)

While it is not specified, it is likely that Hodgson did not himself draw the portraits of the Chepang people holding bows and arrows in the watercolor lithograph that appears in his paper *On the Chepang and Kusunda tribes* published in 1848 in the Journal of the Asiatic Society of Bengalis. The drawings found in Hodgson's publications were mostly done by Newar painter Raj Man Singh and his fellows.

If Hodgson indeed met the Chepang people in person and collected linguistic and ethnographical data, he likely did so with his assistants, details that are not provided by Hodgson (1848: 651; 1874b: 46):

During a long residence in Nepal, I never could gain the least access to the Kusundas, though aided by all the authority of the Durbar; but, so aided, I once, in the course of an ostensible shooting excursion, persuaded some Chepangs to let me see and converse with them for three or four days through the medium of some Gurungs of their acquaintance. On that occasion I obtained the accompanying ample specimen of their language; and, whilst they were doling forth the words to my interpreters, I was enabled to study and to sketch the characteristic traits of their forms and faces.

Hodgson (1848: 651; 1874b: 46) reports the presence of Gurung acquaintances of the Chepang people he met; this suggests that the location of this particular Chepang group was in the Northwestern parts of today's Chitwan district, South of the border of

today's Dhading district; that is, around the village of Dharechok in Ichchhakamana Rural Municipality, or further South to the North of the village of Kalibas in Bharatpur Municipality. These are the only areas adjacent to the route taken by the Royal family members to go hunting in the Chitwan National Park, and which correspond to an area of language contact with the Gurung community – which does not live further East, a region mainly occupied by the Chepangs and the Tamangs; this possible location is represented in Map 2, with the Gurung population in yellow (which expands beyond the arrows), the Tamang population in blue (which expands beyond the arrows), the road from Kathmandu to the Chitwan National Park in brown, the Chepang settlements of Ichchhakamana Rural Municipality in red; see also Map 6 for the areas of language contact between Gurung and Chepang speaking communities.

Map 2. Possible location of Hodgson's encounter with the Chepangs
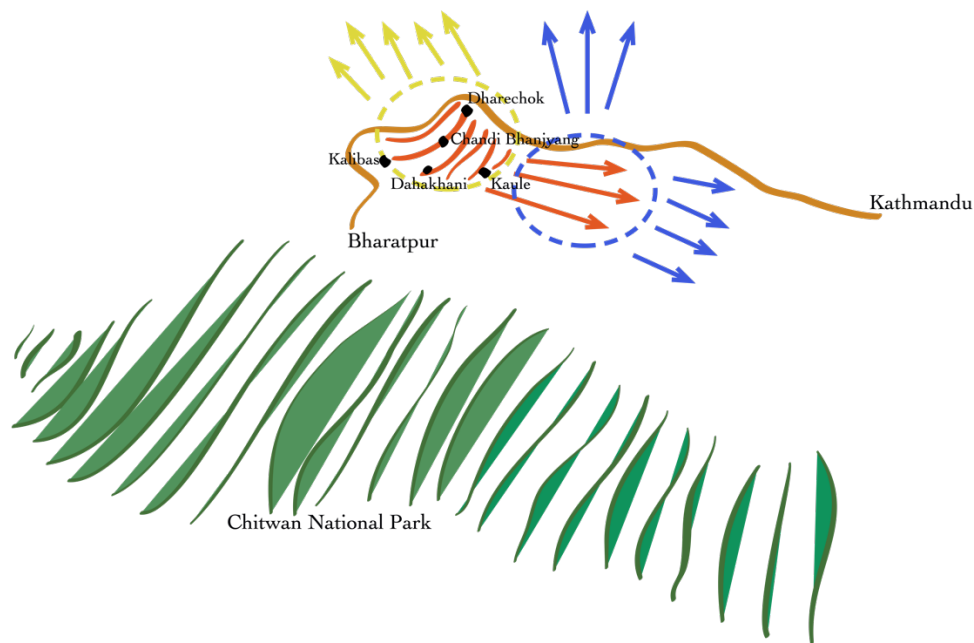
In fact, by the mid-19th century, no settlement of Chepang community existed in the plain of Tarai (down the hills of Chitwan and Makawanpur), which was occupied by the Tharu community; Chepang migrations to the Tarai region likely go back to the second half of the mid-20th century. There were however commercial exchanges that likely already existed in the 19th century and beyond, and which continued until very

recently (childhood of today's people in their thirties), which consisted of Chepang people coming down from the hills along the rivers (Lothar and Manahari Rivers for instance) to Tarai. The trade consisted in selling goods that would grow in the hills and not in the plain, such as millet, buckwheat, Indian butter tree oil or butter and seeds (roasted and eaten, or transformed into poison to catch fish), lentils (*Vigna mungo*), vetch (*Dolichos uniflorus*), and a little less than a dozen of medicinal herb varieties, in order to buy rice and salt, unavailable in the hills.

Whether Hodgson had ever met the Chepang people in Chitwan while on the way to a hunting excursion with the Royal family members remains under question for two main reasons. Reaching the Chitwan National Park at that time was long and hazardous and highly organized in the context of a Royal hunting expedition. It is thus questionable that the Royal family members would have waited for four days on their way to the hunt to let Hodgson spend time with the Chepang people.

Further, Hodgson (1848: 651; 1874b: 46) mentions that the collection of the vocabulary was led "through the medium of some Gurungs of their acquaintance," and that "whilst they were doling forth the words to [his] interpreters, [he] was enabled to study and to sketch the characteristic traits of their forms and faces." This suggests that he did not transcribe the collected vocabulary himself but had it done by a trained assistant, likely the same who transcribed the vocabulary of other languages in absence of Hodgson; that his assistant did so through the presence of Gurung local villagers (Chepangs' acquaintances) and that they all were not using Nepali as a lingua Franca; that Gurung was used between his assistant and the Gurung villagers and that either the Chepangs of this locality were bilingual in Gurung but not conversing directly with Hodgson's assistant, or that the Gurung villagers were bilingual in Chepang and that the Chepangs did not speak Gurung. More likely, he sent his assistants to this area where people knew there were Chepang villages, as he did to collect vocabulary on Kusunda for instance or other languages.

Tachibana (2009: 188–190) emphasizes that it is impossible to determine how much of Hodgson's account of the Chepangs is based on imagination and how much on observation. He concludes that the only certainty in Hodgson's description of the Chepangs is that Chepang's life was not in as natural a state as Hodgson expected.

46

The following literature that appeared in the 19th century mentions, or discusses the Chepang language and/or culture based on Hodgson's observations: Bunsen (1854), Müller (1854), Logan (1856), Latham (1860; 1860; 1862), de Charencey (1862), Beames (1868), Forbes (1877), and Cust (1878).

The second first-hand account of the Chepang people is a manuscript written a century after Hodgson by anthropologist von Nebesky-Wojkowitz (1959) after Nepal opened its border to foreigners in 1951. This paper does not mention anything about the language, but for the first time mentions the location of the Chepangs encountered (1959: 81): villages situated in the area of Upardang Gadhi[5] to the North-West of Shaktikhor in the Municipality of Kalika (wards 11, 9) in Chitwan district.

After von Nebesky-Wojkowitz, the first Nepali scholar to write on Chepang culture was Dor Bahadur Bista in an area situated South of Dhading.

The work of Jest (1966; 1965a; 1965b; 1965c; n.d.) follows and provides the first recordings of the language which date from 1965; Jest mentions as well the location where his study took place: Kanrang Gadhi[6], North of the Municipality of Rapti (ward 13) in Chitwan district.

The work of Caughley starts in March 1969 and a first paper on Chepang phonology is published the same year; the villages where Caughley collected data that year, presented in a report (1969), are all situated in the district of Makawanpur: Chainpur, Bhodibari, Maisirang, Pambung, and Silinge in Raksirang Rural Municipality, and Dumbre and Dhokota in Kailash Rural Municipality.

The locations where the first linguistic research was conducted on Chepang between the 19th and 20th centuries are represented in Map 3.

---

[5] Von Nebesky-Woljkowitz (1959: 81) provided the following GPS location: 84º 37' E 27 º 45' N.

[6] The following GPS location provided by Jest (n.d.) shows a place further East than Kanrang Gadhi: 84º 30' E 27º 45' N.

Map 3. First linguistic research on Chepang - 19th-20th centuries

Linguistic research on Chepang XIX-XXth
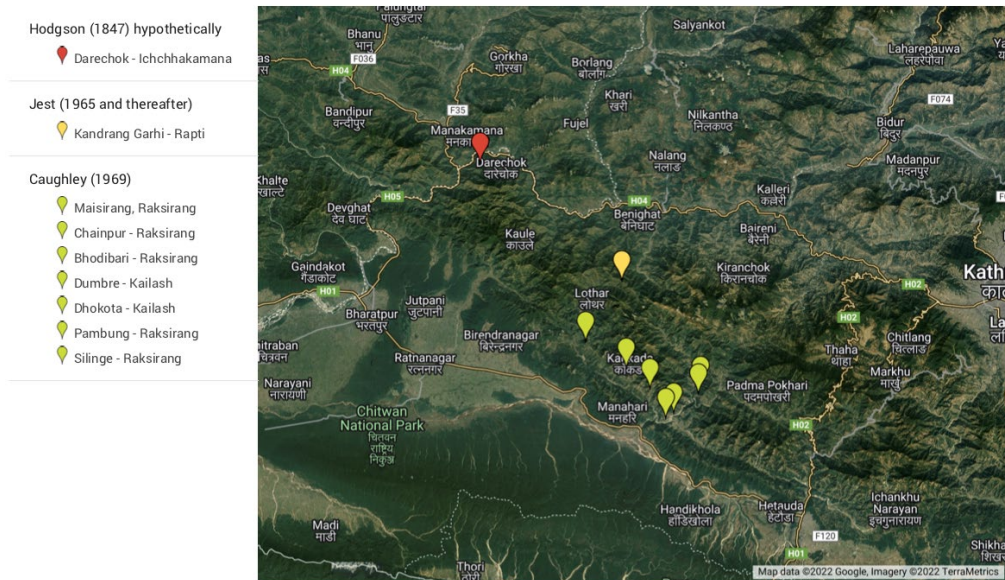
The following literature published at the end of the 20th century focuses on cultural and social aspects of the Chepang community: Upreti (1966), Gurung (1970 and thereafter), Thapa (1974), Bista & Rai (1975), Popham (1978), Rai (1985), Thapaliya (1987).

1.3.2. On the Chepang language

In this section, I briefly present and describe the available earlier literature on the Chepang language, including vocabulary presented or not in comparison with other languages (§ 1.3.2.1), close lexical comparative analyses done between Chepang and other languages (§ 1.3.2.2), analyses of Chepang phonology and morphosyntax (§ 1.3.2.3) and sociolinguistic studies (§ 1.3.2.4).

1.3.2.1. Vocabulary and compared vocabularies

Hodgson (1848) published a first paper 'On the Chépáng and Kúsúnda Tribes of Népál' with a 356-word list including native and borrowed vocabulary. He also provides a list of 27 words that he compares with Tibetan and Lhopa. On the basis of this word comparison, Hodgson (1848: 652–654) suggests that Chepang and Lhopa, a language spoken in Bhutan, are related, and that the Chepang people originally come from Bhutan. This paper was later republished (Hodgson 1848; 1857b; 1874b; 1991).

In 'Comparative Vocabulary of the Languages of the Broken Tribes of Nepal,' Hodgson (1857a) provides another 282-word list presented in comparison with the vocabulary of three IA languages, i.e., Dahari (Dahi, Darhi)[7], Dénwár[8], and Kuswár[9], and three TH languages, i.e., Pahari (Pahi)[10], Bhrámu[11], and Váyu (Háyu)[12].

In Hodgson (1874b: 47; 1874b: 47), a footnote is added where Hodgson proposes for the first time a divide between "complex pronominalized" and "non-pronominalized" languages:

> The oldest tribes of Himalaya, as sufficiently proved by their relative condition and location, are the broken tribes driven to the inclement summits or malarious glens of the Himalaya; and these in general have languages of the pronomenalized or complex sub-type, (…)

This footnote is further detailed with the name of the languages pertaining to the division proposed by Hodgson (1880: 314):

---

[7] This IA language is also known as Darai (Paudyal 2003).

[8] This IA language is also known as Danuwar (Rai & Kuegler 1975).

[9] Dhakal (2014) suggests that Kuswar is a name that was used to refer to a variety of an IA language called Majhi.

[10] This TH language (Smith 2022) is also known as 'Pahi' (Grierson 1909; Toffin 1981), 'Padhi,' (Grierson 1909) and 'Phri' (Gautam & Thapa-Magar 1994).

[11] Also known as Baram, this language is said to be TH and spoken by Darai and Baram community members (Yadava 2007; Kansakar et al. 2014).

[12] This TH language is known as both Vayu and Hayu (Michailovsky 1988; 2003).

The complex Himalavan tongues are Limbu, Kiránti, Háyu, Kuswár, Súnwar, Dhimáli, Bhrámu, Chepáng, Kusunda, etc. The simple or nonpronominalised are Newári, Thumi, Pahi, Múrmi, Gúrung, Mágár Khas (mixed), Lep'cha, Palnsen or Syar'pa (Serpa), Bodo, etc.

Hodgson (1848; 1857a; 1874b; 1991) does not transcribe the glottal stop but uses an acute accent on certain vowels; however, no explanation is provided as for what this diacritic represents. The glottal stop was nevertheless already phonological, as also suggested by Caughley (1982: 6) since it is clearly a retention that goes back at least to Proto-Chepang-Bhujel (PCB). It is possible that pitch differences were already perceptible in the realization of the glottal stop (§ 2.8) and that tonogenesis (§ 2.9) had started taking place.

Other descriptive and comparative analyses of Chepang vocabulary based on Hodgson (1848; 1857a; 1874b; 1991) are found in Logan (1856), Hunter (1868) and Forbes (1877; 1878; 1881).

Caughley (2000) compiled the first Chepang dictionary mainly based on the Chepang variety spoken in Maisirang, Raksirang Rural Municipality (ward 8) in the district of Makawanpur. A recent version (2016) of this dictionary is made accessible online through Webonary developed by the Summer Institute of Linguistics (SIL).

Adhikari (2016; 2017) published a dictionary of Chepang based on data collected in different areas, but the name of the villages where the data come from are not specified. The glottal stop is not transcribed.

1.3.2.2.   Close comparative analyses of Chepang vocabulary

Watters (2003) offers the first in-depth comparative study between Kham, Magar and Chepang, at the levels of the lexicon and morphosyntax. He proposes a reconstruction of some nominal, verbal and morphosyntactic forms based on an analysis of retention and a few shared innovations. Watters (2003: 28–29) concludes that the three languages share a common ancestor at a higher level, showing that Kham and Magar

have more in common than Chepang and Kham. Watters (2003: 29) classifies the two clades Kham-Magar and Chepang-Bhujel under a Kham-Magar-Chepang clade that he calls Central Himalayish, a sub-branch of Himalayan along with Kiranti language or East Himalayan, which as he mentions corresponds to Bradley's (1997) proposal.

Regmi (2012) compares lexical variation between Bhujel and Chepang speech communities. He finds that Chepang and Bhujel vocabulary is 32% to 34% similar and suggests that Bhujel and Chepang are two independent languages rather than language varieties.

1.3.2.3.   Phonology and verbal morphology

Earlier phonological analyses are presented in the introduction of CHAPTER II and that of verbal morphology in § 5.8.2.1. As for the analysis of Chepang phonation, Weidert (1987) provides a description of the realization of laryngeal final and pre-final sonorant[13] glottal consonants /h/ and /ʔ/. Weidert's (1987) analysis is based on lexical items he collected himself with a speaker from Maisirang, Raksirang Rural Municipality (ward 8) and Caughley's recordings that were collected in the same location.

Weidert (1987: 10) describes the tonal realization entailed by the presence of the glottal stop as a high pitch. This analysis is conducted in comparison with tonal realizations found in two TH languages' sub-groups, i.e., Kuki-Naga-Chin and Barish. Weidert does not suggest a tonal realization in presence of the glottal fricative /h/.

Finally, based on Caughley (1969), Pittman (1970) published an article on Chepang prosody.

1.3.2.4.   Sociolinguistics

A report on language use and attitudes of Chepang speakers can be found in Sapkota & Uranw (2013). A sociolinguistic survey was conducted by Adhikari (2006) in

---

[13] The glottal consonants /h/ and /ʔ/ are transcribed as following the final sonorant consonants.

the Eastern parts of the Chitwan district. More on language endangerment is discussed in § 1.7.

1.3.3. On the Chepang people and culture

Hodgson (1848; 1857b; 1874b; 1991) published the first anthropological accounts of the Chepang people that he provides along with that of the Kusunda that he portrays as "broken." Hodgson focuses on the description of the physical traits of the Chepangs and makes a lot of assumptions about the Chepang people living in a state of nature. His descriptions are unsurprisingly offensive and colonialist, a reflection of anthropology of the 19th century (1848: 650):

> They toil not, neither do they spin; they pay no taxes, acknowledge no allegiance, but, living entirely upon wild fruits rod the produce of the chase, are wont to say that the Rajah is Lord of the cultivated country as they are of the unredeemed waste. They have bows and arrows, of which the iron arm-heads are procured from their neighbours, but almost no implement of civilization, and it is in the very skilful snaring of the beasts of the field and the fowls of the air that all their little intelligence is manifested. (…) They are, in fact, not noxious but helpless, not vicious but aimless, but morally and intellectually, so that no one could without distress behold their careless unconscious inaptitude.

This type of discourse resulted in much social damage for the community who still faces the consequences. Subsequent biased conceptions tied to such descriptions have spread and remained in people's minds which can translate in several ways: romanticized visions of the community living in harmony with nature, lack of social respect, non-recognition of their language, culture, spirituality, and history as rich as other communities in Nepal. Dismissive attitudes towards Chepang community members occur daily, ingraining the same type of discourse again in society.

In line with Hodgson's discourse, Bista (1967) or else Caughley (1969; 1971; 1982) continue to strengthen the image of a people living in forest and cave, hunting and foraging. Caughley (1969: 79; 1971) notes that the Chepangs "have few words for agricultural productions and tools," that they "started agricultural life very recently" according to "some other anthropologists," that "they learnt it from Tamangs, Chetris and others and use their words for sickle, spade and axe." This description was corrected by Caughley in 2015; he added the following preface note: "This is a very early account of Chepang culture, written by Dahal and Bandhu, edited by Caughley and published in the Journal of the Tribhuvan University Aug. 1969 6/1:77-89. It was compiled only a few weeks after initial contact. For a much fuller and accurate account see Caughley Chepang Cultural Notes (forthcoming)." Through footnotes he corrected his thoughts about the lack of words related to agriculture: "later it was found that they do have their own words for these items."

Such an interpretation of the community's lifestyle is also present in Caughley (1982) who claims that they were living a semi-nomadic life until recently or else that he could not find "evidence to indicate that the Chepangs had in the past any social organisation at a level higher than that of the family, or extended family."

The observations made on the Chepang society through Chepang community members' accounts of their culture, spirituality, and history, along with recently made available information through two official Nepalese documents dated the 18th and 19th centuries are discussed in § 1.6.

The most recent accounts of Chepang culture and spirituality take a different approach, more centered on the community members' discourse, rather than the result of biased interpretations or the presence of intermediary interpreters between the researcher and the Chepang people.

Gurung (1987; 1989; 1994) offers detailed studies on the social organization (at a level higher than that of the family or extended family) of the Chepang and the spiritual aspects of their animistic beliefs.

Another in-depth study of Chepang culture, spiritual beliefs and shamanism can be found in Dhungel (1994) or more recently Adhikari (2010).

Manandhar (1989; 2000) explores the use of medicinal plants within the Chepang community of Makawanpur district.

Regarding the socio-economic status of the Chepang people, several studies were recently conducted. Shrestha (1997) explores the economic use of the Indian butter tree; Yadav (1997) and Pandey (2001) focus on the role of the forest in the economic development of the community; Bohora (2002) and Pokharel (2002) investigate the socio-economic status of the Chepangs living in the district of Chitwan; Shrestha (2004) conducts a case study in the Rural Municipality of Manahari (ward 4); and Khatri (2017) focuses on the effect of Chepang socio-economy on education.

The richest accounts of Chepang shamanism are found in the work of Riboli (1993; 1994a; 1994b; 2000; 2004; 2011; 2014; 2020; Forthcoming) conducted in different areas of Chitwan and Makawanpur districts over the past thirty years. Riboli's descriptions of Chepang shamanism are based on her collaborative work with Chepang shamans, investigating their shamanic practices, the divinities and spirits involved in such practices, the role of the shaman in the world of the alive and the other worlds invisible to human beings, and finally the impact of Christianity on the Chepang shamanism.

Finally, Tachibana (2009) offers the first in-depth anthropological accounts on broad aspects of the Chepang community. This ethnographic study is the result of over three years of field research in areas of Kalika Rural Municipality, Chitwan district. Tachibana discusses the problems and limitations of ethnography which attempts to portray ethnic identity from subjective self-representations by certain ethnic groups and tries to describe the Chepang ontology as a resource for identity construction from the Chepangs' representations of others (2009: 44):

> By depicting such diverse selves of the Chepangs, we can enrich the resources of identity construction for the Chepang people and provide materials for non-Chepang people to reshape their vision of themselves and the world.[14]

---

[14] English translation of Japanese in collaboration with Kenichi Tachibana.

1.4.   The name Chepang, a source of myths-conceptions

The Chepang language was assigned the ISO 639-3 code 'cdm.'[15]

Since Caughley (1980; 1982), the name Chepang, may it apply to the people or to the language, has been the source of intense etymological speculations.

Caughley (1982: 2) first claims that "the term 'Chepang' is the one used by Nepali speakers (the non-Chepangs) and has associations with the Nepali *cepto* flat-nosed and *cepāṅ* frog which are demeaning, so the word is not in common use amongst the people themselves." This statement is problematic at two different levels; first, it does not recognize the variety of pronunciation of the name Chepang amongst the speakers themselves, and assuming as we will see, that the only real or "pure" pronunciation is [tɕjobãŋ] and that [tɕepãŋ] is not a native pronunciation is clearly wrong; second, that the pronunciation [tɕepãŋ] is the result of two folk etymologies spread by Nepali speaking community members is highly unlikely, since such etymology has never been attested so far, coming from within or outside the Chepang community. It is however possible that these assumptions come from the linguist's investigation of the etymology of the name Chepang rather than such etymologies being present in people's mind. I myself looked into all pronunciation forms to see if they could relate to any Nepali word and show an exonymic origin for the name Chepang. I noticed that the form <cepāṅ> corresponds to a species of frog that very few people have ever heard of, that is, that this species is not commonly referred to unless one studies biology or natural sciences. Such exonymic origin is highly unlikely – because why would the Chepangs call themselves with such form if it had been imposed by outsiders in a diminishing attempt?

Indeed, amongst Chepang speakers, several pronunciations of the name Chepang are attested: [tɕepãŋ], [tɕjobãŋ], [tɕjʌbãŋ], [tɕjebãŋ], [tɕjʌpãŋ], [tɕjopãŋ], [tɕjepãŋ]. This led Caughley (1980; 1982) to suggest a possible etymological origin for the pronunciation [tɕjobaŋ] which corresponds to a common pronunciation of the name Chepang in the area of Maisirang, Raksirang Rural Municipality. This etymology, which

---

[15] For a critical discussion of the existence, use, and administration of such standardized code, see Morey et al. (2013).

is characterized by Caughley (1982: 2) himself as "not satisfactory," is the idea that [tɕjobaŋ] could be morphologically broken down into *tɕjo-baŋ* where *tɕjo* means 'top of' and *baŋ* means 'stone,' translating as 'people of the top of the stones.' As shown in § 3.4.6.1.2, this compositional form is not grammatically possible, and the pronunciation of [tɕjobaŋ] is merely one amongst others.

In the meantime, Caughley (1982: 2) chose to use the form *Cyobang* or *Cyo'bang* [tɕjobaŋ] as the main phonological representation of the name Chepang in particular in the dictionary (Caughley 2000; Caughley 2016) which resulted in strengthening this particular etymology which largely spread in subsequent literature: Rai (1985), Khanal (2014a: 4), Viel (2020: 230). In addition, this etymology gave rise to further aberrant claims that the Chepang people considered themselves coming from the top of the stones (Rai 1985), which led to further negative consequences at the social level.

This etymological interpretation is contested by all the Chepang people with whom a discussion on the etymology of the name Chepang took place. As further mentioned by Riboli (2000: 28), such idea that the Chepangs would come from the top of the stone or even from the stone itself was reinforced by other communities in positions of power: "the remaining members insisting that the story of the stones had been invented by the Brahmin or Chhetris[16] to denigrate them."

Caughley (1982: 2–3) proposes another possible etymology for the name Chepang which consists in breaking down the word into *tɕoʔ*, which means 'child' but that he translates as 'person' since it is also found in the compounds *goj-tɕoʔ* 'boy, man' and *moʔm-tɕoʔ* 'girl, woman,' and *paŋ*, which he interprets as similar in form to the Tibetan suffix *-pa* present in the name of Tibetan communities, such as "Khampa, Sherpa, Lhopa, and Horpa."

This latter hypothesis is indeed more likely. The suffixing form *-pa* does not only have cognates in Tibetan but also in languages likely closer to Chepang at lower levels, such as Wambule or Bahing. In Wambule, *-pa ~ wa ~ wo* is used as a male gender suffix

---

[16] Two highest ranked castes in the system imposed through the 1854 Legal Code (Muluki Ain).

(Opgenort 2004: 132–137), reconstructed as such by Benedict (Benedict 1972: 96), *-pa* (masculine)[17].

Another misunderstanding is found in Bista (1967), which again resulted in spreading wrong information about the Chepang community. He (1967: 120) claims that the Chepangs can be divided into two main groups, again tainted with its load of negative social characterizations, i.e., pukunthali and kachhare Chepangs, while he did not account for the meanings and contexts of use of the words पुकुन्थली <pukunthali> and कछाडिया <kachāḍiyā> sometimes pronounced /kʌtɕʰaɾe/:

> Chepangs make the distinction between two economic groups, those who have developed a purely agricultural economy and others who still partly depend upon food-gathering, hunting and fishing. The former group lives in the eastern part of the region and is known as the Pukunthali; the latter lives in the western part and is known as the Kachhare. Kachhare Chepangs are more backward and primitive than the Pukunthalis. The Kachhares like to be called Sunpraja and have no sub-divisions, while the Pukunthalis are called Praja and have a number of exogamous clans.

The word पुकुन्थली <pukunthali> refers to a location in the Municipality of Pukunthali (ward 16) in the district of Dolakha, where sacred places are dedicated to the cult of the divinity Bhumi. This location is present in accounts of shamanic travels and has a symbolic status in Chepang spirituality and community, sometimes referred to as an area of mythical origin for the Chepang people.

The term कछाड् <kachāḍ> refers to the location on the slope of a hill between the top, known as *lekh ~ lek* which means 'ridge of the hill or mountain,' and the lower parts of a slope, known as *besi* which comes from the idea of 'infinite, boundlessness,' that is, the location where the plain extends and continues.

---

[17] Along with *pa* 'father' and *-ma* (feminine) and *ma* 'mother.'

The terms कछाडिया \<kachāḍiyā\> or कछाडे \<kachāre\> are defined as कछाडमा बसने \<kachāḍmā basne\> (Brihat Nepali sabdakosh, Nepal Academy) which means 'those who live in the कछाड \<kachāḍ\>, that is, between the ridge of the hill or mountain and the plain.'

This area in any hill, whether in Chitwan, Makawanpur, Dhading or Gorkha, is usually the area where the Chepang people and other indigenous communities live and practice agriculture. Such a distinction between two groups of Chepangs, one of which would be more advanced than the other is not attested and contested when discussed with community members. Bista (1967: 120–121) goes on enhancing his descriptions of the different characteristics of each of the groups, and none of these are accurate, raising the question of the source and goal of such claims.

As a final word on this matter for now, the etymology of the name Chepang remains to be explored and it is important to keep in mind that ill-formed, unfounded, and derogatory etymologies only result in harming community members and may give birth to further derogatory fake mythologies on the origin of the people themselves. Such tendencies to try to explain the origin of a word with no linguistic, historical linguistic or historical knowledge explaining that such and such meaning could be "possible" was observed with other groups in Nepal, entailing the same social consequences for the people, such as the Thangmi, as discussed by Shneiderman and Turin (2006: 124).

1.5. Chepang and Praja

The name Chepang is often used by community members to refer to their people and language. During the Panchayat period, which started with King Mahendra in 1960 and ended in 1990, a new term came into use to refer to the Chepang people or language, i.e., प्रजा \<prajā\> "Praja." This term was imposed by the rulers to refer to literally their 'subjects.' The name Praja started to be used as a last name within the Chepang community and to appear on identity cards.

For some community members, the use of the name Praja is not a problem and can even be favored by contrast with the name Chepang; for others, it is the opposite

effect, where people reject the use of the term Praja given its semantics of 'subject of the rulers,' felt to be offensive and degrading, preferring to be referred to as Chepang.

Tachibana (2009: 17–22), who began his research in Kalika (Chitwan) in 1989, at the end of the Panchayat period, notes that the Chepang villagers in this area preferred to be called by the title Praja, as named by the King, and disliked the name Chepang considered disparaging. As another manifestation of the rejection of the name Chepang, Tachibana (2009: 17–22) mentions that elders pronounced their name [tɕobãŋ], while young people sometimes avoided the use of the pronunciation [tɕepãŋ] and called themselves [tɕejebãŋ].

Further, Tachibana (2009: 17–22) explains that, when the 1990 democracy movement led to the King's acceptance of a multi-party system and the abolition of the Panchayat, the term "democracy" was commonly translated as प्रजातन्त्र <prajātantra> "prajatantra," literally 'system of the subjects.' This term soon was considered inappropriate to describe democracy, and later came to be referred to as लोकतन्त्र <loktantra> "loktantra," which means 'system of the people.' Accordingly, the term Praja as an ethnic name was also considered inappropriate, and those who originally disliked being called Chepang came to accept it rather than Praja.

1.6.   The Chepang people – inside the forest, the kingdoms

Through § 1.2, 1.3, 1.4, and 1.5, I intended to give an overview of what earlier literature have claimed about the Chepang people and language. Amongst the descriptive accounts, we have seen several misconceptions that have harmed the community and that may have led to a biased vision of the community in subsequent literature.

It is clear, that the Chepangs have lived close to forested areas, that they have practiced hunting, fishing, and foraging, and that such practices are not different than those of other indigenous communities living in the same environment at a same period, such as the Tamangs, the Gurungs, or else the Magars. The Chepangs do not think they were living in caves, as many have pretended, as much as surrounding communities have not either. While grazing the cattle leads any community living in the hills of Nepal to

settle for a few weeks or months in areas of pasture and therefore accommodate their temporary daily life with a shelter in a hut or a cave, this does not make any of them nomads living in the jungle. These practices are reported in Chepang oral literature, as much as stories of a social and judicial organization based around the presence of local kings or rulers. The Chepang vocabulary shows native lexical items dedicated to hunting, fishing, gathering, and to wild botany, as much as it shows native lexical items dedicated to agriculture, raising of cattle, domesticated animals, ancient varieties of crops, sewing and weaving, or else to the description of house parts.

The Chepangs were not living in the forest outside a social system formed with other neighboring communities but were part of such a system, by contrast with what Hodgson (1848) claims. Not only were they already part of a system as any other indigenous communities by 1776 at the time of King Pratap Singh's order (§ 1.2), which reflects a social organization already based on agriculture, but accounts of local ancestries and history situate the presence of Chepang settlements and local kings or rulers before the unification campaign of Nepal by King Prithivi Narayan Shah which started in 1743 and ended in 1769. For instance, a story narrates that a Chepang king of Gundi (RAP-13), King Gunai, was beheaded by two other neighboring kings, King Damu and King Meme, and that King Gunai's skull was buried with a mango pit which grew a mango tree that only died over a decade ago. The lifespan of a mango tree grown in Asia may be up to 400 years and would correspond to the presence of Chepang kings before the unification of Nepal started in 1743 or even beyond, in the 17th century, that is, during the Malla dynasty that reigned in the Kathmandu valley (13th-17th centuries).

It is therefore very much likely that, in the 17th and 18th centuries, existed Chepang local kings or rulers, also sometimes referred to in Nepali as थुम् राजा <thum rājā>, which literally means 'king of the ridge of the hill or mountain,' who held political power in their communities at the local level in the hills of Northeast Chitwan, Northwest Makawanpur and Southwest Dhading; this historical configuration parallels the presence at the same period of other kingdoms ruled by indigenous communities, such as the Chaubisi kingdoms to the North, the Baisi kingdoms to the West, and Kiranti rulers in Eastern Nepal.

1.7.   Language endangerment, preservation, and revitalization

Chepang is not spoken anymore by the youngest generations in many places, as observed in recent studies (Adhikari 2006: 59–60; Khanal 2014b: 13, 22). Different bodies of institutions have estimated its degree of endangerment as "vulnerable" (Moseley 2010) and "threatened" (Eberhard, Simons & Fennig 2020).

In 2004, the Nepal Federation of Indigenous Nationalities (NEFIN) declared 59 indigenous nationalities as legally recognized by the NEFDIN Act of 2002. These groups were classified into five socioeconomic groups based on various indicators: literacy rate and access to higher education, house type, landownership, occupation, language, and population. These groups are as follows: endangered, highly marginalized, marginalized, disadvantaged, and advantaged. The Chepang people and language were then categorized as highly marginalized. The loss of Chepang speakers is mainly due to socio-economic pressures that may lead parents to not speak in Chepang to their children but in Nepali, hoping to give them an education based on Nepali and English that would lead to better social status.

After the end of the Panchayat period in 1990, the Nepal Chepang Association (नेपाल चेपाङ संघ <nepāl cepāṅ saṃgh>) was created with the aim of promoting and preserving the Chepang language and culture, through the development of museums, participation in indigenous celebrations, and organization of Chepang festivals, such as Nwangi[18]. Each Municipality in the areas where the Chepangs live have one to several local Nepal Chepang Association managed by a board of local members.

During the past few years, the Chepangs have shown a growing interest in preserving their language and culture, using social media as a means of language activism, or organizing writing workshops. Such workshops are not developed in a particular organizational setting, but often the result of individual initiatives often associated with a political party. Such writing workshops are often organized at the local

---

[18] Nwangi is a festival observed in the month of Bhadra (August-September) to celebrate the new harvested crops. The selection of the specific date by the shamans is tied to the phases of the moon. It is celebrated through shamanic performances which start in the evening and last until dawn.

level and last two to three days. In parallel, additional initiatives come from the youth, with for instance the creation of the first sitcom in the Chepang language[19] which blends romance and action.

1.8.   Phylogenetic classification of Chepang within Trans-Himalayan

In this section, I present the Trans-Himalayan (TH) (a.k.a. Tibeto-Burman or Sino-Tibetan) language family in broad terms (§ 1.8.1), review earlier classifications of Chepang (§ 1.8.2), and suggest three possible internal phylogenetic classifications for the languages within Central-Himalayan that needs to be further explored (§ 1.8.3).

1.8.1. Trans-Himalayan languages

Trans-Himalayan (TH) is the second largest language family in terms of speakers, with languages spoken in an area that goes from Pakistan through the Himalayas to South-East Asia. The most ancient written records of TH languages exist for Old Chinese, Tibetan, Tangut, Newah, and Burmese. The phylogenetic classification of TH languages and their branches remains controversial (Driem 2011; Driem 2014; Jacques Draft; Jacques Draft). Two main proposals are under debate: in proposal (1), two main branches split off Proto-Sino-Tibetan (PST) or Proto-Trans-Himalayan (PTH), i.e., Sinitic and Tibeto-Burman (Benedict 1972; Matisoff 1991; Matisoff 2003; LaPolla 1992; LaPolla 1994; Sagart et al. 2019; Zhang et al. 2019; Zhang et al. 2020); in proposal (2), regardless of the number and internal position of the branches splitting off PTH, Sinitic is one of them, while Tibeto-Burman is not, and there is no ancestor of Sinitic that is not ancestor to other branches (Klaproth 1823; Shafer 1955; Driem 1997; Driem 2001; Driem 2011; Driem 2014; DeLancey 2010; DeLancey 2015a; DeLancey 2015b).

---

[19] Link to Chepang Official YouTube Channel:
https://www.youtube.com/channel/UC7Y1picaCw51fRMzPOnSlRw

The debate is grounded on the lack of language descriptions and phonological and morphosyntactical reconstructions at lower-level clades, and on the existence of cognate verbal morphology in many of the TH languages but not in Sinitic. Proposal (1) suggests that, given the absence of verbal morphology in Old Chinese and other languages of the Sinitic branch, verbal morphology should not be reconstructed at the level of PST or PTH and that it would be an innovation in the proto-language of the second branch, i.e., Proto-Tibeto-Burman. Proposal (2) by contrast, supports the reconstruction of verbal morphology at the PTH level and posits its loss in Sinitic.

The TH languages have preserved PTH reflexes at the level of phonology and morphosyntax, and therefore should be reconstructed. There are indeed phonological and nominal and verbal morphological cognates between, for instance: Old Chinese and Burmese (Button 2009), Old Chinese, Tibetan and Burmese (Hill 2014); Old Chinese, rGyalrongic, Tibetan, and Himalayan (Jacques 2015a); Old Chinese, rGyalrongic and Tibetan (Jacques 2015b; Shuya, Jacques & Yunfan 2019); Old Chinese, rGyalrongic, Sal, and Kuki-Naga (Jacques 2019), Old Chinese, rGyalrongic, Himalayan, Sal, and Kuki-Naga (Jacques 2018; 2019; Pons 2021). For a summary of Old Chinese cognates with other TH languages, see van Driem (2007), DeLancey (2013) Jacques (2017).

The TH family is schematized in Figure 3 with three possible branches based on DeLancey's (2015b; 2015a) recent proposal; this proposal reflects three branches out of five (Shafer 1966; Bradley 1997; Bradley 2002; Bradley 2018) if Sinitic and Karen languages are added. The internal branching in Figure 3 is largely based on Bradley (1997), van Driem (2011), and DeLancey (DeLancey 2015b; DeLancey 2015a).

The proposed possible internal phylogenies of Himalayan to be explored are presented and discussed in § 1.8.3.

Figure 3. Trans-Himalayan language family

Trans-Himalayan
- **Western** (Bodic)
  - **Bodish**
  - **?Newa**
  - **Himalayan**
- **Central**
- **Eastern**
  - **Macro-Qiangic**
    - **rGyalrongic**
      - **rGyalrong**
        - **Zbu** — Zbu
        - **North** — Japhug
        - **West** — Tshobdun
        - **East** — Cogtse Jiǎomùzú Brag.Dbar
      - **Khroskyabs**
        - Wobzi Siyuewu
    - **Qiangic**
    - **Naic**
  - **Lolo-Burmese**

Trans-Himalayan
- **Western** (Bodic)
  - **Bodish**
  - **?Newa**
  - **Himalayan**
- **Central**
  - **?Nungic**
  - **Sal**
    - **Jinphaw-Luish** **Northern-Naga** **Boro-Garo**
  - **Kuki-Naga**
    - **Kuki-Chin** **Ao** **Angami** **Tangkhul**
- **Eastern**

1.8.2. Traditional phylogenetic classification of Chepang within Himalayan

In this section, I present the classification of Chepang from Shafer (1955) to Schorer (2016). Before Shafer (1955), the classification mentioning Chepang, and based on the language data found in Hodgson (1848; 1857a; 1874a), broadly consists of positioning Chepang within a group of pronominalized languages, by contrast with other non-pronominalized languages (§ 1.3).

Shafer (1955) classifies Chepang along with Magar and Hayu in the West Central Himalayish sub-branch (called section) of the Bodic branch (called division); in this classification, the TH family counts six main branches (or divisions). Benedict (1972) classifies Chepang in an Chepang-Vayu (Hayu) clade within a Bahing-Vayu sub-branch. Thurgood (1984) suggests that Chepang falls in a Rung group along with for instance rGyalrong, Kham, Qiang, Tangut.

Bradley (1997) proposes to classify Chepang within the Central-Himalayan sub-branch of the Western (or Bodic) branch, along with Magar, Kham, Newah, and Raji-Raute. Bradley's (2002) classification of Chepang is like Bradley (1997): within the Western (or Bodic) branch, the Central-Himalayan sub-branch became Western Himalayan, and the Kiranti sub-branch became Eastern Himalayan. In this proposal, Baram and Thangmi are added to the Western (or Central) Himalayan sub-branch.

Based on Bradley's (Bradley 1997; 2002) proposal of a Himalayan sub-branch, Watters (2003) proposes an internal branching for the languages of the Central-Himalayan (or Western Himalayan) sub-branch, based on phonological and morphological cognates (retention and shared innovations). He proposes that Kham and Magar form a clade and Chepang and Bhujel another one. The choice of Watters (2003) to put together Kham and Magar is that Kham is closer to Magar than it is to Chepang and Bhujel. This proposal is to be further investigated along with another one presented in § 1.8.3.

Matisoff (2015) also suggests that Kham, Magar and Chepang are part of a same sub-branch, and places it within a Himalayish branch along with other sub-branches, such as Tibeto-Kanauri, Newar, and Kiranti.

Based on van Driem's (2001) proposal of a Magaric sub-branch formed with Magar and Kham, Schorer (2016) proposes a Greater Magaric clade, which includes Magaric (Magar, Kham), Proto-Dura (Dura and Tandangre) and a third sub-branch, i.e., Chepangic-Raji, which splits into Chepangic (Chepang, Bhujel) and Raji-Raute (Raji, Raute, Rawat).

Besides that of Benedict (1972), the classifications of Chepang presented above have all in common to place Chepang (and Bhujel) together with Magar, or with Kham, or with Magar and Kham. This is exactly from where we now have to start looking into their phylogenetic relationships, that is, how these languages branch out with one another within a clade that would additionally likely include Dura and Tandangre as proposed by Schorer (2016) and Pons (2017).

The other clades often associated with Magar, Kham, Chepang and Bhujel, are Kali-Karnali ("Raji-Raute") and Eastern Himalayan (or Kiranti). The internal branching of the Himalayan clade, splitting into three sub-branches is a possible proposal (Bradley 1997; Bradley 2002; Driem 2001; Rastogi 2012; Krishan 2001; Schorer 2016; Pons 2017), as illustrated in Figure 4.

Figure 4. Central-Himalayan internal phylogenetic classification (1)

66
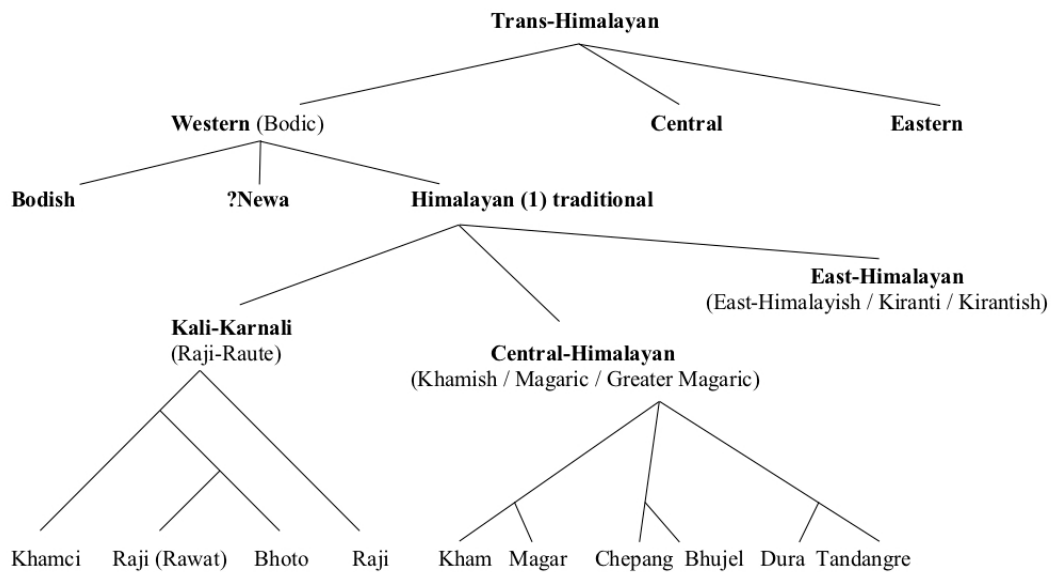
1.8.3. Himalayan internal classification proposals

The reconstruction of lower-level clades is decisive in the understanding of how the languages relate at higher levels. The preliminary analysis of the retention and shared innovation between Magar, Kham and Chepang is important and needs to be developed further.

Watter's (2003) proposal of the internal phylogenetic relationships between Magar, Kham and Chepang within Himalayan is represented in Figure 5, with the addition of a Dura-Tandangre clade (Schorer 2016; Pons 2017). By contrast with the internal classification proposed in Figure 4, there is a common ancestor between Proto-Kham-Magar (PKM) and Proto-Chepang-Bhujel (PCB), that is not shared with the common ancestor of the Dura-Tandangre clade.

As stated in § 1.8.2, Watters finds that Kham is more related to Magar than Kham is to Chepang and Bhujel, which led him to classify them into two separate clades: Kham-Magar and Chepang-Bhujel. However, it is also possible that Kham-Magar-Chepang-Bhujel form a clade and that Kham splits off their ancestor, i.e., Proto-Kham-Magar-Chepang-Bhujel, earlier, and that Magar-Chepang-Bhujel forms a clade splitting later. This proposal is represented in Figure 6.

Both of these proposals are possible and should be further investigated looking at more phonological and morphological cognates, results of retention and shared innovation; while it is demonstrated that Kham and Chepang do not have as much in common as Kham and Magar (Watters 2003), there are still interesting sound changes left to explore between Magar and Chepang-Bhujel, in addition to cognate verbal argument indexation, cognate ancient derivational morphology, and possible shared innovations.

Figure 5. Central-Himalayan internal phylogenetic classification (2)

Trans-Himalayan (2) Watters (2003) for PKMCB as Central Himalayan (did not include Dura)

- **Western** (Bodic)
  - **Bodish**
  - **?Newa**
  - **Himalayan**
    - **Kali-Karnali** (Raji-Raute)
      - Khamci
      - Raji (Rawat)
      - Bhoto
      - Raji
    - **Central-Himalayan** (Khamish / Magaric / Greater Magaric)
      - Kham
      - Magar
      - Chepang
      - Bhujel
      - Dura
      - Tandangre
    - **East-Himalayan** (East-Himalayish / Kiranti / Kirantish)
- **Central**
- **Eastern**
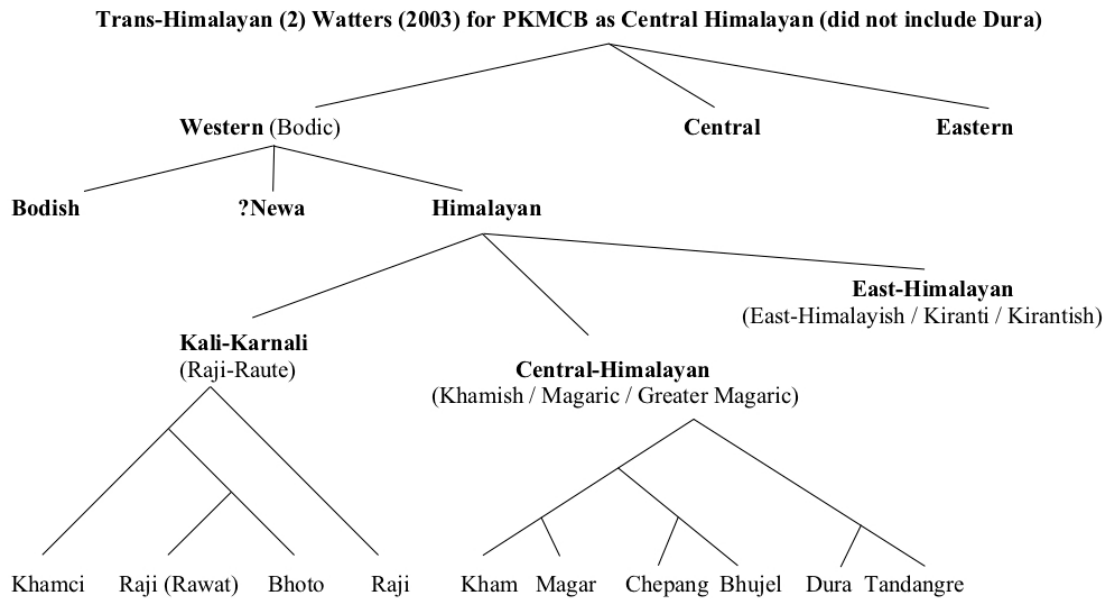
Figure 6. Central-Himalayan internal phylogenetic classification (3)

Trans-Himalayan (3) More retention and shared innovation in PMCB that between Kham and PCB

- **Western** (Bodic)
  - **Bodish**
  - **?Newa**
  - **Himalayan**
    - **Kali-Karnali** (Raji-Raute)
      - Khamci
      - Raji (Rawat)
      - Bhoto
      - Raji
    - **Central-Himalayan** (Khamish / Magaric / Greater Magaric)
      - Kham
      - Magar
      - Chepang
      - Bhujel
      - Dura
      - Tandangre
    - **East-Himalayan** (East-Himalayish / Kiranti / Kirantish)
- **Central**
- **Eastern**
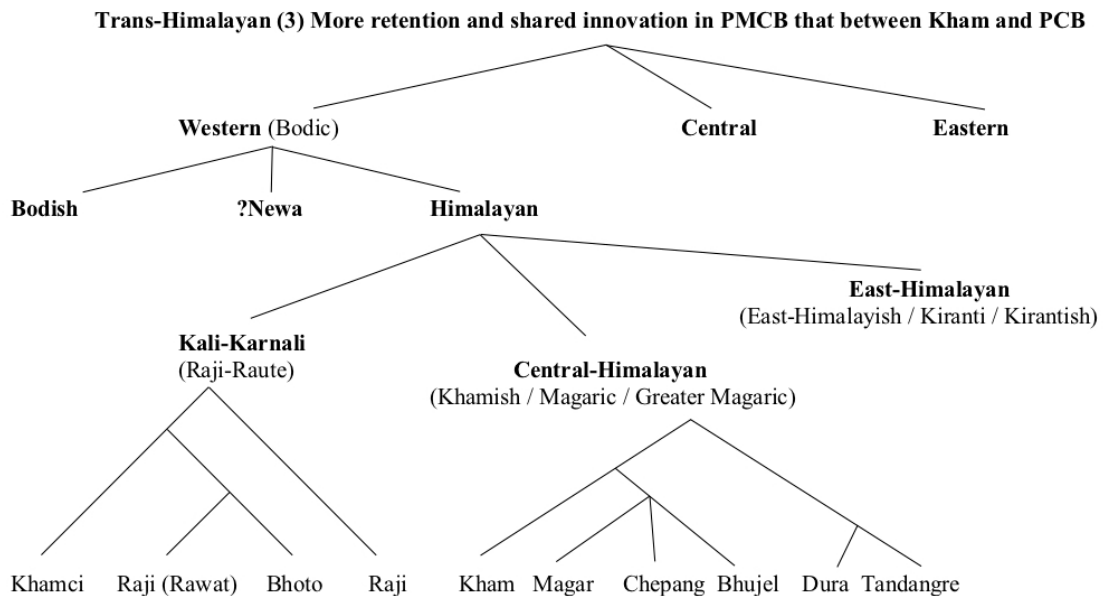
1.9. Language variation

The present dissertation primarily focuses on the description of the Chepang language varieties spoken in the basin of the Lothar River, in between the districts of Chitwan to the West, and Makawanpur to the East. In addition, the varieties of Lothar are further compared with three other varieties, specifically: Manahari, Rapti, and Handikhola. While this study seeks to embrace variation, many other varieties still need to be acknowledged and explored, and aside from the varieties that have been documented and examined for this project, much remains to be explored.

The motivation behind describing variation in the phonological and grammatical structures of Chepang is threefold: first, it aims at giving Chepang community members a more accurate representation of the diversity in their language, rather than focusing on a single variety spoken in one village; second, it allows us to posit more-informed historical hypotheses regarding the evolution of Chepang and the reconstruction of Proto-Chepang (PC) and beyond; third, it provides a solid base for future investigations of Chepang and its varieties.

Language variation is present in Chepang at all levels of the language structures: phonology, lexicon, morphology, and syntax. This can be the result of the internal evolution of Chepang native features and constructions, or due to the external influence of intense language contact with Nepali, yielding to the borrowing of words, morphemes, or constructions through calquing, or changes affecting for instance phonological and prosodic features.

The Chepang community members who have closely collaborated on this project often affirm that, despite variation, they do not have great difficulties understanding each other. Community members are aware of language variation; they can identify the geographical areas where a specific form or feature is used and describe the specific forms or features that are characteristic of a particular geographical area.

A recurrent observation by Chepang community members of the Lothar River is that the language varieties spoken in Maisirang (Raksirang-8), i.e., a Manahari variety, or in the areas of Kaule (Icchakamana-1) and Siddhi (Kalika-11), significantly differ from theirs, while remaining understandable. The varieties of Kaule (Icchakamana-1) and

Siddhi (Kalika-11), which are spoken in Chitwan further west of the Lothar River, along the Kayar River to the border with the district of Gorkha, have not been investigated so far. However, the language spoken in Maisirang has been documented and described by Ross Caughley since 1968 (1982:8). His findings are mostly based on the language spoken by Bhabikan Chepang who started to work with him when he was 23 years old. For the present project, linguistic data were also collected in Maisirang, and were compared with that analyzed by Caughley and other varieties.

1.9.1. Language varieties and areas of study

The names given to the over-arching divisions between the studied language varieties were chosen to reflect the geographical areas where these varieties are spoken. Since people's settlements are found on the sides and crests of hills born from the course of the rivers in the hills, or along the rivers in the plain, the names of the rivers were chosen to label these varieties. I divide the studied language varieties into three main groups: Lothar, Manahari, and Rapti and Handikhola. This division is based on the linguistic differences observed between the varieties, considering the late drastic changes that have taken place in the varieties spoken in the plain (Rapti and Handikhola) because of language contact with Nepali.

This divide does not imply that variation does not exist within each group. On the contrary, variation is also present within each of these groups. However, beyond the linguistic differences that exist amongst the language varieties of a group, these may have more in common with one another than with the language varieties of another group, that is, they may share specific commonalities that are not found in other groups, or that may differ more greatly with that of other groups. The linguistic features characteristic of each group, based on the above mentioned criteria are described in the following sections: § 1.9.1.1 for Lothar, § 1.9.1.2 for Manahari, and § 1.9.1.3 for Rapti and Handikhola. Finally, this section ends with an account and discussion of the dialectal divisions proposed by Caughley (§ 1.9.2) and by a brief description of our present approach to describing language variation in this grammar (§ 1.9.3).

Map 4, which is based on Map 1, represents the study area of the Chepang language varieties. It is marked by a red square.

The locations of the 28 villages where the linguistic data have been investigated for this project are shown in Map 5.

As mentioned, the language varieties primarily described in this dissertation are that of Lothar and Manahari. The Lothar variety corresponds more specifically to the varieties spoken in 14 villages situated in the Municipality of Rapti and in the Rural Municipality of Raksirang. These villages are marked in red on Map 5. Villages that represent the varieties of Manahari, and Rapti and Handikhola, are respectively marked in green and light blue on Map 5.

Map 5 is an impressionistic representation of the locations of the villages with regard to the rivers. The exact locations of the villages are shown on a Google Map created with GPS data collected in each village, since most of these villages do not appear on Google Map. This map is made accessible through the following link: https://www.google.com/maps/d/u/0/edit?mid=184syxxVtcrv_CHTEDchg-54doVw4ez0&usp=sharing

Finally, the abbreviations of the names of the villages used in Map 5 and more generally in this study are explicated in Table 2, along with the abbreviations relative to the current (Municipalities and Rural Municipalities) and former (Village Development Committees (VDC)) administrative divisions, current Municipalities and Rural Municipalities sub-divisions (ward numbers), and districts where these villages are situated.

Map 4. Study area of the Chepang language varieties

Map 5. Chepang language varieties studied
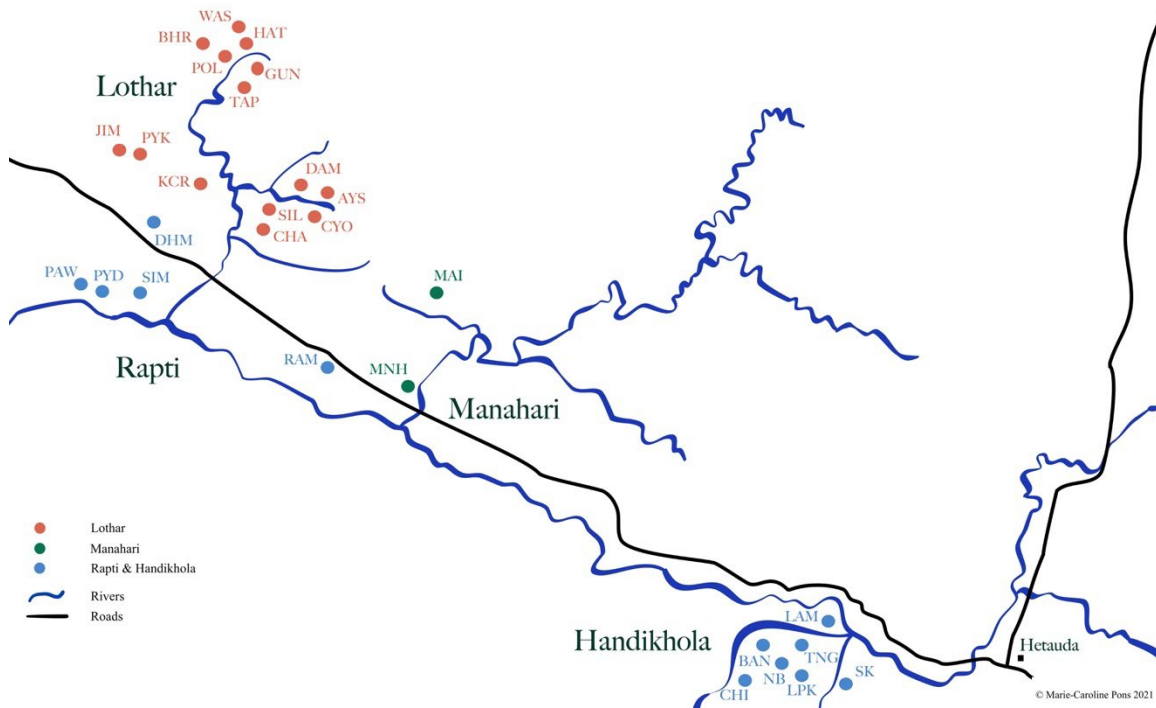
Table 2. Chepang studied language varieties

Varieties	municipality, ward no & villages		ABBR.	former VDC	district
Lothar					
	Municipality	**Rapti-13**	**RAP-13**	Lothar	Chitwan
		Bhantarang	BHR		
		Gundi	GUN		
		Hatti Sunde	HAT		
		Polkim	POL		
		Tapang	TAP		
		Wasbang	WAS		
	Rural Municipality	**Raksirang-6**	**RAK-6**	Kankada	Makawanpur
		Chapala	CHA		
		Cyorang	CYO		
		Silinge	SIL		
	Rural Municipality	**Raksirang-7**	**RAK-7**	Kankada	Makawanpur
		Aisirang	AYS		
		Dambarang	DAM		
	Municipality	**Rapti-11**	**RAP-11**	Korak	Chitwan
		Jimling	JIM		
		Panyak	PYK		
		Kuccur	KCR		
Rapti					
	Municipality	**Rapti-6**	**RAP-6**	Bhandara	Chitwan
		Pawari	PAW		
		Pyari Dap	PYD		
	Municipality	**Rapti-3**	**RAP-3**	Piple	Chitwan
		Dhameli	DHM		
	Municipality	**Rapti-5**	**RAP-5**	Piple	Chitwan
		Simara	SIM		

Varieties	municipality, ward no & villages		ABBR.	former VDC	district
	Rural Municipality	**Manahari-7** Ramantar	**MAN-7** RAM	Manahari	Makawanpur
Manahari					
	Rural Municipality	**Raksirang-8** Maisirang	**RAK-8** MAI	Kankada	Makawanpur
	Rural Municipality	**Manahari-7** Manahari	**MAN-7** MNH	Manahari	Makawanpur
Handikhola					
	Rural Municipality	**Manahari-4**	**MAN-4**	Handikhola	Makawanpur
		Bankarya Tol	BAN		
		Chisopani	CHI		
		Lamitar	LAM		
		Lampakha	LPK		
		Naya Basti	NB		
		Siddha Kali	SK		
		Tongra	TNG		

1.9.1.1.   Lothar varieties

Lothar refers to the language varieties spoken in villages of the Municipality of Rapti (wards 13, 11) in the district of Chitwan, and in villages of the Municipality of Raksirang (wards 6, 7) in the district of Makawanpur.

According to community members' memory of their ancestry, the settlements of people living in Rapti-13, Rapti-11, and Raksirang-6 can go back to a minimum of six to ten generations for the latest, that is 150 to 250 years (25 years per generation) or back to about 1870 to 1770. Community members' accounts of their local history indicate that their settlements may be even older, as they narrate events that took place before the

Gorkha conquest which started around 1743 in the area followed by the unification of Nepal by Prithivi Narayan Shah in 1769.

The language varieties of Lothar are generally well preserved. This can primarily be seen through the predominant use of native lexical items rather than borrowings. The variety that has lost the most native lexical items in favor to the use of borrowings is Rapti-11, although elder speakers still have knowledge of their existence and forms.

One of the main features shared by these varieties is a widespread tonal realization of the glottal stop, rather than through glottal constriction and release which is nevertheless also preserved. Tonal realization is attested amongst both male and female speakers. However, this does not mean that glottal constriction and release is not observed at all in people's speech. While this needs further investigation, there are positions in prosodic and/or morphosyntactic constructions that may favor the occurrence of a constriction and release (§ 2.10.5). Speech rate in the varieties of Lothar may differ, with a speech rate somewhat slower to very slow observed in the village of Dambarang (RAK-6). Native verbal morphology is very well preserved in all Lothar varieties, and some variation occurs, such as the marking of 1st person past tense in for instance the villages of Polkim and Syamrang (RAP-13) where the morpheme *=kaŋ* is used and not *=alaŋ*.

1.9.1.2.   Manahari varieties

The varieties of Manahari correspond to the language spoken in villages of the Municipalities of Raksirang (ward 8) and Manahari (ward 7), in the district of Makawanpur.

According to the community members of Maisirang (RAK-8), the first settlements of the people in the area go back to six generations, that is around 150 years (25 years per generation), or to about 1870. This is corroborated by the personal story of the eldest community member of Maisirang, Gopal Chepang, 102 years old (as for 2021 CE or 2078 VS); he was born in 1976 VS (1919 CE). He is the elder brother of Bhabikan Chepang. At age 14, Gopal Chepang witnessed the 1990 VS (1933 CE) earthquake in Maisirang. According to him, his parents and grandparents were also born in Maisirang, probably

around 1870 and 1880, respectively. According to the community members of Maisirang, the first settlers of Maisirang arrived from a village called Pambung (RAK-8), situated further West in the hills of Makawanpur.

Some communities more recently migrated from the hills to Manahari (MAN-7) further South in the plain of Tarai.

The Manahari varieties are also well preserved. Some native roots still attested in Raksirang (ward 8) and Manahari (ward 7) have been replaced by Nepali borrowings in other varieties, such as *kaɾ* vs. *taɾa* (<N.) to mean 'star.' In these varieties, the realization of the glottal stop is also tonal, similarly to the Lothar varieties, in addition to stronger and more widespread glottal constriction and release than in Lothar. Stress is observed to be more marked, in the sense that the initiator of stress is stronger, inducing like the presence of a glottal stop at the onset of syllables of certain types (§ 2.10.5).

1.9.1.3.   Rapti and Handikhola varieties

The developments of the Rapti and Handikhola varieties resulted from a more intense language contact with Nepali and Tharu speaking communities. This entails a more frequent use of Nepali as lingua franca.

The settlements of the speakers of Rapti (RAP-3, RAP-5, RAP-6, MAN-7) and Handikhola (MAN-4) varieties are born of recent waves of migrations that go back to the mid-19th century, during and after the Panchayat period (1960-90). These migrations were mainly due to natural catastrophes (landslides, floods), environmental conservancy planning forcing people to leave forested areas, or to search for new land to practice agriculture or access labor opportunities.

The Panchayat period also entailed the migration of Chepang communities from their villages in the hills of Chitwan to work in the fields of Nepali-speaking landlords settled in the area of Talti (Benighat-Rorang Rural Municipality) in the district of Dhading. After the Panchayat, some stayed while others left to settle along the Handikhola River (MAN-4) and later the Rapti River (RAP- 3, RAP-6). The migrations that took place after the Panchayat period were mainly from the hills to the plain of Tarai, or to the forests of the Chitwan National Park and the Parsa Wildlife Reserve.

Community members of Rapti (wards 3, 5, 6) and Manahari (ward 4) identify the following Municipalities and Rural Municipalities in the hills as the locations from where their ancestors come: Raksirang and Kailash Rural Municipalities (Makawanpur), Rapti Municipality (Chitwan), Benighat-Rorang Rural Municipality (Dhading).

However, the settlements in Rapti Municipality (wards 3, 5, 6) are more recent than that of Handikhola (MAN-4), which led to different linguistic settings regarding the preservation of the language.

The Rapti and Handikhola varieties show a lower number of speakers in older generations (below 50 years old). In daily activities, Chepang community members who speak the language often use Nepali to communicate amongst each other, rather than Chepang. The prosodic patterns of Chepang are completely lost in particular in the varieties spoken in Handikhola, which has completely borrowed the prosodic features of Nepali. In addition, the glottal stop may be completely absent with no tonal contrast left either. Native verbal morphology is not as preserved as in the Lothar and Manahari varieties and more vocabulary was borrowed from Nepali.

In the varieties of Rapti, the situation varies. Either community members do not speak the language at all, or some speak the language in a way that reflects the language spoken in their original area. This is due to the fact that the Chepang settlements in this area are the latest and that they were formed through migrations from various places.

1.9.2. Dialectal division by Caughley (1982)

Caughley (1982: XV, 9–10, 182–191) suggests the existence of three Chepang dialects: Eastern, South-Western, and Western. Both Eastern and South-Western correspond to Chepang, while Western refers to Bhujel, the most closely related language of Chepang. At the time, Bhujel (also called Bhujeli) was not officially recognized as a language of its own. Caughley (1982: XV, 9–10, 182, 188) considered Bhujel as a Western sub-dialect of Chepang along with Mid-Western, spoken north of the Kayar River and Mahabharat Range.

For Caughley (1982), the Eastern dialect (which thus concerns Chepang specifically), is further divided into North-Eastern (also referred to as Northern), and

South-Eastern. The North-Eastern dialect corresponds to what I call Lothar varieties, and the South-Eastern dialect corresponds specifically to the variety spoken in Maisirang (Raksirang-8), which is the variety to which his descriptions mostly rely on; the variety of Maisirang belongs to what I refer to as Manahari varieties. Finally, the South-Western dialect corresponds to the language varieties spoken along the Kayar River further west of the Lothar River. Caughley (1982: 190) notes that this latter dialect presents linguistic features of both Western and Eastern dialects and suggests that it is the result of borrowing from the Western dialect, that is, from Bhujel. The Kayar variety (South-Western) has not been investigated for this project but recordings accessible through the work of anthropologist Kenichi Tachibana who has been working with the Chepang community since 1989 are currently being processed for future projects. Rather than having borrowed features from Bhujel, it is possible that the Kayar variety reflects a clade that would have split earlier from Proto-Chepang (PC), after the split of Proto-Chepang-Bhujel (PCB).

Caughley (1982) does not specify the exact locations of the villages from where his data come from regarding the North-Eastern, South-Eastern, South-Western and Western dialects, except for his main area of study of the South-Eastern dialect, i.e. Maisirang (Caughley 1982: 8): "The Maiserang dialect is taken as the standard for this thesis. However further material was obtained from villages to the North, West and East of Maiserang, which itself lies in the southern and central part of the Chepang region and also from Bhujeli."

Caughley's dialectal division is based on the phonological and morphological differences he observes between the dialects (1982: 9–10, 182–191). These observations are reported in Table 3 and concern North-Eastern, South-Eastern, and South-Western dialects.

Table 3. Phonological and morphological linguistic differences found in Caughley's (1982) dialectal division of Chepang: North-Eastern, South-Eastern, and South-Western[20]

		South-Eastern (Maisirang)	North-Eastern (Lothar)	South-Western (Kayar)
Phonology				
(1)	/e/, /o/	lower	higher	higher
(2)	/ə/ vs. /ʌ/	higher central /ə/	lower back /ʌ/	lower back /ʌ/
(3)	alveolar cluster	no	yes	yes
(4)	ʔ Ci sonorants	no	no	yes
(5)	ʔ realization	falling or high-falling pitch	absence of glottal	high pitch on following syllable
Morphology				
(1)	Past 2/3 DITR/TR	*-ʔakan*	*-kan*	?
(2)	Past 1SG	*-ʔalaŋ*	*-kaŋ*	?
(3)	Non-Past INV 3>3SG	*-naʔ-thəy*	*-naʔ-təy*	?
(4)	Non-Past DIR/3O	*-wʔ*	*-na-w*	*-nə-wʔ*
(5)	Past 2/3PL INTR	*-ʔaka-y*	*-ʔaka-y*	*-ʔala-y*

In the remaining of this section, I briefly compare what I observed in the Chepang varieties studied in this project, i.e., Lothar and Manahari varieties, with Caughley's (1982) findings on Eastern dialects. I do not comment on the differences he observes between Chepang (Eastern and South-Western) and Bhujel (Western), and only when possible I comment on the differences between Eastern (Lothar and Manahari) and South-Western (Kayar), since no analysis has yet be done on the Kayar varieties.

As for phonological features: (1) height differences in the pronunciation of the vowels /e/ and /o/ are still present between the variety of Maisirang (RAK-8) and Lothar varieties, but a socio-linguistic study would help understand the discrepancies between their realizations amongst speakers since it is not presently observed with all speakers of Maisirang (RAK-8); (2) a variation between the mid central vowel /ə/ and mid back vowel

---

[20] Caughley's (1982) transcriptions are preserved in this table, while I use my glosses for the description of the morphemes, since at times, it may differ from Caughley's.

/ʌ/ is possible, but would also need further investigation; (3) the presence of consonant clusters formed with alveolar stops is attested in some Lothar varieties which correspond to a few villages of RAP-13: Polkim and the neighboring villages of Syamrang, Sarling and Yuiling (§ 2.3.5); (4) although the varieties of the Kayar River (South-Western) remain to be investigated, initial glottalized sonorants seem unlikely, particularly because they are not attested in other Himalayan languages and because, in the single minimal pair Caughley provides, it seems that one of the roots may have been borrowed from Nepali: *myan* 'hair' vs. *ʔmyan* 'net-bag'; *myān* 'long sheath, scabbard' (<N.); (5) the glottal stop /ʔ/ may be realized as a constricted segment in Lothar (North-Eastern) additionally realized with a raising (mid to high) tone (§ 2.8), by contrast with Caughley's observations; the glottal realization needs further research regarding the variety of Maisirang (RAK-8).

As for morphological features: (1) Caughley describes the morpheme *-kan* as found in North-Eastern (Lothar) and *-ʔakan* as found in South-Eastern (Maisirang), while I find that both *=ka=n* 2/3.PST=DIR/TR and *=aka=n* 2/3.PST=DIR/TR are attested in all Lothar varieties; (2) Caughley states that the 1st person past morpheme *-kaŋ* is attested in North-Eastern (Lothar) while *-ʔalaŋ* is used in South-Eastern (Maisirang), while I observe that both forms *=alaŋ* and *=akaŋ ~ kaŋ* are used in all Lothar varieties, with the presence of *=kaŋ* predominantly attested in a few villages of RAP-13: Polkim and the neighboring villages of Syamrang, Sarling and Yuiling (§ 5.7.11); (3) the combination of the inverse morpheme *-ta* and the 3rd person singular morpheme *-i* in non-past tense is described by Caughley as being non-aspirated in North-Eastern (Lothar), giving *-təy*, and aspirated in South-Eastern (Maisirang), giving *-tʰəy*, while both forms are attested in free variation in all the studied varieties; (4) the difference in pronunciation described by Caughley between the vowel [a] and [ə] in the realization of the non-past marker *-na* is also observed in free variation in all varieties; this is a change from [a] to [ʌ] (or [ə]) observed in front of the 3rd person object or direct marker morpheme *=u* (§ 5.8.3.5); (5) the form *-aka-y* in past tense with 2nd and 3rd person in intransitive constructions is attested in all studied varieties and no example has been found so far with the form *-ala-y* as observed by Caughley for South-Western (Kayar).

1.9.3. Present approach to describing language variation

The phonology chapter includes descriptions of the morphophonology. All the sound changes described as morphophonological are conditioned by the environment within which they occur. The sound changes affecting vowel and consonant phonemes, or vowel glide sequences can occur within a single morpheme, at morpheme boundaries within a word (or internal sandhi) or between words (or external sandhi).

I have classified the type of sound changes according to the type of variation attested in three categories: regular, free variation, sporadic. The sound changes classified as "regular" are attested with all roots, morphemes, or words that meet the described conditions for the sound change to take place, across all the different studied varieties and regardless of sociolinguistic differences between individuals. Some sound changes can be considered to exist in "free variation." In this case, it means that they are not necessarily attested with all speakers of the studied varieties. This type of sound change often reflects sociolinguistic variation. While the limits of this project did not allow an in-depth sociolinguistic analysis of this type of variation, I provide major leads to pursue such analysis in the future, and in particular accounting for the geographical area where these changes are attested. Finally, some of these sound changes are called "sporadic," in the sense that they are restricted to specific roots, morphemes, or words. This type of sound changes is attested in free variation amongst speakers.

1.10. Language policy and language contact

Language contact between the Chepangs and other communities varies geographically. The Chepangs have historically been living in the hills of the Mahabharat Range (§ 1.1). Their presence in the plains of Chitwan and Makawanpur is the result of later waves of migration (§ 1.9). The information presented in this section is based on the observations of the Chepangs living in the discussed areas and demographic data from the 2011 Census by the Nepal Central Bureau of Statistics (CBS).

One of the languages with which the Chepangs have been in contact is Nepali. Nepali, also referred to as खस कुरा <khas kurā>[21], has been the lingua franca of Nepal since the 14th century (Bandhu 1989; Bista 1991). It was originally spoken in the Western parts of Nepal, in the Karnali region, where the Khas kingdom was established in the 11th century. Nepali translations are attested as early as the 15th century (Bandhu 1989). But it was after the unification of Nepal was completed in 1769 by King Prithivi Narayan Shah that Nepali was hegemonically imposed in all domains of society. Nepali literature flourished from the end of the 18th century through the 19th century[22], and Nepali became the official language of Nepal in the 1930s (Bandhu 1989). This period also marks the use of the term Nepali to refer to the language (Pradhan 1991; Burghart 1996). The term Nepali was made popular through its usage in British English (Burghart 1984). It derives from the endonym नेपाल भाषा <nepāl bhāṣā> used by the Newah community to refer to their own language[23] (also known as नेवारि <nevāri>) (Regmi 1961). The imposition of Nepali as the language of the nation with the goal of reinforcing civic nationalism (Shrestha 2007) participated in maintaining the political and social domination of Nepali native speakers, i.e., high Hindu castes, over indigenous people (Gaige 2009; Toba, Toba & Rai 2005; Lawoti 2005; Shrestha 2007; Lawoti & Hangen 2012).

Before the 1950s, up until the end of the Rana regime (1846-1951), access to formal education was quasi limited to the children of the Rana family and other high Hindu caste families (Wood 1965; Bista 1991; Eagle 1999; Weinberg 2013). The

---

[21] Nepali is also called पार्बतिया <pārbatiyā>, which means 'belonging to the mountains or hills' (Cust 1878). The term गोरखालि <gorkhāli> started to be used by the Newah community (the oldest inhabitants of the Kathmandu valley) to refer to the language spoken by the Gorkhali invaders (Burghart 1996), the soldiers and inhabitants of the Gorkha kingdom (part of a confederation of 24 states known as Chaubisi Rajya ruled by King Prithivi Narayan Shah at the time of the expansion of his kingdom that led to the unification of Nepal which started in 1743 (Burghart 1984; Pradhan 1991).

[22] The first Nepali translation of the Sanskrit epic poem Rāmāyaṇa (रामायण) took place in the second half of the 19th century with the posthumous version of Bhanubhakta Acharya published in 1887 by Nepalese poet Motiram Bhatta.

[23] Nepāl Bhāṣā was also called Deśa Bhāṣā (Regmi 1961).

Panchayat period (1960-1990) that follows the Rana regime marks the promotion of a unique language, Nepali, and a unique religion, Hinduism, with the goal of unifying the nation as a Hindu society. A cultural and linguistic assimilation towards a unique Nepali identity began, illustrated by the slogan *Ek bhāṣā, ek bheṣ, ek dharma, ek des* meaning 'One language, one dress, one religion, one nation' (Gellner, Pfaff-Czarnecka & Whelpton 1997; Lawoti & Hangen 2012; Moronval 2017). 44.6% of the population reported Nepali as their mother tongue in 2011 (Nepal Census by CBS).

The Chepangs have been in contact with Nepali more intensively over the past 50 years, with the creation of the first national primary schools outside the Kathmandu valley in the 1960s[24]. The government established the National Education Board in 1953 and soon after that, the Nepal National Educational Planning Commission (NNEPC) advised by Dr. Hugh B. Wood, a Professor of Education at the University of Oregon (Wood 1965; Weinberg 2013). This commission determined the future of education in Nepal with the implementation of Nepali as the sole medium of instruction, supporting the development of monolingual Nepali speakers while pushing towards the decline of the use of native languages, as illustrated by the 1956 NNEPC report (1956: 96):

> It should be emphasized that if Nepali is to become the true national language, then we must insist that its use be enforced in the primary school…Otherwise, Nepali, though learned, may remain a "foreign" language rather than the child's basic, thinking language. Local dialects and tongues, other than standard Nepali, should be vanished from the school and playground as early as possible in the life of the child.

In the hills of Makawanpur and Chitwan, settlements of Nepali native speakers are rare. During the Panchayat period, Bahun-Chhetri (high Hindu castes) populations were living nearby the Chepang communities in the South of Dhading and Gorkha districts, and in the hills of the northeastern parts of Chitwan, and many abandoned their

---

[24] The first primary schools to open in their respective areas were in 1965 (2022 VS) in Silinge (RAK-6), and in 1969 (2026 VS) in Syamrang (RAP-13).

land at the end of the Panchayat period in the 1990s to settle in the plain of Tarai. The Chepangs who live in the hills have thus mainly been in contact with Tamang communities, in addition to the Gurungs, Magars, Bhujels, and Kamis.

The Tamangs, Gurungs, and Magars speak TH languages in addition to Nepali. The Kamis are native speakers of Nepali. The Bhujels (also referred to as Gharti Bhujel) who still speak their native language mainly live in the district of Tanahun, in areas inhabited by the Magars and Chepangs[25]. The Bhujels speak Bhujel (also referred to as Bhujeli or Gharti), the TH language the most closely related to Chepang. As noted above, Bhujel was considered by Caughley to be a dialect of Chepang (Caughley 1982: 1, 8). The Bhujels also use the native word Puhgal to refer to their language.

In the plain of Chitwan and Makawanpur, in addition to Tamang and Nepali, more languages are spoken, such as Tharu, Darai, Kumal, Majhi, Damai, Danuwar, Newah, Bhojpuri, and Hindi. These languages have had no noticeable influence on the Chepang language spoken by the communities living in the plain or in the hills. By contrast with the Chepang varieties spoken in the hills, that spoken in the plain has changed more drastically, having been remarkably affected by phonological and prosodic changes (§ 1.9.1.3). These changes are clearly the result of the influence of Nepali.

Map 6 represents language contact between Chepang and the four other main TH languages spoken in an area covering five districts (Chitwan, Makawanpur, Tanahun, Gorkha, and Dhading): Tamang, Magar, Gurung, and Bhujel. This map does not represent Indo-Aryan (IA) languages in contact with Chepang, such as Nepali, and does not say anything about the languages spoken outside the represented areas of contact. Finally, note that, while the four TH languages represented in Map 6 are spoken by populations with whom the Chepangs have mainly been in contact in the hills of the Mahabharat Range. The situation in the plain (the border between the hills and the plain being roughly represented by a black line on the map) is a little different, intermingling more language diversity, with communities speaking other IA and TH languages.

---

[25] Bhujel speakers are also present in some areas of Chitwan where there is no Chepang speakers. In Dhading, 20 speakers were reported in Benighat-Rorang Rural Municipality and in Gorkha, 15 speakers in Sundarbazar Municipality (Nepal Census 2011 by CBS).

Map 6. Language contact between Chepang and other TH languages
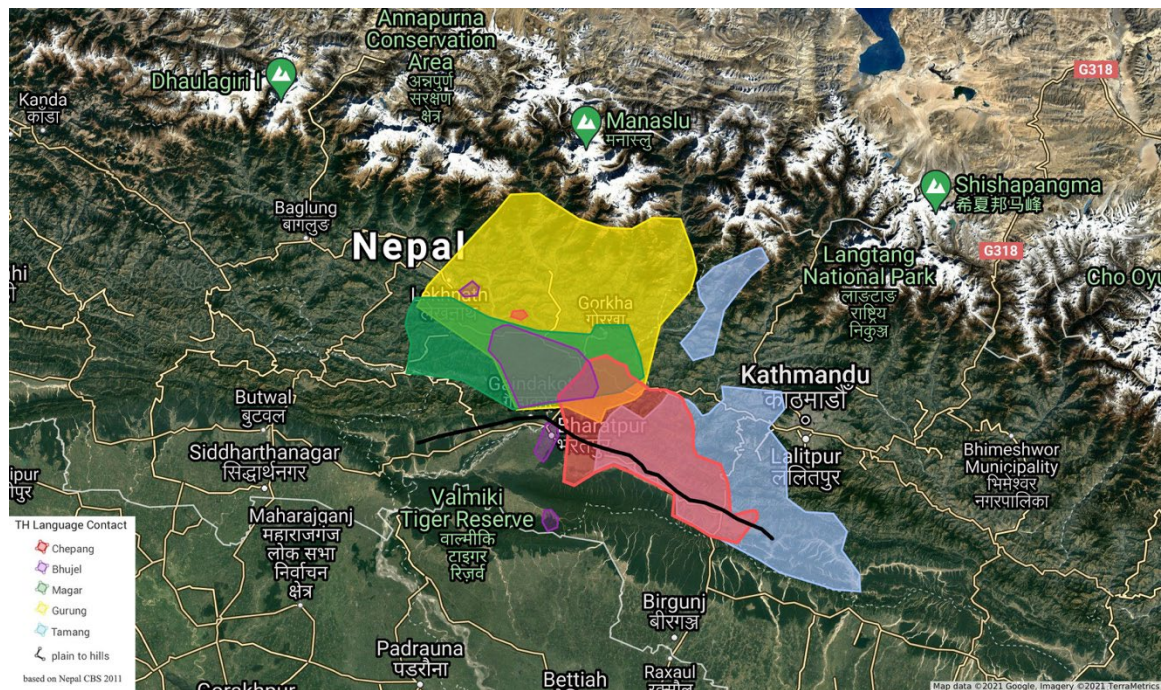
In Table 4, I present the number and percentage of Nepali and Chepang speakers in the four districts of Chitwan, Makawanpur, Dhading and Gorkha.

What can be observed is that Chepang speakers are more numerous in the districts of Chitwan and Makawanpur, representing 4.12% and 3.85% of the total population, respectively, by contrast with the districts of Dhading and Gorkha. Finally, the district of Makawanpur shows the lowest number of Nepali speakers, with 41.49% of the total population while they represent more than 70% of the population in Chitwan, Dhading and Gorkha.

These numbers are nevertheless insufficient to have a good understanding of the impact of Nepali on the Chepang language at the local level. This should be analyzed in a more dynamic way over time and space, by comparing for instance the numbers from 2011 with those that will come out from the next Nepal Census, and by comparing these numbers on a smaller scale, over small areas representing a few villages. This will allow us to see how the number of speakers of Chepang and Nepali have decreased or increased over time, and in which areas exactly the Chepangs thrive speaking their language with little to no contact at all with Nepali.

Table 4. Number of Chepang speakers in Chitwan, Makawanpur, Dhading, and Gorkha

Nepal CBS 2011	Total population	Chepang speakers' number	Chepang speakers' percentage	Nepali speakers' number	Nepali speakers' percentage
Chitwan	521,054	21,469	4.12 %	399,420	76.66 %
Makawanpur	415,601	16,004	3.85 %	172,422	41.49 %
Dhading	334,292	8,437	2.52 %	235,784	70.53 %
Gorkha	268,942	1,625	0.60 %	196,295	72.99 %
Total	**1,539,889**	**45,535**	**3.09 %**	**1,003,921**	**65.19 %**

The Chepang communities who live in the hills of the Lothar River basin have mostly been exposed to Tamang and Nepali languages. In some places, the Chepang population outnumbers the Tamangs and Kamis. In RAP-13, RAP-11, RAK-6, and RAK-7, there are many villages only occupied by the Chepangs, by contrast with Caughley's general observations (1982: 4–5).

I present the contact situations found in the Municipalities of RAP-13, RAP-11, RAK-6, and RAK-7 in Table 5, Table 6, Table 7 and Table 8, respectively.

This information was collected with Santosh Praja, resident of Rapti-13, Moti Lal Praja, resident of Rapti-11, Singh Lal Chepang, ward Chairman and resident of Raksirang-6, and Saphal Chepang, Chairman of Nepal Chepang Association of the Rāpti Municipality.

In these tables, the symbol ✓ marks which community is present in the village and the symbol + signifies which community outnumbers the other(s). The names of the villages are transcribed in the Chepang orthography based on the Roman alphabet proposed in § 2.11 since there are village names that are native to Chepang.

Table 5. Contact situations in Rapti-13 (RAP-13)

Rapti-13	Chepang	Kami
Dāngtes	✓	
Drungbāng	✓	
Gundi	✓	
Gāmerāng	✓	
Hātti Sudde	✓	
Kālitār	✓	
Kānrāng	✓	
Kumitār	✓	
Lo'ling	✓	
Mairāng	✓+	✓
Polkim	✓	
Rokrāng	✓+	✓
Sārling	✓	
Syāmrāng	✓	
Tāpāng	✓	
Thandānā	✓	
Wāsbāng	✓	
Yuiling	✓	

Table 6. Contact situations in Rapti-11 (RAP-11)

Rapti-11	Chepang	Tamang	Kami
Āmpāni	✓	✓+	
Barentār	✓	✓+	
Cisopāni	✓	✓+	
Dumrikhāri	✓	✓+	
Hātti Dāp	✓	✓+	
Jimling	✓		
Kālpāni	✓	✓+	
Kāmitār	✓	✓+	
Mudebās	✓	✓+	
Onibāng	✓+	✓	
Payāk	✓+	✓	
Sarkundi	✓+	✓	
Satisāl	✓	✓+	
Simal Dādā	✓		
Thākaltār	✓	✓+	✓

Table 7. Contact situations in Raksirang-6 (RAK-6)

Raksirang-6	Chepang	Tamang	Kami
Aynātār	✓	✓+	
Bānge	✓+	✓	
Bāngrāng	✓	✓+	
Bhuibisāuna		✓	
Bujarāng	✓	✓+	
Cāpala	✓		
Citure	✓		
Cyorāng	✓+	✓	
Dungthali		✓	
Dusarāng	✓		
Gidarbāng	✓		
Jāsetār		✓	
Jyārāng	✓+	✓	
Karāse	✓+	✓	✓
Khāmbadhani	✓	✓+	
Kharkande	✓		
Kwitan	✓		
Māgini	✓		
Mākaldamār	✓	✓+	
Pālāse	✓		
Sidhām	✓+	✓	
Silinge	✓+	✓	✓
Silādhuni			✓
Wālgādi	✓+	✓	

Table 8. Contact situations in Raksirang-7 (RAK-7)

Raksirang-7	Chepang	Tamang
Āysarāng	✓	
Dagabāng	✓	
Dambarāng	✓	
Darāng	✓	
Dhirāng	✓	
Gāibāng	✓	
Garling	✓	✓+
Kharling	✓	✓+
Lāngkā	✓	
Rinsarāng	✓	
Tiruwā	✓	
Yomkacār	✓	

I observed that in RAK-6, where language contact with Tamang communities is the greatest of the four areas presented here, the Tamangs fluently speak Chepang, while the Chepangs do not necessarily speak Tamang unless there had been marriage alliances with the Tamangs in someone's family, leading to bilingualism. This may reflect a social dominance of the Chepang community over the Tamangs in these areas, and contradict the general view that the Chepangs have been considered as an inferior caste by the Tamangs (Jest 1966: 178). The idea of the existence of superior and inferior ethnic identities was imposed by the high Hindu castes' ideological principles in Nepal at the beginning of the 19th century. This was effected by the establishment in 1854, of a caste system codifying ethnicity by Prime Minister Jang Bahadur Rana through the Muluki Ain (Legal Code) (Höfer 1979; Levine 1987; Gellner, Pfaff-Czarnecka & Whelpton 1997). The Chepangs, along with other communities, like the Bhujels (or Ghartis), Tibetans, Hayus, and Tharus were categorized as 'enslavable alcohol drinkers (*māsinyā matvāli*),' below the Tamangs, Gurungs, Sunwars, and Magars, classified as "non-enslavable alcohol drinkers (*namāsinyā matvāli*)" (Regmi 1977; Höfer 1979: 44–45).

Caughley (1982: 4) suggests that the Tamangs may have come from the North to settle in areas formerly occupied by the Chepangs since the Tamangs "have had very little linguistic or cultural influence on the Chepangs." Caughley (1982: 4) states that this contact situation between the Chepangs and the Tamangs may not go back further than "a few generations." The question of the linguistic and cultural influence of the Tamangs over the Chepangs, and inversely of the Chepangs over the Tamangs, in addition to the timespan within which both the Chepangs and Tamangs have been living close to each other, remains to be explored. Indeed, they may not be of the same type. While the Tamang language has had little to no influence on the Chepang language – beside a few attested borrowings of proper nouns (§ 3.3.5), the Tamangs and the Chepangs are both represented in shared narratives that relate their respective animistic and shamanistic beliefs.

In Chepang and Tamang oral literature that was transmitted to us by several Chepang and Tamang shamans of RAP-6 (ancestors coming from Dhading), RAK-6, and RAP-13, one can find narratives that trace both the origin of shamanism and that of the people themselves to two brothers of a same line of descents: Tungsuri Pan (guru of Chepang shamans) and Urgyen (guru of Tamang shamans). Tungsuri Pan (also referred to as Tungsuri Ban or Tungsuri Bon) and Urgyen decided to see who was the fastest to reach both the realm of the skies, called *lāngkā*, and the underworld, called *patāl*. When Urgyen realized that Tungsuri Pan had almost reached *lāngkā*, he went faster and hit Tungsuri Pan's drum. From then on, only one membrane remained on a Chepang shaman's drum instead of two.

While the accounts of their actions present some variation, they show how Tungsuri Pan and Urgyen have together changed the course of Chepang and Tamang shamanic practices, such as: the loss of one membrane on the Chepang shamans' drums as a result of Ugryen's anger; access to the underworld (*patāl*) denied to the Tamang shamans by Tungsuri Pan as a revenge of losing a drum's membrane; division of healing and funeral practices to two different authorities for the Tamangs, i.e., *bompo* and *lāmā*, respectively, while both healing and funeral practices can be carried out by the Chepang shaman. Versions of this story and additional stories were also collected by Riboli

(1994b; 2000). For instance, Riboli (2000: 86) further reports that a Tamang shaman became shaman through the guidance of Chepang divinities in their dreams.

The division of shamanic domains between the Chepangs and the Tamangs, the fact that the Chepang shaman can travel in the *patāl*, i.e., the underworld, realm of divinities, while the Tamang shaman only travels in the skies, another realm of divinities to which the Chepang shamans have also access is also reported by Riboli (2000: 85–87). Attempts to diminish the power of a shaman by destroying their drum's membrane are reported in shamanic myths of other communities, such as the Tungusic people of Siberia (Lot-Falck 1961: 26).

There are nevertheless some differences between what we found and some of Riboli's accounts (1994b; 2000) regarding, for instance, the names and roles of Tungsuri Pan and Urgyen. Riboli (2000: 10) uses the term *Tunsuriban* for Tungsuri Pan[26], and *Urghsuriban* for Urgyen. She further considers that the former represents all Chepang shamans and the latter all Tamang shamans. However, they are said to be shaman gurus (spiritual teachers) of Chepang and Tamang shamans, respectively. They live in the jungle, separately from the world of humans, in the realm of the divinities and spirits, and may be also called *ban jhānkrī* 'jungle shaman.' As gurus, they are basically the shamans from whom Chepang and Tamang shamans receive their training in their dreams.

These accounts reflect a somewhat deeper connection between the Tamangs and the Chepangs than that presupposed by Caughley (1982: 4). The Chepangs have been in contact with the Tamangs for a long time, and although exogamy does not commonly happen between them, it is likely that the Tamangs settled into Chepang territory coming from the North, as suggested by Caughley (1982: 4) and anthropologists (Holmberg 2005; Tautscher 2007).

---

[26] The variation between Ban, Pan and Bon, may be analogical, since Ban /bʌn/ means 'jungle,' where these gurus live, and Bon /bon/, Bonpo or Bompo is used in Tamang shamanism and Buddhism to refer to the spiritual authority.

1.11.   Methodology and language data

This section lays out the methodology used to collect and analyze the data in this study. I first describe the journey towards collaborative research that allowed the Chepang Language Documentation and Description Project (CLDDP) to take place over a four-year period, from Fall 2017 to Fall 2021 (§ 1.11.1). I then present the digital corpus on which the present description is based (§ 1.11.2). Finally, I show what I set up to give access to the documentation to the Chepang community and beyond (§ 1.11.3).

Within this four-year period, I spent a total of 26 months (2 years and 2 months) in Nepal with Chepang community members of Chitwan and Makawanpur districts. When I was not physically in Nepal, I was working on the analysis of the Chepang data while in constant contact with the community members on a daily basis.

When in Nepal, I spent my time sharing the daily life of the Chepangs, exploring their cultural and spiritual knowledge and practices, learning, recording, transcribing, and analyzing the language, while compiling a dictionary. I worked on my relationship with the Chepang community members, grew and strengthened my network, and developed the project towards collaborative research. In December 2019, I started to constitute and train a team of native speakers who were interested in participating in the project at a higher degree of involvement and responsibility.

1.11.1. Methodology and teamwork

Throughout this research, I concentrated my efforts on developing a collaborative research environment with Chepang community members, giving native speakers the opportunity to engage in leading the documentation and description of their language, carrying the tasks and responsibilities that have traditionally been carried out by a linguist.

The project's journey started in the Rural Municipality of Manahari-4, Makawanpur in Fall 2017 with Chepang community members of seven villages scattered along the Handikhola River basin, at the border with the Parsa Wildlife Reserve, now known as Parsa National Park. This first part of the journey was dedicated to meeting Chepang community members interested in being involved in the project, starting eliciting basic vocabulary and morphosyntactic data, and recording texts, in addition to learning how to speak the language.

The second part of the journey continued in the hills of Makawanpur, in the Rural Municipalities of Raksirang-6 and Raksirang-7, where I worked with community members of five villages. I continued to develop the collaborative aspect of the project, while further investigating the linguistic aspects of the language. I started giving basic linguistic training to a community member of Silinge (RAK-6), Bipana Chepang, in order to work together on the transcription and analysis of the recordings. She also took on the lead in interviewing community members to record conversations. To work with electricity and Internet, I set up an office space in the village of Bhandara situated in the Municipality of Rapti-5, in the plain of Chitwan, at the crossroads from where departures to the hills of Makawanpur and Chitwan take place. There, we kept working on the transcription and analysis of the recordings.

During this period, I travelled to record more speakers from various places, such as Maisirang, in the Rural Municipality of Raksirang-8, where Caughley based his research on Chepang. I also investigated the language spoken by the Chepangs living in the area of Bhandara, and in particular along the Rapti River (RAP-6) at the border with the buffer zone of the Chitwan National Park. Several workshops were held with the

elders and local shamans to work on revising the collected vocabulary related to the flora and fauna, natural phenomena, and shamanism.

I then came to better understand the main linguistic features that differentiate and unite the varieties spoken in the plain and the hills, and how these may have evolved (§ 1.9). I also learned about the migration dynamics from the hills to the plains of Chitwan and Makawanpur that occurred in the past 50 years, the historical connections that exist between the Chepangs living in the basin of the Handikhola River and along the Rapti River, whose roots go back to the same places in the hills, be it RAK-6, RAP-13, or Talti in the South of the district of Dhading. In addition, I also learned about the ancient paths that the Chepangs would take to trade, such as the one along the Lothar River to come down to the plain to buy salt, rice and other products unavailable in the hills, while selling their own products, such as Indian butter tree oil, etc.

The third part of the journey extended the exploration of the Chepang language varieties to the hills of Chitwan, in five villages of Rapti-13 Rural Municipality and three villages of Rapti-11 Rural Municipality.

During the project, I trained six native speakers (between 19 and 35 yo [year old]) in basic recording and language data processing, in conducting interviews with community members, and in lexicography to revise and extend the dictionary while including language variation.

This last period strengthened the collaborative side of the work, with a team constituted by six community members (19-35 yo) trained on different aspects of the documentation and description of Chepang: leading interviews with native speakers; video recording; basic command of ELAN; transcribing and reading the International Phonetic Alphabet (IPA), in addition to reading other transcription modes used by Caughley in his dictionary (a mixed IPA-Roman form and a Devanāgarī form); basic phonetic analysis (focusing on the realizations and positions of the glottal stop and glottal fricative); basic morphological analysis (distinguishing parts of speech, word segmentation, nominal and verbal inflectional morphology).

The team members were from the two main geographical areas where most of the investigation of Chepang has taken place for this project, i.e., Raksirang-6 and Rapti-13, which represent the varieties of Lothar: Dinesh Praja (Tapang), Pan Maya Praja (Gundi),

Prakash Praja (Tapang), and Santosh Praja (Polkim) from RAP-13; Pabitra Chepang (Silinge) and Samjhana Chepang (Khadkare) from RAK-6.

Our team work focused on documenting more language varieties through the recording of natural conversations and traditional narratives, revising multiple times our 3,000+ word lexicon while acknowledging variety, cross-checking it with the vocabulary found in Hodgson (1848; 1857a; 1874a), Caughley et al. (1969; 1970), Caughley (1969; 1970a; 1972; 2000; 2016), Weidert (1987), and Adhikari (2016; 2017), in addition to segmenting, transcribing and translating audio and video recordings in ELAN, and revising earlier transcribed materials.

1.11.2. Digital corpus

The ELAN digital corpus was built in collaboration with the Chepang native speakers involved in the project to allow the description and analysis of the Chepang language. As of June 2022, it consists of over 25 hours of audio and video material: 261 recordings that include texts (narrative and expository), conversations, elicitation sessions (monolingual in Chepang), and songs (folk and shamanic chants).

This corpus presently counts over 138,500 words transcribed and translated in ELAN. The translations are either in Nepali, or in both English and Nepali. The metadata of the speakers who contributed to the realization of the recordings that constitute the digital corpus is given in APPENDIX I. The metadata of the recordings that constitute this digital corpus is given in APPENDIX II.

1.11.3. Documentation and access

Giving access to the collected documentation is important for the empowerment of the Chepang communities at various levels. It allows sustainable preservation and strengthens the transmission of the documented knowledge of the communities. It provides the younger generations with a different manner of rediscovering, learning, and sharing the knowledge passed on by the elders via modern technology, to the point of

envisioning new socio-economical perspectives to sustain their life using their language (such as the practice of journalism in Chepang via online video-sharing platforms). It spreads and consolidates the use of two new orthographies (Devanāgarī and Roman-based) developed through the collaborative work with Chepang community members of Manahari, Raksirang and Rapti Municipalities in Makawanpur and Chitwan.

The collected documentation will be deposited and made more broadly accessible as well in a sustainable archive such as DOBES at The Language Archive of the Max Planck Institute[27].

It is almost impossible for community members to access archived materials, since many do not have email or access to computers. Therefore, I created a website[28] linked to a YouTube channel[29] where people can easily access and download the recordings via their mobile phones. This contributes to preserving and spreading the language documentation to community members and inspires people to continue to record their elders and to participate in the preservation of their culture while empowering the use of their language.

1.12.   Organization of the dissertation

This dissertation is organized in five chapters. The first has been the introduction.

The second chapter provides a detailed description of the phonological system of Chepang attested in most varieties while acknowledging variation. This covers syllable structure, consonants, monophthong vowels and diphthongs. It discusses attested phonemic variations and provides a tentative reconstruction of Proto-Chepang (PC) phonology. An emphasis is given to the analysis of the surface realizations of the final and sonorant pre-final glottal stop. Finally, this chapter includes an overview of the prosodic system of the language.

---

[27] https://dobes.mpi.nl

[28] tinyurl.com/thechepanglanguage

[29] https://www.youtube.com/channel/UC-DHyr9p3DN5xlkiJLC91Rw

Chapter three is dedicated to nouns and nominal morphology. After defining what a noun is, by contrast with a verb, the chapter divides into two main parts: the first focuses on nominal formation processes and the second on nominal morphology. The nominal formation processes attested in Chepang are described through both a synchronic and diachronic perspective, seeking to understand how certain words or parts of words came to mean what they mean. This first part also includes a detailed description of the kinship terms used in Chepang and a list of cognate forms found in other TH languages. The morphological study provides an analysis of the inflectional and derivational morphemes attested with nominal roots.

The fourth chapter presents the pronouns and determiners used in Chepang. Beyond describing their use, it seeks to acknowledge their origin and development.

Chapter five describes verb formation and verbal morphology through both a synchronic and diachronic approach. It sheds light on the derivational functions of ancient morphology and describes the non-canonical Direct-Inverse system of Chepang as the result of historical development and pragmatic triggers in language use.

CHAPTER II

PHONOLOGY

This chapter provides a segmental phonological and phonetic description of the phonemes attested in Chepang: 38 consonants and 6 vowels, as represented in Table 9.

Table 9. Consonant and vowel phonemes of Chepang

consonants								
p	b		t	d			k	g
pʰ	bʰ		tʰ	dʰ			kʰ	gʰ
		ʈ	ɖ		tɕ	dʑ		
		ʈʰ	ɖʰ		tɕʰ	dʑʰ		
m			n				ŋ	
mʰ			nʰ				ŋʰ	
	l		ɾ		j		w	
	lʰ		ɾʰ		jʰ		wʰ	
	s	ʃ				h	ʔ	

vowels		
i		u
e		o
	ʌ	
	a	

The present phonological analysis is based on a lexicon of 3,129 words, collected through elicitation and corpus data over a period of 4 years with Chepang community members of different villages of Makawanpur and Chitwan districts (§ 1.11). This lexicon has been verified and revised multiple times with native speakers of the different studied varieties at various stages of the project. They have also been compared with the Chepang vocabulary found in Hodgson (1848; 1857a; 1874b), Bandhu et al. (1969; 1970), Weidert (1987), Caughley (2000, 2016), and Adhikari (2016; 2017), with which they may differ.

The following descriptions of the phonological segments include details about attested variation in all studied varieties of Chepang. Reference to the vocabulary collected by Caughley (2000; 2016) and to earlier analyses of the phonology is made when appropriate. Previous accounts of Chepang phonological and suprasegmental features include Caughley (1969; 1970a; 1970b; 1980; 1982), Bandhu et al. (1969; 1970), Weidert (1987) and Adhikari (2016; 2017).

A significant difference between the present analysis and earlier ones is that breathy and aspirated consonants were previously considered to be the combination of the glottal fricative /h/ and the consonant. Thus, we find for instance 18 consonants in Caughley (1969; 1970a; 1982; 2000; 2015; 2016; 2016). Only the analysis found in Bandhu et al. (1970) proposes 24 consonants, including breathy sonorants, which were not characterized as breathy, though these were not characterized as breathy but as voiceless. Yet, other breathy or aspirated consonants were not considered as single phonemes in this account. Finally, Adhikari (2016; 2017) proposes 25 consonants, including breathy and aspirated stops, but excluding breathy sonorants and the glottal stop. None of these accounts includes the retroflex stop and post-alveolar fricative consonants, a choice likely motivated by the fact that they are not native to Chepang.

I start by giving an overview of the full inventory of consonant and vowel phonemes found in the language, specifying place and manner of articulation, and distinguishing native phonemes from non-native ones (§ 2.1). Before examining the phonemic and phonetic details of the realizations of the segments, I describe the syllable structure (§ 2.2), since parts of the phonological analysis are tied to the analysis of syllable structure. I then give a detailed description of the consonants (§ 2.3) found at

syllable onset (§ 2.3.1) and coda (§ 2.3.2), along with their allophones in free variation (not conditioned). I show that the laryngeal or glottal segments /h/ and /ʔ/ are in paradigmatic distribution, forming a complex coda in combination with sonorant consonants (§ 2.3.3). While they are part of the consonant phoneme inventory of the modern language, retroflex stop and post-alveolar fricative consonants essentially occur in loan words borrowed from Nepali (IA). I outline some patterns in the treatment of borrowed consonants since it is found that they can sometimes be nativized but not always (§ 2.3.4). I describe the native consonant clusters at syllable onset (§ 2.3.5). I present the morphophonological processes that affect consonants (§ 2.3.6). I propose a reconstruction of Proto-Chepang (PC) consonants and consonant clusters based on the studied varieties (§ 2.4). Monophthong vowels are described and illustrated with vowel charts representative of their pronunciation by 3 female and 1 male speakers from the Lothar variety (§ 2.5). I describe the morphophonological processes that affect the monophthong vowels (§ 2.5.1). I examine the realization and combinations of the sequences formed by a vowel and an approximant or glide, or by two vowels, and discuss their distribution with regard to the syllable structure (§ 2.6). These sequences are also affected by morphophonological processes (§ 2.6.7). Finally, I lay out all the rhymes attested in Chepang with simple and complex nuclei followed by simple and complex codas (§ 2.7). All the descriptions of the phonemes are illustrated with minimal phonemic sets of native roots (otherwise specified), supplemented with waveform and spectrogram figures when appropriate.

Amongst the consonant phonemes of Chepang, the glottal stop /ʔ/ requires special attention (§ 2.3.3). With the glottal fricative /h/ and other consonants, the glottal stop /ʔ/ shows traces of ancient verbal derivational morphology (5.2); amongst the variations attested in the realization of the glottal stop /ʔ/, one is involved in tonogenesis; finally, historically, the question of the reconstruction of /ʔ/ at the Proto-Himalayan (PH) level is to be discussed with regard to the different phonemic statuses that its cognates, if any, hold in other Himalayan languages, such as Hayu (Michailovsky 2003), Limbu (Driem 1987), Belhare (Bickel 2003), Sunwar (Borchers 2008), Yakkha (Schackow 2015) and Khaling (Jacques 2016b).

Finally, I propose the use of two possible orthographies in § 2.11: one based on Devanāgarī and one on the Latin or Roman alphabet.

2.1. Consonant and vowel inventories

Chepang distinguishes 38 consonant phonemes. All but the glottal stop are phonemic at syllable onset (§ 2.3.1), while a set of 13 native phonemes restricted to unaspirated voiceless stops and non-breathy sonorants appears at syllable coda (§ 2.3.2). The consonant phoneme inventory in Table 10 presents both the native and non-native consonants of Chepang. The non-native consonants, noted in italic, are retroflex and post-alveolar fricative consonants.

Table 10. Consonant phoneme inventory

	bil.		alv.		post-alv.	retr.		alv-pal.		pal.	vel.		gl.
stop	p	b	t	d		t	$ɖ$				k	g	ʔ
asp. or breathy	p^h	b^h	t^h	d^h		t^h	$ɖ^h$				k^h	g^h	
fricative			s		ʃ								h
affricate								tɕ	dʑ				
asp. or breathy								$tɕ^h$	$dʑ^h$				
nasal		m		n								ŋ	
breathy		m^h		n^h								$ŋ^h$	
flap				ɾ									
breathy				$ɾ^h$									
lat. approximant				l									
breathy				l^h									
approximant		w								j			
breathy		w^h								j^h			

The vowel phoneme inventory distinguishes 6 monophthongs. They are presented in Table 11. The present analysis does not recognize the existence of diphthongs in Chepang. I describe the combination of a glide and vowel and two vowels in § 2.6.

Table 11. Vowel phoneme inventory

	front	central	back	
	unrounded	unrounded	unrounded	rounded
close / high	i			u
close-mid	e			o
open-mid			ʌ	
open / low		a		

2.2. Syllable structure

I present the syllable structure of Chepang before describing the phonemic segments, since their description will at times refer to their distribution within the syllable. The structure of a native canonical syllable in Chepang can be represented linearly, as in Figure 7, or hierarchically, as in Figure 8. Optional segments are presented in parentheses.

Figure 7. Chepang canonical linear syllable structure

$$\sigma = \quad (C_1) \quad (L/w) \quad V_1 \ (G/V_2) \quad (Lg) \quad (C_2)$$

where   and

$C_1 = [\text{-syll}]$ C = consonant

$V_1 = [\text{+syll}]$ V = vowel

$C_2 = [\text{-syll}]$ L = liquid /l/ or /ɾ/

                V = vowel

                G = glide

                Lg = laryngeal

Figure 8. Chepang canonical hierarchical syllable structure

```
                              σ
                  /                       \
              onset                      rhyme
             /      \              /        |         \
        initial  clustering   nucleus          pre-final   coda
           |         |          /   \             |          |
         (C₁)      (L, w)      V₁   (G/V₂)        (Lg)       (C₂)
```
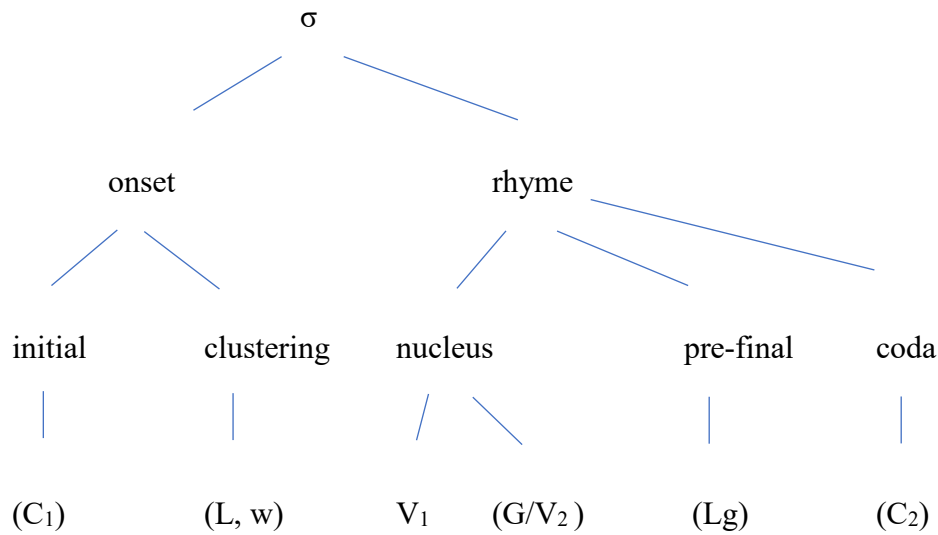
A syllable minimally consists of a vowel nucleus: it can form a rhyme by itself. However, few morphemes feature a syllable whose rhyme only consists of a vowel nucleus. Besides the demonstratives or 3rd person pronouns (*i, o, u*), only two other roots are found; they are bound roots: *e-* 'chestnut (*Castanopsis indica*)', and *a-* 'foxtail millet (*Setaria italica*).' These bound roots *e-* and *a-* need a suffix classifier specifying their category in order to form a word. For instance, the suffix *-si* occurs with *e-* to refer to the category of 'tree:' *e-si* 'chestnut tree (*Castanopsis indica*);' the suffix *-jam* occurs with *a-* to refer to a type of 'grain of cereal' *a-jam* 'foxtail millet (*Setaria italica*)' can be formed.

Apart from consisting of a single vowel nucleus, the rhyme can end with a simple coda (C₂), restricted to unaspirated voiceless stops and non-breathy sonorants, or in a complex coda where sonorants may be preceded by a laryngeal phoneme (Lg): a fricative glottal /h/ or glottal stop /ʔ/. The optional onset can be formed of a single initial consonant or a consonant cluster that consists of a limited set of initial consonants followed by a medial consonant: a liquid /l/ or /ɾ/ (L), or a bilabial approximant /w/ (w) (§ 2.3.5). The cluster formed with the bilabial approximant /w/ is restricted to follow the velar stop consonants /k/ and /g/. The clusters /kw/ and /gw/ are analyzed as possible reflexes of the PTH labiovelar consonant phonemes */kʷ/ and */gʷ/ or clusters /*kw/ and /gw/ (§ 2.3.5.5). The combination of /k/ and /g/ with /w/ could in fact be synchronically

analyzed as two labiovelar consonant phonemes /k^w/ and /g^w/. However, in order not to complicate the development of an orthography, I chose to consider /w/ as part of a consonant cluster rather than two additional labiovelar phonemes.

In addition to palatal and bilabial glide consonant phonemes at onset and coda, the proposed syllable structure accounts for an optional second vowel or glide in the nucleus. Structurally, the presence of glide consonants at onset (C_1) or coda (C_2) and the possibility of a complex nucleus formed by an additional vowel or glide entails that a syllable structure of the type $C_{[+glide]}$ V or V $C_{[+glide]}$ could be equivalent to that of a V_1 G/V_2 type syllable with no initial nor coda consonant. I have analyzed glides as independent consonant phonemes since their distribution shows that they are not merely allophones of the high vowels /i/ and /u/ gliding at onset or coda. There are three main reasons for that: first, two series of glides exist at onset (breathy vs. non-breathy); second, glide consonants at coda can be preceded by a laryngeal consonant; and third, a V_1 G/V_2 type of nucleus can also be preceded by a glide consonant at onset or be followed by a glide consonant at coda. Therefore, it is necessary to distinguish the glides /j/ and /w/ from the vowels /i/ and /u/ phonemically. Further details relative to the distribution of glides and vowel-glide or vowel-vowel sequences are presented in § 2.6. In addition, note that the linear or hierarchical structure represented by V_1 (G/V_2) cannot account for the fact that the first element (V_1) can also hold the position of the glide in the sequence. In other words, I described this sequence as V_1 (G/V_2) but it could as well be (G/V_1) V_2. Finally, this syllable structure implies that the concept of diphthong does not really apply, since it would entail that a glide consonant present in the nucleus would be phonemically different from a glide in syllable initial and coda position. In fact, when it comes to talking about the nucleus consisting of more than a single vowel, I will call it a vowel-glide sequence, or a two-vowel sequence.

I present the possible syllabic structures that apply for the rhyme in Figure 9, and for the onset-rhyme in Figure 10. These are illustrated with corresponding examples of roots.

Figure 9. Rhyme syllabic structures

rhyme	(C₁)	(L,w)	V₁	(G/V₂)	(Lg)	(C₂)			
V₁			e				*e-*	[e]	'chestnut (castanopsis indica)'
V₁C₂			a			p	*ap-*	[ap]	'shout at target'
			a			l	*al-*	[al]	'go'
			a			j	*aj*	[aj]	'mother-in-law
			ʌ			w	*ʌw-*	[ʌw]	'cease (rain)'
V₁G/V₂C₂			u	i		ʔ	*uiʔ-*	[ʉiʔ]	'whistle slowly to attract fish, crabs'
			o	j		k	*ojk*	[ojk]	'flour made of roasted corn' (arch.)
			ʌ	j		ŋ	*ʌ̃jŋ-*	[ʌ̃jŋ]	'hurt after shock in the back (heart)'
V₁LgC₂			o		ʔ	j	*oʔj-*	[oʔj]	'be friable, or floppy, not rigid'
			a		h	m	*ahm*	[ahm]	'porridge, cooked cereal'

Figure 10. Onset-rhyme syllabic structures

onset-rhyme	(C₁)	(L,w)	V₁	(G/V₂)	(Lg)	(C₂)			
C₁V₁	ɾ		o				*ɾo*	[ɾo]	'flower (gen.)'
	m		a				*ma*	[ma]	'yes'
	j		u				*ju-*	[ju]	'melt, dissolve'
	w		a				*wa*	[wa]	'bird (gen.)'
C₁V₁V₂C₂	l		j~e	a		ŋ	*ljaŋ-*	[leãŋ]	'Himalayan raspberry'
	b		j~e	a		w	*bjaw-*	[beaw]	'appear in number (bird)'
	p		j~e	a		k	*pjak*	[peak]	'pig'
	ɾ		u	i		ŋ	*ɾuiŋ*	[ɾʉĩŋ]	'bamboo'
	k		u	e		ʔ	*kueʔ*	[kʉeʔ]	'fish hook'
	j		u	i		n	*juin*	[jʉin]	'bat'
C₁LwV₁	m	ɾ	o				*mɾo-*	[mɾo]	'recover (from illness)'
	k	w	i				*kwi*	[kwi]	'dog'
C₁LwV₁V₂	k	ɾ	u	i			*kɾui-*	[kɾʉi]	'be tangled, curly (hair)'
C₁LwV₁V₂C₂	b	ɾ	a			w	*bɾaw-*	[bɾaw]	'be big, tall'
C₁LwV₁C₂	g	w	e			j	*gwej*	[gwej]	'yam, taro (gen.)'
C₁V₁C₂	ɾ		u			ʔ	*ɾuʔ*	[ɾuʔ]	'fish poison'
	m		a			j	*maj*	[maj]	'flesh, meat'
C₁V₁LgC₂	ɾ		ʌ		h	m	*ɾʌhm-*	[ɾʌhm]	'heal, dry (wound)'
	d		a		ʔ	l	*daʔl*	[daʔl]	'earthworm'
C₁LwV₁LgC₂	k	l	o		ʔ	n	*kloʔn-*	[kloʔn]	'submerge (river bank)'
C₁LwV₁V₂C₂	b	l	a	j		k	*blajk-*	[blajk]	'feel bored, lazy'
	k	ɾ	j~e	a		p	*kɾjap-*	[kɾeap]	'cry'
C₁LwV₁V₂LgC₂	g	l	j	u	ʔ	m	*gljuʔm*	[gljuʔm]	'charcoal, coal'
	p	ɾ	j~e	a	h	j	*pɾjahj-*	[pɾeahj]	'shuck, shell (corn seeds)'

2.3.   Consonants

This section describes native consonants at syllable onset (§ 2.3.1), syllable coda (§ 2.3.2), the distribution of the glottal segments in complex codas (§ 2.3.3), the realization of non-native consonant phonemes (§ 2.3.4), consonant clusters attested in initial position (§ 2.3.5), and the morphophonology of consonants (§ 2.3.6).

2.3.1.   Consonant onsets

All the consonants presented in Table 12 are phonemic at syllable onset. By contrast with Table 10, the consonant onsets in Table 12 are all native to Chepang. Allophones presented here occur in free variation, and sometimes, more than one allophonic realization of a phoneme can be found for a single individual. These allophones can be found with all roots featuring the onset consonant triggering allophony. Some allophones, also attested in free variation, are nevertheless restricted to certain roots. While these allophones are described in the following descriptions of the phonemes, I did not include them in Table 12. Native phonemic consonant onsets and their allophones are presented in Table 12.

Table 12. Native phonemic consonant onsets and allophones in free variation

	bil.		alv.		alv-pal.		pal.	vel.		gl.
stop	p	b	t	d				k	g	
asp. or breathy	pʰ	bʰ	tʰ	dʰ				kʰ	gʰ	
allophone	[ɸ]	[β]						[ᵏx]	[ᵍɣ]	
fricative			s							h
affricate					tɕ	dʑ				
allophone					[ts]	[dz]				
asp. or breathy					tɕʰ	dʑʰ				
allophone					[tsʰ]	[dzʰ]				
nasal		m	n						ŋ	
breathy		mʰ	nʰ						ŋʰ	
flap			ɾ							
breathy			ɾʰ							
lat. approximant			l							
breathy			lʰ							
approximant		w					j			
breathy		wʰ					jʰ			

2.3.1.1.   Bilabial stop onset

Four bilabial stops are phonemic at syllable onset. They are distinguished by voicing and aspiration or breathiness. This is illustrated in Table 13.

The bilabial stops /p/ and /b/ can be fricativized (or spirantized) to [ɸ] and [β] respectively. They are allophones in free variation.

Table 13. Minimal set for bilabial stop onsets

p		*pe-*	'be nice, good'	**b**		*be-*	'be thin'
pʰ [pʰ]~[ɸ]		*pʰe-*	'leave behind, quit'	**bʰ** [bʰ]~[β]		*bʰe-*	'separate (inside family)'

2.3.1.2. Alveolar stop onset

Alveolar stops contrast in voicing and aspiration or breathiness. This is illustrated in Table 14. They do not present any specific allophones.

Table 14. Minimal set for alveolar stop onsets

t	*tuŋ-*	'drink'	**d**	*duŋ-*	'sprout'
tʰ	*tʰuŋ-*	'knock against'	**dʰ**	*dʰuŋ*	'termites larvae's nest'

2.3.1.3. Velar stop onset

Velar stops distinguish as well voicing and aspiration or breathiness. This is illustrated in Table 15. Fricativization can sometimes occur with both /kʰ/ and /gʰ/, but when it does, a velar burst is still perceptible. The velar fricative allophones are thus not fully fricativized and pronounced [ᵏx] and [ᵍɣ]. These allophones are used in free variation.

Table 15. Minimal sets for velar stop onsets

k		*ka-*	'put inside, insert'	**g**		*ga-*	'open mouth'
kʰ [kʰ]~[ᵏx]		*kʰaŋ-*	'heat, boil (milk, water)'	**gʰ** [gʰ]~[ᵍɣ]		*gʰaŋ*	'hole'

2.3.1.4. Fricative onset

Fricative onsets are the voiceless alveolar /s/ and glottal /h/, as shown in Table 16.

Table 16. Minimal set for fricative onsets

s	*saj-*	'grow (plant)'	**h**	*haj-*	'do'

2.3.1.5.　　　Affricate onset

Affricates in Chepang are alveolo-palatal. They have a four-way contrast in voicing and aspiration/breathiness: /tɕ/, /tɕʰ/, /dʑ/, /dʑʰ/. The alveolar allophones [ts] and [dz], and [tsʰ] and [dzʰ] are attested in free variation. They are illustrated in Table 17.

Table 17. Minimal sets for affricate onsets

tɕ [tɕ]~[ts]	*tɕok-*	'grow (plant, tree)'	**dʑ** [dʑ]~[dz]	*dʑok-*	'be fast'
tɕʰ [tɕ]~[tsʰ]	*tɕʰak-*	'pile up (things)'	**dʑʰ** [dʑʰ]~[dzʰ]	*dʑʰak-*	'love, like, be pleased'

Caughley (2015: 34, 36) categorizes Chepang affricates as palato-alveolar and transcribes them as <c> and <j> while using [ţš] and [ḓž] in phonetic transcriptions, later changed to [tʃ] and [dʒ] in a revised version (Caughley 2015). The reason behind categorizing them as alveolo-palatal rather than palato-alveolar is that they are more palatal than English palato-alveolar affricates (/tʃ/, /dʒ/) or French fricatives (/ʃ/, /ʒ/). They are similar to Nepali in this regard. Both alveolo-palatal and palato-alveolar affricates are post-alveolar sounds realized with the tongue touching the alveolar ridge. While palato-alveolar affricates may be both apical or laminal, that is to say, realized with either the tip or the blade of the tongue touching the alveolar ridge, alveolo-palatal affricates are described as essentially laminal in addition to be more palatalized (Ladefoged & Maddieson 1996: 90, 150).

When not alveolar, affricates in TH languages spoken in Nepal are often categorized as palato-alveolar. This is also the case in IA languages, such as Nepali. Masica (1991:94) notes that affricates in IA are nevertheless different from English palato-alveolar affricates, pronounced with the "blade touching the hard palate" and suggests that a different transcription should be used. Acharya (1991: 22) uses the term "palatal" to describe affricates in Nepali. Only two Himalayan languages, Hayu (Michailovsky 1988; 2003: 519) and Limbu (Driem 1987: 4) are described as not exhibiting palato-alveolar affricates (/tʃ/, /dʒ/), but lamino-palatal and lamino-postalveolar affricates respectively, in both cases transcribed /tɕ/ and /dʑ/. In Sunwar (Borchers 2008: 35), affricates are described as palatal (/c/, /ɟ/). For Newah, Genetti has different

descriptions of affricates that she transcribes <c> and <j>: in her grammar (2007: 38) she uses the term palato-alveolar and in the grammar sketch published in the Sino-Tibetan Languages (2003: 356) she describes them as alveolo-palatal. Hargreaves (2003: 372) uses the term "alveo-palatal" and uses as well <c> and <j> to transcribe them.

Some languages of the region exhibit complementary distribution in the realization of affricates, such as in Chantyal (Bodish) where alveolo-palatal affricates ([tɕ], [dʑ]) are attested in front of front vowels and alveolar affricates ([ts], [dz]) elsewhere (Noonan 2003: 317).

The variation attested in free variation for the Chepang affricates is also observed in Nepali, and may have developed under language contact. Masica (1991:94) points out that there is a tendency in IA languages to pronounce alveolo-palatal affricates as alveolar affricates.

2.3.1.6.   Nasal onset

Two series of nasal consonants (non-breathy and breathy) contrast at three places of articulation: bilabial, alveolar, and velar. A minimal set of non-breathy nasals which contrast in place of articulation is presented in Table 18. A complete minimal set distinguishing non-breathy vs. breathy nasals has not been found in our data, only examples of near minimal pairs. These are presented in Table 19.

Table 18. Minimal set for non-breathy nasal onsets

m	*mo*	'Indian spotted eagle'	n	*no*	'ear'	ŋ	*ŋo-*	'feel dizzy, drunk'

Table 19. Minimal and almost minimal sets for non-breathy versus breathy nasal onsets

m	*maŋ-*	'swell'	mʰ	*mʰaŋ-*	'wake sb up'
n	*na-*	'COP, to be born'	nʰ	*nʰap-*	'lay sth down (mat, clothes)'
ŋ	*ŋa*	'1SG'	ŋʰ	*ŋʰaɾ-*	'feel full, not able to eat more'

Breathy nasals are mainly produced as a breathy consonant, i.e. [m̥], [n̥], [ŋ̊], rather than a breathy vowel exclusively, as illustrated in Figure 11 and Figure 12 with the words *mʌj-saj* 'banana' and *mʰʌj=sa* 'be lost, disappear.' If one compares the nasal segments in the two figures, the breathy nasal appears to be less voiced than the non-breathy nasal segment. Breathiness on nasals clearly entails devoicing of the nasal segment. This is also observed with other breathy sonorants (§ 2.3.1.7, § 2.3.1.8). In Bandhu, Dahal & Caughley (1970), breathy sonorants were essentially described as voiceless sonorants.

While breathy nasals can be merely described as breathy consonants, they nevertheless have their breathiness spread to the first half of the nucleus (Figure 11). Breathiness spreading through the entire nucleus is also attested with some speakers (Figure 12).Caughley (1969, 1970a, 1982, 2015) considers breathy nasals as the combination of two segments, a glottal fricative /h/ followed by a nasal. He applies the same analysis to other breathy or aspirated consonants. In Caughley (1969, 1970a, 1982, 2000, 2015, 2016), breathy nasals and other breathy sonorants at onset are thus transcribed as: /hm/, /hn/, /hŋ/, /hl/, hr/, /hy/, and /hw/.

Figure 11. Breathy bilabial nasal /mʰ/ vs. /m/ - breathiness spreading on half nucleus

mʌj-saj 'banana' vs. *mʰʌj=sa* 'be lost, disappear'
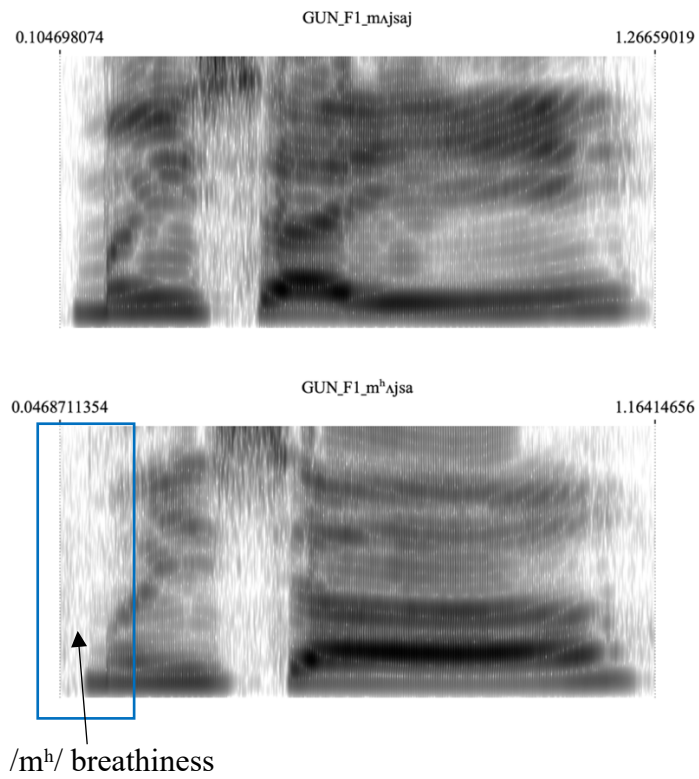
/mʰ/ breathiness

Figure 12. Breathy bilabial nasal /mʰ/ vs. /m/ - breathiness spreading on full nucleus

mʌj-saj 'banana' vs. *mʰʌj=sa* 'be lost, disappear'

/mʰ/ breathiness
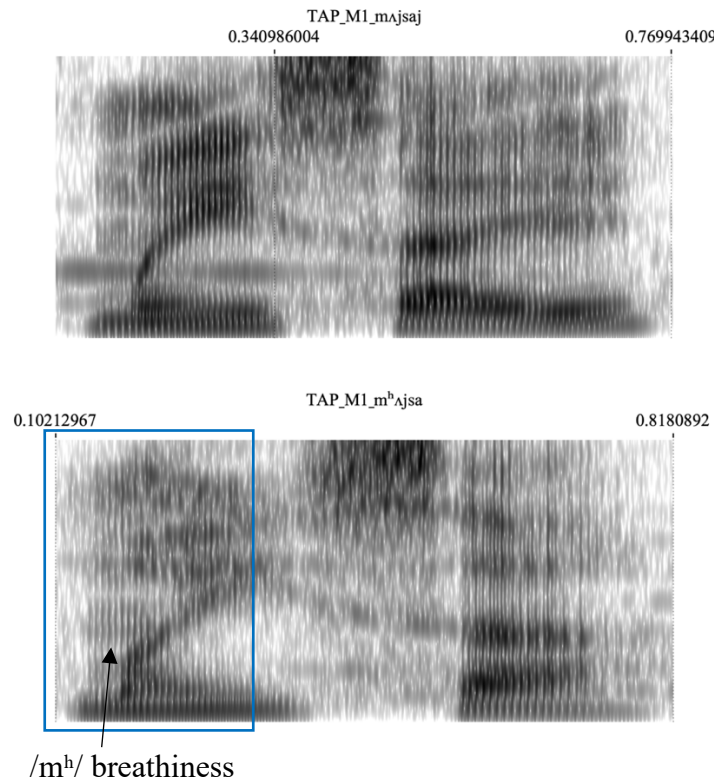
Variation in the realization of breathy nasals is only attested so far in two villages of RAP-13 (TAP, GUN) and in RAK-8 (MAI). These allophones are not predictable, i.e. they are not triggered by any specific phonetic environment. Moreover, they are idiosyncratic in that they do not target all the roots featuring the phoneme triggering the allophony. Finally, it appears that in some cases, more than one allophone of the same phoneme can be used interchangeably in the same geographical area.

The breathy bilabial nasal phoneme /mʰ/ shows two additional allophones, each of which the result of two opposite changes: a glottal fricative [h] from lenition or loss of the nasal segment, and an unbreathy phoneme [m] result of lenition or loss of breathiness. The allophones of the breathy bilabial nasal /mʰ/ are only attested in RAP-13 (TAP, GUN). All the roots formed with the onset /mʰ/, along with their allophones, have been reported in Table 20. While the allophone [h] is found in free variation with [mʰ] in these villages, the allophone [m] is the only one used in TAP and GUN. Spectrograms of the root /mʰaŋ-/

'wake sb up' pronounced as [mʰ] and [h] are presented in Figure 13 and Figure 14 respectively.

Table 20. Idiosyncratic allophones of the breathy bilabial nasal /mʰ/

mʰ	*mʰa-*	'become extinct, disappear'
	mʰok-	'empty a basket carried on the back'
	mʰʌjk-	'have lucid dream'
	mʰʌk-	'be soaking, wet'
mʰ [mʰ]~[h]	*mʰaŋ- ~ haŋ-*	'wake sb up'
	mʰeʔ ~ heʔ	'fire'
	mʰeʔ- ~ heʔ-	'forget'
	mʰotoŋ ~ hotoŋ	'mouth'
	mʰus- ~ hus-	'be burnt (food, rice)'
	mʰʌj- ~ hʌj-	'be lost, disappear'
	mʰʌɾ- ~ hʌɾ-	'wonder'
	mʰjaʔŋ- ~ hjaʔŋ-	'be eager, motivated to'
mʰ [mʰ]~[m]	*mʰʌs ~ mʌs*	'smell of sweat'
	mʰut- ~ mut-	'kindle, ignite (fire)'

Figure 13. Phoneme /mʰ/ allophone [mʰ]

mʰaŋ=sa [mʰaŋ.sa] 'wake sb up'

TAP_M1_mʰaŋsa
0.0578445829   0.94406778
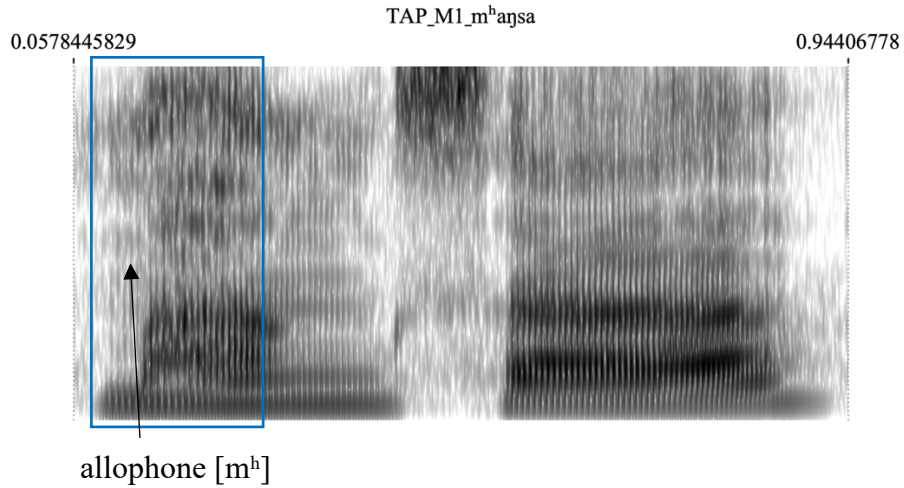
allophone [mʰ]

Figure 14. Phoneme /mʰ/ allophone [h]

mʰaŋ=sa [haŋ.sa] 'wake sb up'

TAP_M1_haŋsa
0.0442055356   0.616807851
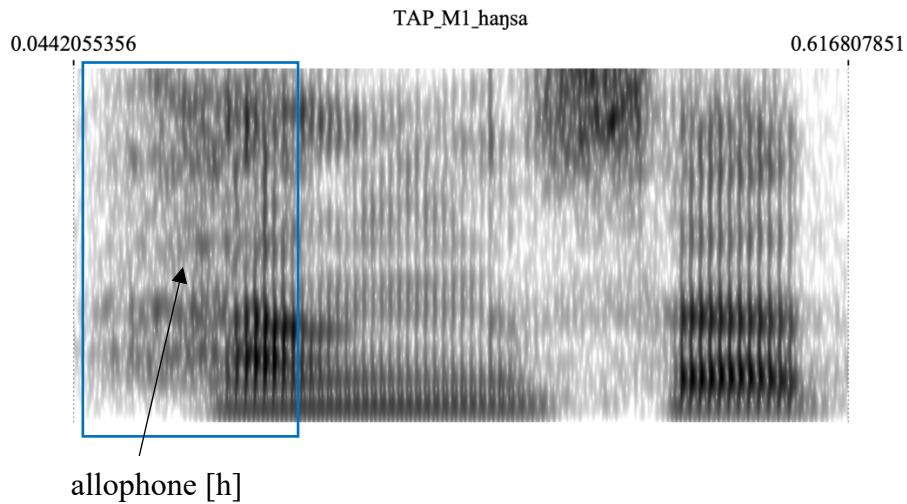
allophone [h]

In our data, only three roots feature the breathy alveolar nasal /nʰ/ at onset and one of them shows a different allophone [h]. These roots are presented in Table 21.

Similarly to /mʰ/, the allophone [h] is found as an allophone of [nʰ]. In this case as well, this is the result of lenition or loss of the nasal segment. This variation in

pronunciation, attested in RAP-13 (TAP, GUN), is only observed with one of the two roots formed with the palatal on-glide of the nucleus /ja/.

Table 21. Idiosyncratic allophones of the breathy alveolar nasal /nʰ/

nʰ	nʰap-	'lay down, lay out sth (mat, sheet, clothes)'
	nʰjaʔj-	'lean over, lean down'
nʰ [nʰ]~[h]	nʰjahm- ~ hjahm-	'be free with no line (place to do work like mill, shop)'

In the data collected from RAK-8 by Caughley (2000, 2016), other roots featuring a breathy alveolar nasal /nʰ/ at onset are reported. Either these roots have been replaced by borrowings from Nepali or their meanings have not been recognized by the speakers of the Lothar varieties.

The phoneme /ŋʰ/ shows a glottal fricative allophone [h] similarly to /mʰ/ and /nʰ/, as a result of lenition or loss of nasality. The root ŋʰjal ~ hjal 'pollen' likely lost its on-glide /j/ of the glide vowel sequence /ja/ to become ŋʰal in RAK-6.

The allophones of the breathy velar nasal /ŋʰ/ presented in Table 22 are again only attested in RAP-13 (TAP, GUN). Elsewhere, /ŋʰ/ is used.

Table 22. Idiosyncratic allophones of the breathy velar nasal /ŋʰ/

ŋʰ	ŋʰur-	'snore'
	ŋʰwehj-	'swing hips (while dancing)'
ŋʰ [ŋʰ]~[h]	ŋʰar- ~ har-	'feel full, not able to eat more'
	ŋʰjal ~ hjal	'pollen'

An additional root with a breathy velar nasal onset is attested in RAK-6: ŋʰik- 'be sweet (cashew, peach seed).' Phonemic and semantic variations attested for this root are shown in Table 23. As we can see, the root ŋʰik- has a different form hik- in RAP-13, result of lenition or loss of nasality, but also a different yet related meaning: 'be very sweet.' While this latter meaning is expressed by the root sjok- in RAK-6, the root ek- is used in RAP-13 to mean 'be fat and sweet (peanut, pork, capricorn beetle larvae, cashew,

peach seed).' This last meaning encompasses that of the RAK-6 root *ŋʰik-* 'be fat and sweet (cashew, peach seed),' while RAP-13 restricts the meaning of *ek-* to 'be fat and sweet (peanut)' and uses the root *sjok-* to mean 'be fat and sweet (pork, capricorn beetle larvae).'

Table 23. Phonemic and semantic variation for the RAK-6 root *ŋʰik-*

	LOTHAR		
	RAP-13		RAK-6
ek-	'be fat and sweet (peanut, pork, capricorn beetle larvae, cashew, peach seed)'	*ŋʰik-*	'be fat and sweet (cashew, peach seed)'
hik-	'be very sweet'	*sjok- (mah)*	'be very sweet'
ek-	'be fat and sweet (peanut, pork, capricorn beetle larvae, cashew, peach seed)'	*sjok-*	'be sweet; be fat and sweet (pork, capricorn beetle larvae)'
		ek-	'be fat and sweet (peanut)'

Additional roots featuring a breathy nasal onset /ŋʰ/ are reported for RAK-8 (MAI). These roots collected by Caughley (2000, 2016) show different onsets in Lothar varieties. Amongst these different onsets, a change from pronouncing /ŋʰ/ as [ŋʰ] to the fricative glottal [h] is also observed on three roots. By contrast with the breathy nasal [ŋʰ] to fricative glottal [h] type of change described above, the particular allophone [h] of the phoneme /ŋʰ/ in these three roots is present in all studied Lothar varieties, i.e., not merely restricted to RAP-13 (TAP, GUN). This shows that, assuming the same directionality of change, the development from a breathy nasal to a fricative glottal [h] has potentially taken place in a different way and at a different time.

In addition to the glottal fricative allophone [h], a sound correspondence between /ŋʰ/ and /kʰ/ is attested with one root: *hŋayk-* ~ *khayk-* 'wrench off, harvest corn.' It is here unclear whether the directionality is from [ŋʰ] to [kʰ], since both roots are used in RAK-8 (MAI). It is possible that, in this particular case, /kʰ/ became [ŋʰ] as a result of anticipatory assimilation of voicing given the presence of the palatal off-glide /j/ in the nucleus. Indeed, no such correspondence is attested with other roots formed with an

aspirated velar stop /kʰ/ at onset and a non-aspirated velar stop /k/ at coda, when it comes to roots featuring a monophthong nucleus.

Finally, the root *hŋənə-* 'be wobbly, unsteady (post, tree)' is not attested in all the studied Lothar varieties, but only the root *rʰunʌ-* 'shake, move (tree, stone).' This meaning is similar to that of *hŋənə-*. The morphology of native dissyllabic roots formed with an open-mid back vowel /ʌ/ at the nucleus of the second syllable, is rare, and *rʰunʌ-* is the only root attested with an alveolar nasal at onset of the second syllable. It is indeed possible that *hrunuʔ-* and *hŋənə-* are different transcriptions of the same root, since different transcriptions of the same morpheme (with sometimes a slightly different meaning) are observed elsewhere in Caughley (2000, 2016).

In Table 24, I present all the roots with a breathy nasal onset /ŋʰ/ found in RAK-8 (MAI) in comparison with the forms attested in the Lothar varieties.

Table 24. Additional variations attested for the RAK-8 onset /ŋʰ/

MANAHARI		LOTHAR	
RAK-8			
hrunuʔ-	'shake object, cause to vibrate'	*rʰunʌ-*	'shake, move (tree, stone)'
hŋənə-	'be wobbly, unsteady (post, tree)'	-	
hŋayk- ~ khayk-	'wrench off, harvest corn'	*kʰayk-*	'twist and break off (corn)'
hŋik-	'creak, grate, squeak; grind teeth'	*hikʌ-*	'grind teeth, saw'
hŋi- ~ hni-diŋ	'daytime'	*hi-diʔŋ*	'afternoon (1-3pm)'
hŋis-	turn (head) away (disdain)'	*hil-*	'turn away from, stretch back'

The breathy alveolar and velar nasals /nʰ/ and /ŋʰ/ appear to be rare in Chepang, by contrast with the breathy bilabial nasal /mʰ/. Cognate roots with /mʰ/ onset are found in Bhujel[30], but roots featuring /nʰ/ or /ŋʰ/ onsets in Bhujel have different forms in Chepang. These forms possibly suggest later innovations in Bhujel, such as the Chepang root *neh* 'nose' having developed to *nʰe* 'nose' in Bhujel. The only cognate in Bhujel that shows a similar breathy nasal onset in Chepang is the root *nʰi* 'day' which occurs in the compound *hi-diʔŋ* (Lothar) ~ *hŋi- ~ hni-diŋ* 'afternoon' in the variety of RAK-8 (MAI).

---

[30] Data from Bhujel, as found in Caughley (1998) and Regmi (2007).

What is interesting in the variation attested for breathy nasals in Chepang, and in particular the change that consists in the lenition or loss of the nasal segment resulting in a glottal fricative [h], is that this change does not apply to all roots featuring the phoneme which entails the variant pronunciation. This either suggests the existence of different phoneme sources for the allophones, or that this sound change has taken place at different stages, being triggered by lexical frequency. Yet, comparing Chepang and Bhujel cognate roots with those found in other Himalayan languages would highlight the historical status of breathy nasals in Chepang and Bhujel, and help reconstruct their proto-forms at the Proto-Chepang-Bhujel (PCB) level and beyond.

2.3.1.7.   Liquid onset

Similarly to nasals, the two series of liquids /l/ and /ɾ/ contrast in phonation, i.e. non-breathy vs. breathy. The four liquids /l/, /ɾ/, /lʰ/ and /ɾʰ/ are alveolar. The phonemes /l/ and /lʰ/ are lateral approximants and /ɾ/ is a flap. Minimal sets contrasting liquids /l/, /ɾ/, /lʰ/ and /ɾʰ/ are presented in Table 25. Both breathy /lʰ/ and /ɾʰ/ are pronounced like breathy consonants ([l̤], [ɾ̤]) with breathiness more or less spreading to the first half of the nucleus as observed with nasals.

Table 25. Minimal sets for liquid onsets

l	*la*	'rope'	ɾ	*ɾa*	'bamboo winnowing-tray'
l	*lok-*	'be far'	ɾ	*ɾjaŋ-sa*	'smoke and dry (meat)'
lʰ	*lʰok-*	'send, lend, send away'	ɾʰ	*ɾʰjaŋ-sa*	'have large eye and eyelid'
lʰ	*lʰjut-*	'peel scratching'	ɾʰ	*ɾʰjut-*	'drag, pull along'

2.3.1.8.   Approximant or glide onset

The palatal approximant or glide /j/ and the bilabial approximant or glide /w/ are distinguished by their phonation, i.e., non-breathy vs. breathy. Table 26 illustrates their contrast in place of articulation, and Table 27 in phonation.

Table 26. Minimal set for non-breathy approximant or glide onsets

j	*ja-*	'one'	**w**	*wa*	'bird, hen'

Table 27. Minimal sets for breathy approximant or glide onsets

j	*jom*	'bear'	**jʰ**	*jʰom*	'flea (of hen)'
w	*wa*	'bird, hen'	**wʰ**	*wʰat-*	'pierce (ear, nose)'

2.3.2. Consonant codas

The set of native consonant codas comprises 13 consonants: 4 stops, 2 fricatives, 2 liquids, 3 nasals, and 2 glides. Only voiceless stops, voiceless fricative and non-breathy sonorants occur at syllable coda. They are presented in Table 28.

Table 28. Native phonemic consonant codas and main allophones

	bil.	alv.	pal-alv.	pal.	vel.	gl.
stop	p	t			k	ʔ
allophone	[pˈ] [pʰ]	[tˈ] [tʰ]			[kˈ] [kʰ]	
fricative		s				h
nasal	m		n			ŋ
flap				ɾ		
lat. approximant		l				
glide		w		j		

121

2.3.2.1. Stop coda

Only four stops can be found in coda position; they are voiceless and contrast in place of articulation: /p/, /t/, /k/ and /ʔ/. They are illustrated in Table 29.

Table 29. Minimal set for stop codas

CV	*ju-*	'melt, dissolve'			
p	*jup-*	'catch sb with the arm'	**t**	*jut-*	'caress'
k	*juk*	'rhesus macaque'	**ʔ**	*juʔ*	'mouse'

The voiceless stops /p/, /t/ and /k/ at syllable coda have two additional allophones attested in free variation. These allophones either show absence of release or aspiration. The unreleased allophones [p�], [tˈ], and [kˈ] are illustrated with a sound wave and spectrogram of /t/ at coda in Figure 15. In this context, the root *nemet* 'ant' is uttered in the following sentence: *ja=dʑo nemet dal=kʰa mu=na* 'An ant is in the lentils.'

When a root is pronounced in isolation, voiceless stops can surface as well with the presence or absence of the release, but can also be aspirated. This is illustrated in Figure 16 with the root *tuk* 'stomach,' pronounced in isolation by the same speaker.

The aspirated allophone of a voiceless stop at coda is not restricted to roots uttered in isolation. The aspirated realization of final stops also occurs in identificational constructions in absence of copula. This allophone may be triggered by the absence of a morpheme following a root which holds the final position in a sentence in addition to a fall in F0. The condensed airflow during the closure is more strongly liberated, resulting in a puff of air at the release. This is shown in Figure 17 the phoneme /p/ realized [pʰ] in the root *pop ~ pʰop* 'lung' uttered in the following sentence: *didi, i pʰop* 'Elder sister, this is a *lung*.'

Figure 15. Phoneme /t/ allophone [tˀ]

nemet [nemetˀ] 'ant'
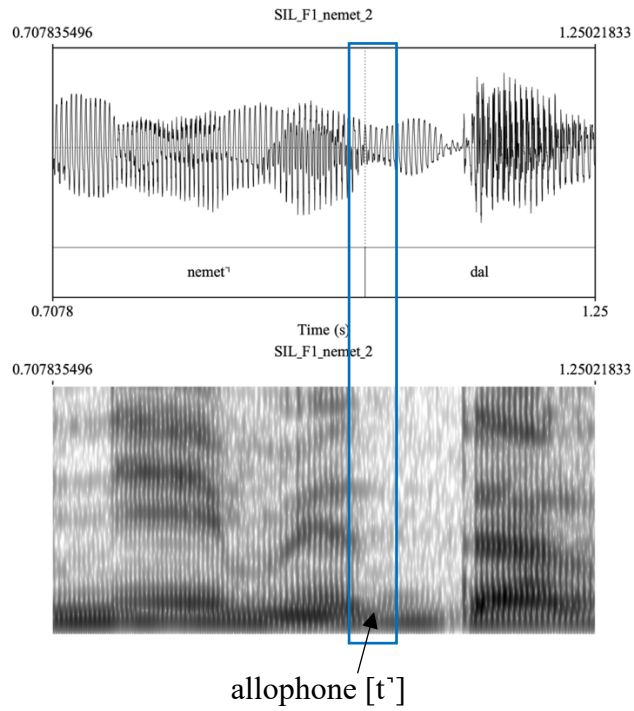
allophone [tˀ]

Figure 16. Phoneme /k/ allophone [k], [kˀ], and [kʰ]

tuk [tuk], [tukˀ], [tukʰ] 'stomach'

allophone [kˀ]   allophone [k]   allophone [kʰ]
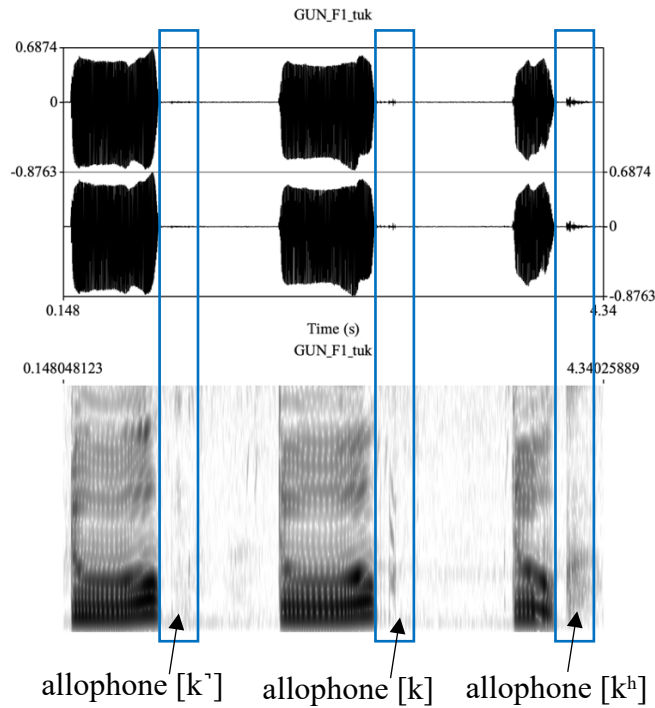
Figure 17. Phoneme /p/ allophone [pʰ]

pop ~ pʰop [pʰopʰ] 'lung'

allophone [pʰ]
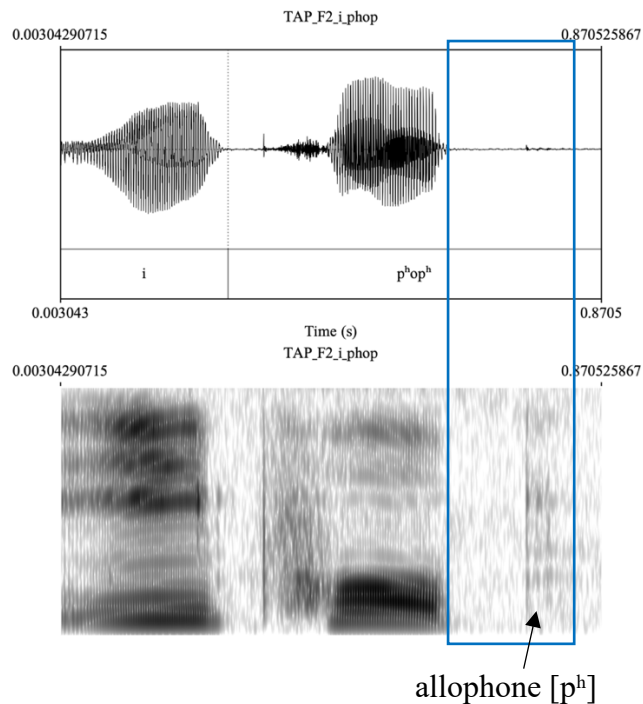
2.3.2.2. Fricative coda

Both fricative /s/ and /h/ contrast in syllable final position. A minimal pair illustrates their contrast in Table 30.

Table 30. Minimal set for fricative codas

CV	*mu-*	'COP, be'			
s	*mus*	'cloud'	**h**	*muh*	'reproductive growth on yam vine, bulbil'

2.3.2.3. Nasal coda

Nasal consonants /m/, /n/, and /ŋ/ contrast at coda, as illustrated in Table 31.

Table 31. Minimal set for nasal codas

CV	*du-*	'be red'						
m	*dum*	'calabash'	**n**	*dun-*	'be thick, dense'	**ŋ**	*duŋ-*	'sprout'

2.3.2.4.   Liquid coda

Liquids /l/ and /ɾ/ are contrastive in final position, as shown in Table 32.

Table 32. Minimal set for liquid codas

CV	ga-	'open mouth'			
l	*gal-*	'be black'	**ɾ**	*=gaɾ*	'a little' (verbal clitic)

2.3.2.5.   Approximant or glide coda

The approximants or glides /j/ and /w/ are contrastive in final position, as shown in Table 33.

Table 33. Minimal set for glide codas

CV	ŋa	'1SG'			
j	*ŋaj-*	'find, meet'	**w**	*ŋaw-*	'roast'

2.3.2.6.   Glottal coda

The glottal fricative /h/ and glottal stop /ʔ/ share the same position at syllable coda, as illustrated in Table 34.

Table 34. Minimal set for glottal codas

CV	la	'rope'			
h	*lah*	'moon'	**ʔ**	*laʔ*	'arrow'

2.3.3. Glottal segments in complex codas

In addition to contrast at coda, the glottal segments /h/ and /ʔ/ share as well a pre-final sonorant position, such that they can form part of a complex syllable coda. Minimal sets of /h/ and /ʔ/ in pre-final sonorant position are presented in Table 35.

The pre-sonorant position of the glottal segments is restricted to syllable coda, i.e., before sonorants in final position. The fact that such combination does not exist in initial position is not the sole reason for considering it as the result of two independent phonemic segments combined rather than a single phonemic segment. Indeed, the independence of these segments primarily lies in the fact that metathesis occurs at morpheme boundary when the glottal fricative /h/ is in pre-final sonorant position of a root encliticized with a morpheme whose onset is a vowel. The glottal fricative /h/ shifts to set at the onset of the encliticized vowel initial morpheme (§ 2.3.6.2.1).

The presence of glottal segments preceding sonorants at syllable coda are ancient traces of derivational verbal morphology (§ 5.2).

The realization of the glottal stop in final and pre-final sonorant positions presents variation. These realizations are described in § 2.8.

Table 35. Minimal sets for pre-glottalized nasal sonorant coda

m	*kum-*	'tie weaving in circle'			
hm	*kuhm-*	'lean head on sth'	**ʔm**	*kuʔm-*	'make sth to lean head on'
n	*dan-*	'ferment greens'			
hn	*dahn-*	'press on'	**ʔn**	*daʔn-*	'repeat'
ŋ	*tuŋ-*	'drink'			
hŋ	*tuhŋ-*	'be drained out'	**ʔŋ**	*tuʔŋ*	'foot (of tree)'
l	*dol*	'crest of hill'			
hl	*dohl-*	'light up (fire)'	**ʔl**	*doʔl*	'capricorn larvae'
ɾ	*juɾ-*	'squeeze furuncle'			
hɾ	*juhɾ-*	'sink into, be shoved under'	**ʔɾ**	*juʔɾ-*	'extract juice'
j	*aj*	'mother-in-law'	**j**	*saj*	'fruit, seed'
hj	*ahj-*	'be soft'	**ʔj**	*saʔj-*	'listen, hear'

| **w** | *ʌw-* | 'cease (rain)' | **w** | *tʌw-* | 'Indian gooseberry' |
| **hw** | *ʌhw-* | 'threaten sb physically' | **ʔw** | *tʌʔw-* | 'be bent back and up' |

2.3.4. Nativization of non-native consonants

Nativizing Nepali borrowings targets morphemes that feature phonemes that either are non-native to Chepang or phonemes that are native to Chepang but whose position in the syllable does not correspond to the syllable structure of Chepang.

Retroflex stops /ʈ/, /ʈʰ/, /ɖ/ and /ɖʰ/ and post-alveolar fricative /ʃ/ are not native to Chepang. Geminate consonants are also not present in Chepang natively. These are found in words borrowed from Nepali (IA), the main source of borrowings for Chepang. Another source of borrowings for Chepang is English. When English vocabulary enters the Chepang lexicon, it happens via Nepali. When Nepali borrows words from English, alveolar stops /t/ and /d/ in syllable initial position are pronounced like unaspirated or unbreathy retroflexes, such as with the words /ʈim/ 'team' or /ɖens/ 'dance.' Chepang will preserve the presence of retroflexes in such words borrowed from English via Nepali.

The treatment of Nepali borrowed roots is subject to variation in two direction: adaptation of the borrowed root to the phonology of Chepang, i.e., nativization, or preservation of the original phonological realization of the root, i.e. pronounced similarly to Nepali. It is possible that nativized borrowings are older than those preserving the Nepali pronunciation. This is also the intuition of the native speakers with whom this question was discussed. Chepang populations did not have access to an education in Nepali before 1960 but have been in greater contact with Nepali speaking communities and others (like the Tamangs in the hills or Tharus in the plain) at least since the mid-19th century. It is possible that older borrowings have been nativized by contrast with recent ones, since it is observed that a single individual can treat differently the same phoneme in different roots.

When the retroflex stops /ʈ/, /ʈʰ/, /ɖ/ and /ɖʰ/, post-alveolar fricative /ʃ/, geminate consonants and other phonemes (which are native but do not fit Chepang syllable structure) are nativized, the following changes are observed: retroflex stops become alveolar in initial and final position; voiced retroflexes and stops are devoiced in final

position; aspiration or breathiness of retroflexes or stops is lost in final position; the post-alveolar fricative is pronounced like an alveolar fricative; voiced and voiceless affricates are pronounced like an alveolar voiceless fricative; geminates undergo degemination.

These nativizing processes observed with borrowed roots from Nepali are summarized in Table 36.

Table 36. Nativization processes with Nepali borrowings

Nepali borrowing	Chepang nativization	
non-native phoneme	**syllable initial**	**syllable final**
ʈ	t	t
ɖ	d	t
ʈʰ	tʰ	t
ɖʰ	dʰ	t
ʃ	s	s
non-native syllabic structure	**syllable initial**	**syllable final**
voiced stop	voiced stop	devoicing
aspirated or breathy stop	aspirated or breathy stop	loss of aspiration or breathiness
voiced affricate	voiced affricate	s
voiceless affricate	voiceless affricate	s
geminate	single consonant	degemination

Nativized pronunciation can be seen as a sign of earlier borrowings while non-nativized borrowings as more recent, as a result of a greater bilingualism. This observation nevertheless needs to be investigated through a socio-linguistic study examining the impact of bilingualism on borrowings or that of the speakers' attitude towards Nepali with regard to their native language. Indeed, some speakers will not treat borrowings differently but preserve either the original Nepali pronunciation or conversely nativize all borrowings. This may be triggered by the degree of education or literacy of the speakers, or by their attitude towards the borrowed language. Elder speakers who have not been exposed to learning how to read and write Nepali in their childhood, or who have started speaking Nepali later in life are likely to treat borrowings differently from younger bilingual speakers. In addition, the question of whether nativization of

Nepali borrowings is an effect that is deliberately avoided or sought by speakers is another open question.

Finally, some roots, which have also likely been borrowed from Nepali, are treated yet differently than other roots. For instance, the Nepali root *koḍo* 'millet (finger),' which does not show a great difference with the modern phonology or syllable structure of Chepang, besides the presence of the voiced retroflex stop /ḍ/, is pronounced /kʌdʌw/ and not [kodo] as expected given the nativization processes observed. This difference can find an historical sociolinguistic explanation in the fact that the Nepali root *koḍo* can have been one of the roots borrowed from Nepali at earliest stages of language contact. This could possibly go back to a time when the Chepangs started to rely on this type of cultivated millet for food, and that it was likely acquired from Nepali speaking communities or others practicing an irrigated agriculture since it could not – and still cannot in many places – be cultivated through a traditional slash and burn type of agriculture practiced in the hills. It is indeed possible that the phonological or phonotactic system of Chepang was different at the time of this borrowing, or that the pronunciation of this Nepali root itself was different, or that the root was borrowed from another non-native Nepali speaking community (like the Tharus in the plain), who had nativized it in their own way. Another example of unexpected treatment of Nepali borrowings is the Nepali root *sag* 'greens' nativized /sagʌ/ while the pronunciation /sak/ would be expected here.

While there is variation in the treatment of borrowings, which can merely be seen as individual until we are able to observe significant patterns through an in-depth (historical) socio-linguistic study, there are nevertheless certain roots that are observed to be consistently nativized through the processes described in Table 36, even amongst speakers treating Nepali borrowed roots differently. These roots correspond to those that we suggest are likely to have been borrowed at earlier stages. Some examples of these are reported in Table 37.

Table 37. Nativized borrowed roots

Chepang	meaning	from Nepali
dut	'milk'	*dut- ~ dutʰ*
tʌpʌɾi	'plate made from leaf'	*tʌpʌɾi*
got	'cowshed'	*gotʰ*
bap	'vapor'	*bapʰ*
sʌjok	'help'	*sʌjog*
pãs	'five'	*pãtɕ*
pjas	'onion'	*pjadz*
asa	'hope'	*aʃa*
biswas	'trust'	*biʃwaʃ*
pʌtɾʌ	'letter, layer'	*pʌttɾʌ*
tɕitʌl	'chital'	*tɕittʌl*

Other roots show more variation; they are observed with borrowings from Nepali and from English via Nepali.

Roots featuring the bilabial voiced fricative /β/ can be preserved or nativized to /b/. Epenthesis of a vowel can occur when the coda and onset of two contiguous syllables do not correspond to Chepang phonotactics. Some roots are presented in Table 38.

Table 38. Other nativized borrowed roots

Chepang	meaning	from Nepali	from English via Nepali
silaβʌɾ	'fake silver'		*silaβʌɾ*
pʰɾi ~ pʰiɾi	'free'		*pʰɾi ~ pʰiɾi*
pʰoto	'photo'		*pʰoto*
puɾbʌ	'East'	*puɾβʌ ~ puɾbʌ*	
ɾatnitik	'politics'	*ɾadznitik*	
mʌɾit	'black pepper'	*mʌɾidz*	

2.3.5. Consonant cluster onsets

Chepang has native consonant clusters in syllable-initial position formed with two liquid medial consonants, the lateral approximant /l/ and flap /ɾ/. These consonant clusters are presented in Table 39. /ɾ/ is attested in all clusters, following two series of stops (unaspirated/unbreathy voiceless and voiced) at three places of articulation (bilabial, alveolar and velar), the fricative /s/, and one series of nasals (unbreathy) at two places of articulation (bilabial and velar). However, /l/ is only observed with the two series of stops at two places of articulation (bilabial and velar) and does not combine with the fricative and nasal series.

In addition to the consonant cluster onsets formed with /l/ and /ɾ/, a few roots present breathy initial clusters. These breathy clusters only occur with /ɾ/.

Table 39. Consonant cluster onsets formed with liquids

	bilabial		alveolar	velar	
stop	pl	pɾ	tɾ ~ kɾ	kl	kɾ
	bl	bɾ	dɾ ~ gɾ	gl	gɾ
fricative			sɾ		
nasal		mɾ			ŋɾ ~ gɾʰ

As described in § 2.2, in addition to /l/ and /ɾ/, the position of medial consonant at syllable onset is also shared with the bilabial approximant or glide /w/. The glide /w/ is attested following a more restricted set of consonants than /l/ and /ɾ/; it is only found in combination with the velar stops /k/ and /g/. These specific clusters are analyzed as possible reflexes of PTH labiovelar consonants /*kʷ/ and /*gʷ/ or glide clusters /*kw/ and /*gw/. The position of medial consonant at syllable onset held by /w/ is not shared with the palatal approximant or glide /j/. Indeed, the position of /j/ belongs to the nucleus or the rhyme, since it can follow as well consonant clusters formed with /l/ and /ɾ/. Further, /j/ can follow any consonant, by contrast with /l/, /ɾ/ and /w/ which are restricted to a specific set.

Bilabial consonant cluster onsets are described in § 2.3.5.1, alveolar consonant cluster onsets in § 2.3.5.2, and velar consonant cluster onsets in § 2.3.5.3. Breathy cluster onsets are presented in § 2.3.5.4. Velar consonant clusters formed with a bilabial approximant are described in § 2.3.5.5. Consonant clusters followed by a sequence formed by a glide and a vowel or two vowels are presented in § 2.3.5.6. Finally, a brief account of the consonant clusters reconstructed back to PTH will be presented in § 2.3.5.7.

2.3.5.1.   Bilabial consonant cluster onsets

There are five bilabial consonant cluster onsets including stop and nasal manners of articulation: /pɾ/, /bɾ/, /pl/, /bl/, /mɾ/. They are illustrated with a set of examples in Table 40. As mentioned, nasals do not combine with the lateral approximant /l/, leaving /mɾ/ the only possible bilabial nasal cluster onset.

Table 40. Bilabial consonant cluster onsets

pl	*plek*	'thigh'	pɾ	*pɾek-*	'cut (meat)'
bl	*blʌk-*	'sink, set (sun, moon)'	bɾ	*bɾʌk-*	'together'
			mɾ	*mɾahŋ*	'echo'

Bilabial clusters are stable across the studied varieties, in the sense that they do not present much variation. The only marginal variation observed with the bilabial onset is a correspondence /pɾ/ ~ /kɾ/ for the root *pɾʌʔ-* ~ *kɾʌʔ-* 'cut all around the base of a tree (this process will make the tree die).' The root *pɾʌʔ-* is so far only attested in POL (RAP-13). This change may be seen as idiosyncratic since it is not observed with other words whose onset cluster is formed with a velar or a bilabial initial consonant.

2.3.5.2.   Alveolar consonant cluster onsets

There are three alveolar cluster onsets formed with stop and fricative manners of articulation: /tɾ/, /dɾ/, /sɾ/. They are illustrated in Table 41.

Table 41. Alveolar consonant cluster onsets

tɾ ~ kɾ	*tɾom- ~ kɾom-*	'patter, fall in small drops'
dɾ ~ gɾ	*dɾop ~ gɾop-*	'be sticky, tacky'
sɾ	*sɾap-*	'whip, slash, slap (rope, clothe)'
sɾ ~ ɾʰ	*sɾek- ~ ɾʰek-*	'smell good (of cooked food: nettle, orchid flowers, bat meat)'

As initial segments of these clusters, voiceless and voiced alveolar stops are only attested so far in two villages of RAP-13 (POL, SYAM). In the other studied varieties, they are realized with an initial velar stop. I present all the roots attested with the alveolar stop clusters in Table 42. It is possible that the alveolar ~ velar onset cluster correspondence is the result of a merger of the alveolar cluster onset with the velar cluster onset. This allophony does not seem to be conditioned by any particular phonetic environment and could be considered an innovation. They are found in combination with any type of vowels. In addition, the stop and nasal consonants in final position of these words do not show any feature that could condition this change. The analysis of the directionality of this merger from alveolar to velar is also supported by the fact that velar clusters also exist in the RAP-13 varieties of POL and SYAM. Finally, the alveolar clusters /*tr/ and /*dr/ have been reconstructed back to PTH (§ 2.3.5.7).

Table 42. Words attested with alveolar consonant cluster onsets

tɾ ~ kɾ	*tɾek- ~ kɾek-*	'tear, rip off'
	tɾeŋ- ~ kɾeŋ-	'kick, push with foot'
	tɾuk- ~ kɾuk-	'wring, milk'
	tɾok- ~ kɾok-	'spray on (urine)'
	tɾom- ~ kɾom-	'patter, fall in small drops'
	tɾʌt- ~ kɾʌt-	'be branched (tuber)'
	tɾak ~ kɾak	'testicle'
	tɾaw- ~ kɾaw-	'be sour (homemade beer kept more than one day)'
dɾ ~ gɾ	*dɾim- ~ gɾim-*	'beat in panic when startled, walking alone (heart)'
	dɾehm- ~ gɾehm-	'be articulated, clear (speech)'
	dɾop ~ gɾop-	'be sticky, tacky'

The cluster /sɾ/ exists in most of the studied varieties, but some roots starting with /sɾ/ present an allophone /ɾʰ/ with others. In the varieties of MAN-4 the phoneme /ɾʰ/ is essentially used. Roots starting with the cluster /sɾ/ are more predominant in the varieties spoken in RAP-13 and RAK-6. The cluster /*sɾ/ is reconstructed back to PTH (Benedict 1972: 37; Matisoff 2003: 59; Jacques 2015), as in the word for 'louse': /*s-rik/, /*srik/ or /*śrik/ (Benedict 1972: 13–14; Matisoff 2003: 78). Jacques (2015: 216) shows additional reflexes of /*sr/ in Japhug, such as /zrɯɣ/ 'louse.' Chepang has also preserved the consonant cluster /*sr/ as with the root /sɾʌjk/ 'louse.'

The four sound correspondences involving /sɾ/ or /ɾʰ/ are presented in Table 43.

It seems clear that we can reconstruct the cluster /*sɾ/ in correspondence 1, and the phoneme /*ɾʰ/ in correspondence 4. It is likely that correspondence 1 is the result of a merger of /sɾ/ with /ɾʰ/ in MAN-4, where this change shows no exception, since the cluster /sɾ/ has entirely disappeared: /sɾʌjk/ 'louse' became /ɾʰek/ (with an additional change in rhyme). If /*sɾ/ is reconstructed as well for both correspondence 2 and 3, it is possible that an allophony /sɾ/ ~ /ɾʰ/ is expending with specific roots towards a merger of /sɾ/ with /ɾʰ/, result of lexical frequency. Yet, the question of whether the change is happening with less or more frequent roots remains open.

It is possible that the change from /sɾ/ to /ɾʰ/ has taken place earlier in villages of the plain under intense contact with Nepali (like in Handikhola River basin) and that this change still gradually spreads North towards the hills along the Manahari and Lothar River basins. Another possibility is that the correspondences 2 and 3 come from two different proto-phonemes since other additional fricative clusters are reconstructed back to PTH, specifically: /*śr/ and /*źr/ (§ 2.3.5.7).

Table 43. Sound correspondences involving /sɾ/ and /ɾʰ/

	LOTHAR		MANAHARI	HANDIKHOLA	examples	
	RAP-13	RAK-6	RAK-7-8	MAN-4	root	gloss
1 *sɾ	/sɾ/	/sɾ/	/sɾ/	/ɾʰ/	sɾak- ~ ɾʰak-	'spray, slash'
2	/sɾ/	/sɾ/	/sɾ/ ~ /ɾʰ/	/ɾʰ/	sɾʌs- ~ ɾʰʌs-	'choke'
3	/sɾ/ ~ /ɾʰ/	/sɾ/ ~ /ɾʰ/	/sɾ/ ~ /ɾʰ/	/ɾʰ/	sɾek- ~ ɾʰek-	'smell good'
4 *ɾʰ	/ɾʰ/	/ɾʰ/	/ɾʰ/	/ɾʰ/	ɾʰʌp-	'fall, step into'

2.3.5.3.　　　　Velar consonant cluster onsets

There are five alveolar cluster onsets formed with stop and nasal manners of articulation: /kɾ/, /gɾ/, /kl/, /gl/ and /ŋɾ/. They are illustrated in Table 44. Similarly to bilabial consonant clusters, the velar nasal is only attested with the phoneme /ɾ/. Compared with /mɾ/ for which there are more examples of roots, only two roots are found with a nasal velar consonant cluster onset. The root *ŋɾʌs-* 'stick to pot or pan (food)' is only attested in SIL (RAK-6). Elsewhere, the meaning of *ŋɾʌs-* has been encompassed by the more general one of the root *jʰum-* 'be burnt, smell like burnt.' The second root, *ŋɾjok-* 'collapse at knees,' is attested in variation with the form *gɾʰjok-*. These allomorphs are presented in § 2.3.5.4.

Table 44. Velar consonant cluster onsets

kl	*klam-*	'be perpendicular to the trunk (branch)'	**kɾ**	*kɾam-*	'heave'
gl	*glam-*	'be hard (soil)'	**gɾ**	*gɾom-*	'hug'
			ŋɾ	*ŋɾʌs-*	'stick to pot'

In the RAK-8 (MAI) variety described by Caughley (1982, 2016), the velar nasal cluster /ŋɾ/ appears in more roots but shows up to two allophones: a nasal allophone /mɾ/ and a velar allophone /gɾ/. I present all of the roots featuring /ŋɾ/ or its allophones in Table 45, as found in Caughley (2016).

In addition to cluster allophony, there is a great deal of variation in vowel quality as well. Whether the form of these roots varies targeting the cluster or the vowel, they were analyzed as being different roots with very similar meanings. This analysis seems unlikely for two reasons.

First, while all these forms presumably come from the area of MAI (RAK-8) (since not mentioned otherwise), one of their variants or a different root may be recognized as being used by speakers of a different variety the others may not. The same is observed when it comes to their meanings. This is the case for example with the root *ŋɾə-* (and allomorphs) 'partly recover (from illness),' realized *mɾo-* in RAP-13 and RAK-6. In this case, the speakers did not recognize the other forms of the root, either in terms of their pronunciation, or in terms of their variation in meaning. The second reason for

considering these apparently different roots as variants of a single root, concerns vowel quality. Chepang does show vocalic variation when the nucleus consists of a glide-vowel or a two vowels sequence. This is discussed in § 2.6.7.

The cluster allophony /ŋɾ/ ~ /mɾ/ ~ /gɾ/, /ŋɾ/ ~ /mɾ/, or /ŋɾ/ ~ /gɾ/ is nevertheless not attested with the roots *ŋɾək-* ~ *ŋrik-* 'groan, moan,' *ŋɾəs-* 'stick to pot or pan (burnt food),' *ŋɾaʔ-* 'mix together (cooked grains),' and *ŋren-* 'be listless, without appetite' in the vocabulary collected by Caughley (2016). Amongst these, only *ŋɾʌs-* (*ŋɾəs-*) 'stick to pot or pan (burnt food)' is used in SIL (RAK-6), while the cluster /ŋɾ/ is not attested in RAP-13.

As for the roots showing cluster allophony, I present in Table 46 the roots found in RAP-13 and RAK-6 with regard to their variants found in RAK-8 (MAI).

Table 45. Velar nasal cluster attested in RAK-8 as found in Caughley (2016)

ŋɾ ~ mɾ ~ gɾ	*ŋɾə-*	'partly recover (from illness)'
	ŋri-	'begin to recover (small child)'
	ŋro-	'recuperate, be less bedridden after sickness'
	ŋroŋ-	'recover, regain some strength'
	ŋrya-	'recover partly'
	ŋryo-	'recuperate'
	ŋre-	'be recuperating'
	mroŋ-	'recover'
	grya-	'recuperate, recover'
ŋɾ ~ mɾ ~ gɾ	*ŋɾəʔ-*	'be nearly similar, agreeable'
	ŋriʔ-	'agree a little (ideas, speech)'
	ŋryaʔ-	'agree, be similar (ideas, speech), be fluent (speech)'
	ŋreʔ-	'agree (ideas, speech), be fluent (speech)'
	mryaʔ-	'be compatible, in complete agreement, harmony'
	gryaʔ-	'agree, be similar (ideas, speech)'
ŋɾ	*ŋɾək-*	'groin, moan (with pain)'
	ŋrik-	'moan with pain (small child)'
ŋɾ ~ gɾ	*ŋruk-*	'fall, tumble with neck bent'
	ŋryok-	'collapse at knees'

	gryok-	'collapse at knees'
	gryoh-	'collapse at knees'
ŋɾ	*ŋɾəs-*	'stick to pot (burnt food)'
ŋɾ	*ŋɾaʔ-*	'mix together (cooked grains)'
ŋɾ ~ mɾ	*ŋɾayk-*	'pull with twist, to be caught by decease'
	mɾuyk-	'break off with twist (small objects)'
	mɾoyk-	'twist to break off'
ŋɾ	*ŋɾen-*	'be listless, without appetite'

Table 46. Velar nasal cluster variation in RAP-13 and RAK-6 with regard to RAK-8

	MANAHARI	LOTHAR		
	RAK-8	RAP-13	RAK-6	
ŋɾ ~ mɾ ~ gɾ	*mroŋ-*	*mɾo-*	*mɾo-*	'recover (from illness)'
ŋɾ ~ gɾ	*ŋɾyok-*	*ŋɾjok-*	*ŋɾjok-*	'collapse at knees'
	gryok-	*~ gɾʰjok-*	*~ gɾʰjok-*	
	gryoh-	*~ kɾjok-*	*~ kɾjok-*	
		~ kɾjuk-	*~ kɾjuk-*	
		~ gɾjuk-		
ŋɾ	*ŋɾəs-*	*jʰum-* (dif. root)	*ŋɾʌs-*	'stick to pot (burnt food)'
ŋɾ ~ mɾ	*mɾuyk-*	*mɾujk-*	*mɾujk-*	'twist, wring together'

By contrast with other liquid consonant clusters, there is more variation attested for the cluster /ŋɾ/. Nevertheless, this variation is more likely to be recent since the cluster /*ŋɾ/ has been reconstructed back to PTH (§ 2.3.5.7).

2.3.5.4. Breathy cluster onsets

Very few examples of onset clusters show an additional feature, breathiness. Half of the roots that feature these breathy clusters have allomorphs. All the examples found with this type of clusters are reported in Table 47.

Table 47. Marginal breathy consonant cluster onsets

br^h	*br^{h}jak-*	'become mature and leave (like bird leaving nest)'
	br^{h}jam- ~ b^{h}jam-	'get lost (not able to find the way back)'
pr^h	*pr^{h}eŋ- ~ p^{h}eŋ-*	'be spacious, clear, or with thin inner surface (tube)'
kr^h	*kr^{h}jat*	'shrub - *hypericum uralum*'
gr^h	*gr^{h}jok- ~ ŋrjok-*	'collapse at knees'

While breathy consonant cluster onsets are particularly rare in Chepang, there is no particular environment that could trigger the presence of breathiness. Hence, these breathy consonant clusters should be reconstructed back to Proto-Chepang (PC) (§ 2.4).

2.3.5.5. Velar and bilabial approximant cluster onsets

As introduced in § 2.2, the distribution of the bilabial approximant /w/ in the syllable structure is that of initial, medial and final consonant. As a medial consonant, it only occurs following the velar consonants /k/ and /g/. This differs from the palatal approximant /j/ which, although it occupies as well the initial and final positions of the syllable structure, it can occur after any type of consonant, as well as follow consonant clusters. That is, /j/ belongs to the nucleus while /w/ does not.

Two vowels sequences formed with /u/ are attested in the nucleus: /ui/ and /ue/. Phonetically, they are pronounced [ʉi] and [ʉe], respectively. These vowels sequences do not entail diphthongization. They can be characterized as being pronounced with a synizesis, that is, as two distinct vowels in a single syllable without forming a diphthong and without the presence of a consonant.

The two vowels sequences /ui/ and /ue/ contrast with /wi/ and /we/, although no strict minimal pair is found. For instance, the root /kwi/ [kwi] 'dog' contrasts with /kɾui-/

[krʉi] 'be tangled, curly,' and the root /gwej/ [gwej] ~ /go/ [go] 'yam, taro' with /kueʔ/ [kʉeʔ] 'fish-hook.' Only one strict minimal pair is attested contrasting /wi/ and /ui/ in initial position: /wiʔ/ 'blood' vs. /uiʔ-/ 'whistle slowly to attract prey (fish, crab).' I present this pair along with some almost minimal pairs contrasting /wi/ with /ui/, and /we/ with /ue/ in Table 48.

Table 48. /ui/ vs. /wi/ and /ue/ vs. /we/

	LOTHAR		meaning
	RAP-13	RAK-6	
wi	*kwi* [kwi]	id.	'dog'
	wiʔ [wiʔ]	id.	'blood'
	swi [swĩ] (<N.)	id.	'injection'
ui	*krʉi-* [krʉi]	id.	'be tangled, curly (hair)'
	uiʔ- [ʉiʔ]	*us-* [us] (< *us-*)	'whistle slowly to attract prey (fish, crab)'
	sʉiʔ [sʉiʔ]	id.	'knife'
we	*gwej-* [gwej] ~ *goj-* [goj]	id.	'yam, taro (gen.)'
ue	*keʔ* [keʔ]	*kueʔ* [kʉeʔ]	'fish hook'

The roots /kwi/ 'dog' and /gwej/ ~ /go/ 'yam, taro,' are part of a series of roots that feature other glide vowel sequences, such as /wa/ and /wo/, which only occur following the velar stops /k/ and /g/. These roots are possibly reflexes of PTH labiovelar consonant phonemes */kʷ/ and */gʷ/ or PTH velar clusters formed with /w/ as a medial consonant, such as /*kw/ and /*gw/. Roots featuring this structure have been reconstructed by Benedict (1972) and Matisoff (2003), some of which have clear reflexes in Chepang.

In Table 49, I present possible reflexes of these roots in Chepang. All the roots featuring /kw/ and /gw/ attested in the studied varieties are reported in Table 50.

Table 49. Chepang reflexes of PTH /*kʷ/ or /*kw/ and /*gʷ/ or /*gw/

Chepang	meaning	PTH	meaning	source
kwi	'dog'	*d-kʷəy-n	'dog'	M 03
		*kwiy, *kwəy	'dog'	B 72
gwa- ~ go-sja	'wild cat (edible)'	*gwa	'fox'	B 72, M 03
kwaŋ	'chin, jaw'	*kwaŋ	'mouth, palate, gums'	M 03
		*gwaŋ	'neck, throat'	M 03
		*(m-)ka ~ (s-)ka	'chin, jaw'	B 72
gwej ~ goj	'tuber, taro (gen.)'	*g/s-rwa	'yam, taro'	M 03
		*kywiy	'yam'	B 72

Table 50. Sequences /wi/, /wa/, /we/, /wo/

	LOTHAR		MANAHARI		meaning
	RAP-13	RAK-6	MAN-4	RAK-8[31]	
wi	kwi [kwi]	id.	kwi [kwi]	kwiʔ	'dog'
wa	kwal- [kwal]	id.	-	kwalʔ-	'beckon sb'
	kwaŋ [kwaŋ]	id.	kwaŋ [kwaŋ]	kwaŋ	'chin, jaw'
	gwa- ~ go-sja [gwa ~ go.sja]	id.	-	goʔ-sjaʔ	'wild cat'
	gwalaŋ [gwa.lãŋ]	id.	-	gwalaŋ	'large sickle'
wo	kwoɾ- [kwoɾ]	id. ~ gwoɾ- [~ gwoɾ]	-	kwərʔ-	'be huddled up'
we	gwej ~ goj [gwej ~ goj]	id.	gwej [gwej]	goyʔ	'tuber, taro (gen.)'

2.3.5.6.   Consonant clusters and glide vowel or two vowel sequences

Bilabial and velar consonant clusters may be followed by a glide vowel sequence formed with the palatal approximant or glide /j/ as an on-glide forming the nucleus, as shown in Table 51 and Table 52 respectively. There is no example attested for its combination with /pl/, nor are there examples of /j/ following an alveolar stop cluster. Only one example is

---

[31] Data from Caughley 2000; Caughley 2016. Some roots show a related but different meaning from the one I understood.

found with /j/ following the fricative alveolar cluster /sɾ/. This word yet presents an allophone /ɾʰ/: *sɾjas-* ~ *ɾʰjas-* 'have an uncontrolled urinary or fecal discharge.'

Table 51. Bilabial consonant cluster onsets and on-glide /j/ nucleus

bl	*bljat-*	[bleat]	'come out, become (flower)'
pɾ	*pɾjahj-*	[pɾeæhj]	'shuck, shell (corn seeds)'
bɾ	*bɾjaj-*	[bɾeæj]	'be spaced up (people, planted seeds)'
mɾ	*mɾja-*	[mɾea]	'be regenerated, have had enough sleep'

Table 52. Velar consonant cluster onsets and on-glide /j/ nucleus

kl	*kljut-*	[kljut]	'peel, shell (onion, garlic, banana)'
gl	*gljun-*	[gljun]	'bring down, take out'
kɾ	*kɾjuhn-*	[kɾjuhn]	'loose shape'
gɾ	*gɾjahm-*	[gɾeahm]	'act upon two things at once inadvertently'
ŋɾ	*ŋɾjok-*	[ŋɾjok]	'collapse at knee'

The bilabial approximant or glide /w/ is not present following consonant clusters, but the two vowels sequence /ui/. I present in Table 53 all the roots attested with the sequence /ui/ following consonant clusters.

Table 53. Consonant cluster onsets in combination with /ʉi/

mɾ	*mɾuik-*	[mɾʉik]	'twist, to wring together'
pl	*pluik-*	[plʉik]	'wear inside out (clothe)'
kɾ	*kɾui-*	[kɾʉi]	'bend sharply to break, change direction, mind'
kl	*kluik-*	[klʉik]	'be tangled, curly, matted (hair)'

Similarly to the sequences presented above, consonant clusters may be followed by glide vowel sequences formed with /j/ as an off-glide. Some examples of this type of root are reported in Table 54.

Table 54. Consonant cluster onsets and off-glide /j/ nucleus

pl	*plajk -*	[plæjk]	'cross ridge, reach summit and be out of sight'
pɾ	*pɾajk -*	[pɾæjk]	'pick up with fingers (to put it in mouth)'
mɾ	*mɾajk-*	[mɾæjk]	'stretch, twist (body)'
sɾ	*sɾʌjk*	[sɾʌjk]	'louse'
kɾ	*kɾajŋ-*	[kɾæjŋ]	'bump together'

When following a consonant, glide vowel sequences formed with a palatal on-glide belong to the nucleus, since they can also follow liquid consonant clusters like any other monophthong vowels or any other type of sequences forming the nucleus. That is, by contrast with /l/, /ɾ/, and /w/, the palatal glide is not analyzed as a medial consonant forming a cluster.

2.3.5.7. PTH consonant clusters

The consonant clusters formed with liquids /r/ and /l/ reconstructed by Benedict (1972: 37) and Matisoff (2003: 69) are presented in Table 55. All of these clusters but the ones highlighted in grey in the table, i.e. /*śr/, /*zl/, /*zr/, /*źr/, and /*ml/ are attested in Chepang and can safely be reconstructed back to Proto-Chepang (PC). Amongst those that are not attested, there could be a possible source for the different correspondences found with /sɾ/, as suggested in § 2.3.5.2.

Table 55. PTH liquid consonant clusters

	bilabial			**alveolar**		**velar**	
stop	*pl	*pr		*tr		*kl	*kr
	*bl	*br		*dr		*gl	*gr
fricative				*sr	*śr		
			*zl	*zr	*źr		
nasal	*ml	*mr					*ŋr

Jacques (2016a) reconstructs Proto-Kiranti verbal roots based on internal reconstructions of verbs of four languages: Wambule, Khaling, Bantawa, Limbu. Following Michailovsky's (1994) classification of Kiranti into four subgroups, one language was chosen for each subgroup: Wambule (Opgenort 2004a), Khaling (Jacques et al. 2012, 2015), Bantawa (Doornenbal 2009) and Limbu (van Driem 1987, Michailovsky 2002). In this paper (2016a: 17), he proposes the reconstruction of the following bilabial and velar stop clusters formed with the liquids /r/ and /l/, including a bilabial and velar aspirated stop: /*pr/, /pʰr/, /*pl/, /*br/, /*bl/, /*kr/, /*kʰr/, /*kl/, /*gr/, /*gl/ (2016a: 17). Amongst the verbal roots he reconstructs, only *kʰraːp 'cry, weep' has a Chepang cognate, the verb kɾjap- 'cry, weep.' An additional cognate is found with the reconstructed form of the Proto-Limbu verb *pʰaːks 'untie,' which corresponds to pʰas- or pʰan- (arch.) 'untie' in Chepang. This root is reconstructed with a cluster onset in Proto-Kiranti: *pʰraːk. While it is clear that complex onsets formed with liquids /r/ and /l/ can be reconstructed back to PTH, a narrower comparison of roots between lower level clades is necessary to provide a better understanding of the proto-forms and their evolutions.

In PTH, and TH languages the status of the approximants /j/ and /w/ with regard to the syllable structure is problematic. Matisoff (2003: 62) poses the question of whether they are part of the initial or the rhyme. He answers saying that "they are intrinsically "Janus-headed," looking backwards and forwards at the same time" and gives examples of his reconstruction of Proto-Lolo-Burmese. Chepang's structural distribution of the approximants /j/ and /w/ is interesting as it also sheds light on this issue. In Chepang, only liquid clusters and labiovelar consonants or velar clusters formed with /w/ can be reconstructed back to Proto-Chepang (PC). This contrasts with /j/, which belongs to the nucleus of the rhyme and not the onset, when preceded by a consonant or consonant cluster (§ 2.2, § 2.3.5.5, § 2.6).

Benedict (1972: 37-38) and Matisoff (2003: 63, 65) reconstruct two series of clusters formed with /j/ and /w/ as medial consonants. They are presented in Table 56. The only ones that I reconstruct for Proto-Chepang (PC) are /*kw/ (or /*kʷ/) and /*gw/ (or /*gʷ/) (§ 2.4).

Table 56. PTH approximant consonant clusters

	bilabial		alveolar		palatal	velar		glottal
stop	*pw	*py	*tw	*ty		*kw	*ky	
	*bw	*by	*dw	*dy		*gw	*gy	
fricative			*sw	*sy				*hw *hy
			*zw	*zy				
			*tsw	*tsy				
			*dzw	*dzy				
nasal	*mw	*my	*nw	*ny		*ŋw	*ŋy	
liquids			*rw	*ry				
			*lw	*ly				
approximant					*yw			

2.3.6. Morphophonology of consonants

In this sub-section, I present the morphophonological processes attested for consonants within roots (§ 2.3.6.1) and at morpheme boundaries within words (§ 2.3.6.2).

2.3.6.1. Within roots

Within roots, the following morphophonological change is attested:
- Alveolar Fricative Palatalization (free variation)

2.3.6.1.1. Alveolar Fricative Palatalization (free variation)

When the alveolar fricative /s/ is followed by a high front vowel /i/, it may entail palatalization and be pronounced like an alveolo-palatal fricative [ɕ].

(1)   s > ɕ / ___ i, j

        ɾam-siʔ   [ɾam.ɕiʔ]   sickle-CL

        sjaw=sa   [ɕjaw.sa]   become=NMZ1

2.3.6.2.   At morpheme boundaries within words (internal sandhi)

At morpheme boundaries, the following morphophonological changes are attested:

        - Metathesis (regular)

        - Alveolar Fricative Palatalization and Aspiration (free variation)

2.3.6.2.1.   Metathesis (regular)

The glottal fricative /h/ may undergo metathesis with the following consonant when preceding a final sonorant consonant. This is not observed with the glottal stop.

(2)   h C$_{[sonorant]}$ > C$_{[sonorant]}$ h / (C) V ___ C$_{[sonorant]}$ + V

        dʑahŋ-alaŋ   [dʑãŋ.ha.lãŋ]   do-1.PST

2.3.6.2.2.   Alveolar Fricative Palatalization and Aspiration (free variation)

When the alveolar fricative consonant /s/ follows an alveolar nasal or stop consonant, it may become palatalized and aspirated. This change is only so far observed in RAP-13 (GUN, TAP, POL).

(3)   s > tɕʰ / C$_{[alveolar][+nasal/+stop]}$ + ___

        bun-si   [bun.tɕʰi]   *Ougeinia.oojeinense*-tree

        ten-sjahŋ   [ten.tɕʰjahŋ]   today-tomorrow ('nowadays')

        got-sa   [got.tɕʰa]   call-NMZ

        got-si   [got.tɕʰi]   orchid-tree

2.4. Proto-Chepang consonants and consonant clusters

Based on the Chepang consonant phonemes and clusters attested in the studied varieties, I suggest the reconstruction of the Proto-Chepang (PC) consonants and consonant clusters as presented in Table 57 and Table 58, respectively. At syllable coda, I reconstruct voiceless non-aspirated stops, fricatives, and non-breathy sonorants are reconstructed. The glottal stop is only reconstructed in coda position.

Table 57. Proto-Chepang consonants

* Proto-Chepang	bil.		alv.		alv-pal.		pal.	vel.		gl.
stop	*p	*b	*t	*d				*k	*g	*ʔ
asp. or breathy	*pʰ	*bʰ	*tʰ	*dʰ				*kʰ	*gʰ	
fricative			*s							*h
affricate					*tɕ	*dʑ				
asp. or breathy					*tɕʰ	*dʑʰ				
nasal	*m		*n						*ŋ	
breathy	*mʰ		*nʰ						*ŋʰ	
flap			*ɾ							
breathy			*ɾʰ							
lat. approximant			*l							
breathy			*lʰ							
approximant	*w						*j			
breathy	*wʰ						*jʰ			

Table 58. Proto-Chepang consonant clusters

* Proto-Chepang	bilabial			alveolar	velar			
stop	*pl	*pɾ	*pɾʰ	*tɾ	*kl	*kɾ	*kɾʰ	*kw
	*bl	*bɾ	*bɾʰ	*dɾ	*gl	*gɾ	(*gɾʰ)	*gw
fricative				*sɾ				
				*ʔ / other change				
nasal	*mɾ				*ŋɾ			

2.5. Vowels

This section describes the vowels of Chepang. I start by presenting the monophthong vowel phonemes in § 2.5.1, and I then describe the morphophonological changes targeting monophthong vowels in § 2.5.2.

2.5.1. Monophthong vowel phonemes

There are 6 monophthong vowel phonemes in Chepang: two front vowels /i/ (close or high) and /e/ (mid-close); one central vowel /a/ (low), and three back vowels /u/ (close or high, rounded), /o/ (close-mid, rounded), and /ʌ/ (open-mid, unrounded). They are presented in Table 59. The phonemic status of these vowels is illustrated through a minimal set of roots in Table 60. These roots are formed with a flap consonant /ɾ/ at onset and a voiceless velar stop consonant /k/ at coda.

Table 59. Vowel phoneme inventory

	front	central	back	
	unrounded	unrounded	unrounded	rounded
close / high	i			u
close-mid	e			o
open-mid			ʌ	
open / low		a		

Table 60. Minimal set for vowels in /ɾ _ k/ environment

i	*ɾik-*	'tie split cane to make roof'
e	*ɾek-*	'notify sb that sth if off or wrong on their body or in their work'
ʌ	*ɾʌk-*	'pull out entirely (grass clump, stones)'
a	*ɾak-*	'close sth with lock or hinge (door, window, box)'
o	*ɾok-*	'chili, be hot, spicy'
u	*ɾuk-*	'follow after running, chase, hunt'

To show the distribution of the 6 vowel phonemes within the vowel space, I plotted the vowel formant values measured for the roots forming the minimal set presented in Table 60. I added to this set the vowel formant values found through another set of roots distinguishing vowels in the same environment at coda (voiceless velar stop /k/). The only root that may not per se fit this environment is the dissyllabic word *tokɾak* 'toad' which can be analyzed either as an open [to.kɾak] or closed [tok.ɾak] syllable. This additional set of roots is presented in Table 61.

Table 61. Almost minimal set for vowels in /t _ k/ environment

i	*tik-ŋol*	'hornet (lesser banded)'
e	*tek-*	'land and rest on sth (after flying)'
ʌ	*tʌk*	'small mattock'
a	*tak-*	'pierce, let happen'
o	*tokɾak*	'toad'
u	*tuk*	'belly'

Figure 18, Figure 19, and Figure 20 illustrate the acoustic vowel space of 4 speakers of the Lothar varieties (3 female speakers (GUN, TAP, SIL) and 1 male speaker (TAP), age between 19 and 35 yo). The three figures present the result of the formant values extracted from 12 words repeated three times in isolation, i.e., a total of 144 individual vowel tokens or 24 tokens per vowel. Vowel formant information was extracted with a Praat (R Core Team 2021) script. F1 and F2 values were captured at vowel midpoint. The formant settings were set at 5500 Hz max for the female speakers and 5000 Hz for the male speaker. The formant values were plotted in R (Wickham et al. 2021) using two packages, DPLYR (Wickham 2016) and GGPLOT2 (Wickham 2016). F1 values were plotted on the *y*-axis and F2 values on the *x*-axis. Figure 18 and Figure 19 present the result of raw vowel formant values: Figure 18 shows all individual vowel tokens and Figure 19 the mean value of these tokens per vowel. Figure 20 presents normalized vowel formant values by speakers.

Figure 18. Chepang vowel space (raw F1 and F2 formant values)
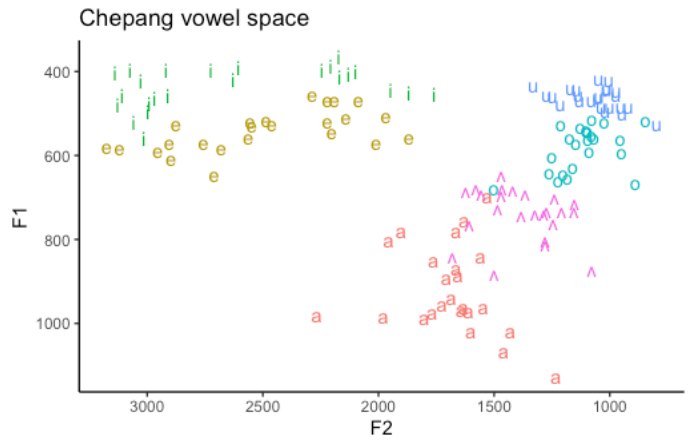
Figure 19. Chepang vowel space (mean of raw F1 and F2 formant values)
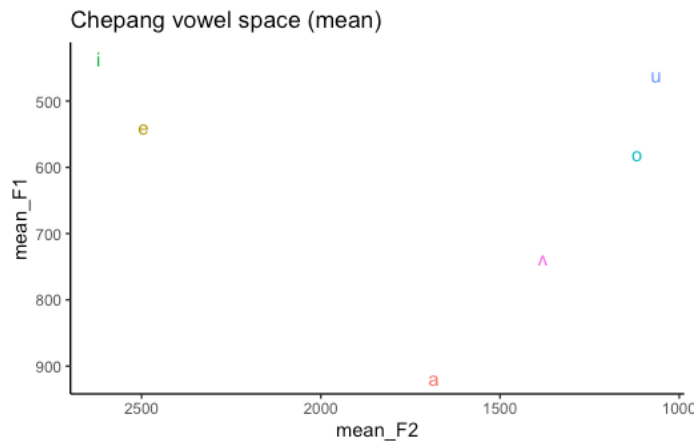
Figure 20. Chepang vowel space (normalized F1 and F2 formant values)
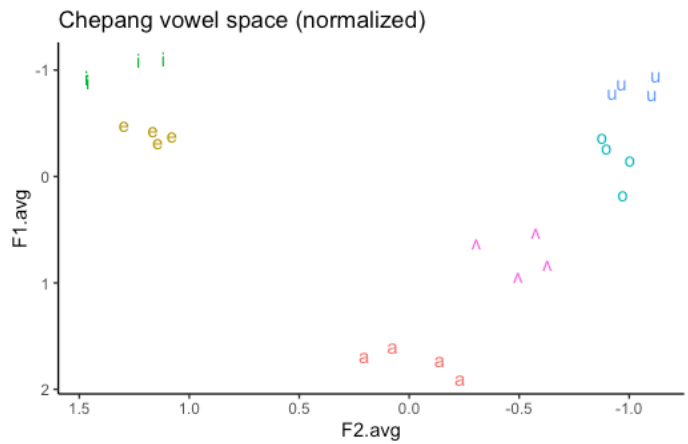
2.5.2. Morphophonology of monophthong vowels

As explained in § 1.9, morphophonological changes can be characterized as different types: regular, in free variation, or sporadic. In this section, I describe the morphophonological changes that affect vowels, and specify the type of change. The morphophonological changes attested for vowels in combination with glides are described in § 2.6.7.

2.5.2.1. Within roots

Within the root, the following morphophonological changes are attested:
- Nasalization (regular)
- Raising and backing (sporadic)

2.5.2.1.1. Nasalization (regular)

Nasalization is not contrastive. Vowels are nasalized when followed by a nasal velar consonant /ŋ/ at syllable coda, as in (4). Nasalization is also present on words borrowed from Nepali, as illustrated in (5).

(4)   V > Ṽ / (onset) __ C[nasal][velar]

     niŋ   [nĩŋ]   2PL

     tʰeŋ- [tʰẽŋ] 'count'

     paŋ   [pãŋ]   'husband's younger brother'

     huŋ   [hũŋ]   'yam (*Dioscorea hispida dannst*)'

     ɾʰoŋ- [ɾʰõŋ] 'empale, stab'

     bʰʌ̃ŋ- [bʰʌ̃ŋ] 'collapse, fall down; destroy, demolish'

(5)   *tʰaw* [tʰãw] 'place' ठाउँ <ṭhāuṃ> (<N.)

2.5.2.1.2.   Raising and backing (sporadic)

The low central vowel /a/ forming the nucleus of the enclitic morpheme /=ma/ ADD can rise and be pronounced more back, like a close-mid back unrounded vowel /ʌ/. This change often happens when the clitic /=ma/ ADD is unstressed, as in (6).

(6)   a > ʌ / =m __ [unstressed]
         =ma   [ma~mʌ]   ADD

2.5.2.2.   At morpheme boundaries within words (internal sandhi)

At morpheme boundaries, the following morphophonological changes are attested:
- Synizesis versus approximant or glide epenthesis (free variation)
- Palatalization (sporadic)
- Fronting (regular)
- Fronting and raising (and deletion) (sporadic)
- Raising and gliding (and deletion and de-gliding) (sporadic)
- Vowel harmony (and deletion) (sporadic)
- Vowel harmony (and deletion) and/or palatalization (sporadic)
- Vowel lowering and vowel harmony (sporadic)
- Regressive vowel harmony (sporadic)
- Loss of nasalization (regular)

2.5.2.2.1.   Synizesis versus approximant or glide epenthesis (free variation)

Open syllable roots suffixed or encliticized with a morpheme formed of a single vowel or with a vowel in initial position can be realized with a synizesis. The sequence of two vowels at morpheme boundary forms a single rhyme or nucleus. This is the main pattern attested in careful speech. Epenthesis of an approximant /j/ or /w/ may

nevertheless occur as well. However, epenthesis does not entail a hiatus or diaresis, i.e., resyllabification. The observations that follow are summarized in Table 62.

Table 62. Two vowels at morpheme boundaries

root	ERG, INST	2SG.IMP.INTR	PST	NMZ/PERF	3O/DIR/2SG.IMP.TR
(onset) V	=i	=ʌ	=a	=o	=u
i	[i ~ ː]	[ʌ ~ ʲʌ]	[a ~ ʲa]	[o ~ ʲo]	-
e	[i ~ j]	[ʌ]	[a]	[o]	[u]
ʌ	[j]	-	[a]	[o]	-
a	[i ~ j]	[ʌ]	[a]	[o]	[u]
o	[i]	[ʌ]	[a ~ ʷa]	[o ~ ʷo]	[u ~ ʷu]
u	[i]	[ʌ]	[a ~ ʷa]	[o ~ ʷo]	[u ~ ʷu]

Palatal approximant or glide epenthesis can occur when a root ending with a high front vowel /i/ is suffixed or encliticized with a morpheme presenting the following central or back vowels in initial position: /a/, /ʌ/ or /o/, as in (7).

Palatal approximant or glide epenthesis is not attested after the high front vowel /i/, but synizesis or vowel lengthening, as in (8). In this context, there is no example of a root followed by a morpheme formed with the high back rounded vowel /u/, which is specifically the morpheme =u 2SG.IMP.TR, since there is no transitive root featuring /i/ in open syllable found in our data.

(7)   Ø > ʲ / (onset) i ___ + V[central, back]
        di=a   [dia ~ diʲa]   rest=PST
        di=o   [dio ~ diʲo]   rest=NMZ:REL/PERF
        di=ʌ   [diʌ ~ diʲʌ]   rest=2SG.IMP.INTR

(8)   *ti*=i   [tii ~ tiː]   water=INST

Bilabial approximant or glide epenthesis can occur when a root ending with a back rounded vowel /u/ or /o/ is suffixed or encliticized with a morpheme presenting the following central or back vowels in initial position: /a/, /o/ or /u/, as in (9).

Approximant or glide epenthesis is not attested with the high front vowel /i/ or open-mid back unrounded vowel /ʌ/, but synizesis, as in (10). The feature of roundedness of the vowels /o/ and /u/ of the root triggers the epenthesis of [ʷ] before a suffix or enclitic that presents roundedness with the high and close-mid back vowels /u/ and /o/ and openness with the low central vowel /a/.

(9)   ∅ > ʷ / (onset) o, u ___ + V[rounded, open]

    to=a   [toa ~ toʷa]   tell_say=PST

    to=o   [too ~ toʷo]   tell_say=NMZ:REL/PERF

    to=u   [tou ~ toʷu]   tell_say=2SG.IMP.TR

    dʑu=a   [dʑua ~ dʑuʷa]   win=PST

    dʑu=o   [dʑuo ~ dʑuʷo]   win=NMZ:REL/PERF

    dʑu=u   [dʑuu ~ dʑuʷu]   win=2SG.IMP.TR

(10)   *ɾo=i*   [ɾoi]   flower=INST

    ŋo=ʌ   [ŋoʌ]   be drunk=2SG.IMP.INTR

    ɾu=i   [ɾui]   snake=ERG

    bʰu=ʌ   [bʰuʌ]   buzz_vibrate=2SG.IMP.INTR

When a root presents a close-mid front vowel /e/ or a low central vowel /a/ in an open syllable followed by a morpheme formed with a high front vowel /i/, such as the ergative or instrumental morpheme =i or verbal plural morpheme =i, the vowel /i/ can be palatalized, as in (11). In this context, by contrast with /e/ or /a/, roots featuring the open-mid central unrounded vowel /ʌ/ do not show synizesis, as in (12).

(11)   =i > j / (onset) e, a + ___

    me=i   [mei ~ mej]   tail=INST

    la=i   [læi ~ læj]   rope=INST

(12) *rʌ=j* [rʌj] liana=INST

When the roots presenting a vowel /e/, /ʌ/ or /a/ in an open syllable are followed by suffixes or enclitics formed by a single vowel or a vowel in initial position, the sequence is realized with synizesis. There is no example found in our data of a root featuring an open-mid back unrounded vowel /ʌ/ in open syllable suffixed or encliticized by a morpheme formed with the vowel /ʌ/ or /u/. These morphemes are specifically the morphemes =ʌ 2SG.IMP.INTR and =u 2SG.IMP.TR. The two verbal roots featuring /ʌ/ in open syllable do not take on such inflectional morphology: *lʌ-* 'be envious' and *prʌ-* 'be loose, break open.' This is illustrated in (13).

(13) *te=a* [tea] beg=PST

   te=o [teo] beg=NMZ:REL/PERF

   te=u [teu] beg=2SG.IMP.TR

   me=ʌ [meʌ] sing=2SG.IMP.INTR

   lʌ=a [lʌa] be.envious=PST

   lʌ=o [lʌo] be.envious=NMZ:REL/PERF

   ta=a [taa] throw=PST

   ta=o [taa] throw=NMZ:REL/PERF

   ta=u [tau] throw=2SG.IMP.TR

   tʰa=ʌ [tʰaʌ] appear=2SG.IMP.INTR

2.5.2.2.2.   Palatalization (sporadic)

The close-mid front vowel /e/ of the verbal root *dʑe-* 'eat' can undergo palatalization when suffixed or encliticized with a morpheme starting with a central or back vowel: /a/, /o/ or /u/. This pattern contrasts with the synizesis that is exclusively observed with other roots featuring a vowel /e/ in open syllable (§ 2.5.2.2.1). This is possible that frequency plays a role in this case, since this verbal root is considerably used. This is illustrated in (14).

(14) e > j / ___ + V[back]

　　　 dʑe=a　[dʑea ~ dʑja]　eat=PST

　　　 dʑe=o　[dʑeo ~ dʑjo]　eat=NMZ:REL/PERF

　　　 dʑe=u　[dʑeu ~ dʑju]　eat=2SG.IMP.TR

2.5.2.2.3.　Fronting (regular)

The low central vowel /a/ in an open syllable root is fronted when encliticized by a morpheme presenting a high front vowel /i/, as in (15).

(15)　a > æ / (onset) ___ + =i [i, j]

　　　 ŋa=i　　[ŋæi ~ ŋæj]　　1sg=ERG

　　　 ɾa=i　　[ɾæi ~ ɾæj]　　bamboo.tray=INST

　　　 =na=i　[næi ~ næj]　　=COP=PL

2.5.2.2.4.　Fronting and raising (and deletion) (sporadic)

The low central vowel /a/ of the non-past morpheme *=na* can further entail raising following fronting (§ 2.5.2.2.3) when encliticized with the plural morpheme *=i*. This morphophonological change can also be followed by the deletion of the plural morpheme *=i*. All three patterns are attested in free variation, resulting in four allomorphic constructions: *=na=i ~ =na=j ~ =ne=j ~ =ne*. This is illustrated in (16).

(16)　1. a > æ / =n ___ + =i [i, j]

　　　 =na=i [næi ~ næj]　　NPST=PL

　　　 2. æ > e / =n ___ + =i [j]

　　　 =na=i [nej]　　　　　NPST=PL

　　　 3. =i [j] > Ø / =ne + =__

　　　 =na=i [ne]　　　　　NPST=PL

2.5.2.2.5. Raising and gliding (and deletion and de-gliding) (sporadic)

The low central vowel /a/ of the imperfective morpheme =*na* can undergo raising when encliticized with the 3rd person object morpheme =*u*. The vowel /a/ becomes [ʌ] and the vowel /o/ glides to become [w]. This change can additionally be followed by the deletion of the vowel [ʌ] and the de-gliding of [w] which then becomes [u]. Both patterns are attested in free variation, resulting in three allomorphic constructions: =*na=u* ~ =*nʌ=w* ~ =*n=u*. This is illustrated in (17).

(17) 1. a > ʌ / =n ___ + =u
      (+) u > w / =nʌ + =___
      =*na=u* [nʌw]   impf=3SG.O
      2. ʌ > Ø / =n ___ + =w
      (+) w > u / =n + =___
      =*na=u* [nu]   impf=3SG.O

2.5.2.2.6. Vowel harmony (and deletion) (sporadic)

Progressive vowel harmony can occur on the locative morpheme =*haŋ* when encliticized to the distal pronoun *o*. In this case, the low central vowel /a/ of =*haŋ* becomes a close-mid back vowel [o]. This can be followed by the deletion of the distal pronoun *o* in addition to the deletion of the initial glottal fricative /h/ of the locative morpheme =*haŋ*. All three patterns are attested in free variation, resulting in four allomorphic constructions: *o=haŋ* ~ *o=hoŋ* ~ *hoŋ* ~ *oŋ*. This is illustrated in (18).

(18) 1. a > o / o + =h ___ ŋ
      o=haŋ [o.hõŋ]   DIST=LOC1
      2. o > Ø / ___ + =hoŋ
      o=haŋ [hõŋ]   DIST=LOC1
      3. =h > Ø / o + = ___ oŋ
      o=haŋ [õŋ]   DIST=LOC1

156

2.5.2.2.7.   Vowel harmony (and deletion) and/or palatalization (sporadic)

When the locative morpheme =*haŋ* is encliticized to the proximal pronoun *i*, this latter can undergo lowering and be realized as a close-mid front vowel [e]. This change can further entail a progressive vowel harmony on the morpheme =*haŋ* whose low central vowel /a/ becomes [e]. Similarly to the change attested for the distal pronoun /o/ encliticized with the locative morpheme =*haŋ* (§ 2.5.2.2.6), deletion of the proximal pronoun and of the initial glottal fricative of the morpheme =*haŋ* can occur as well. In addition to these, palatalization of the pronoun *i* [i ~ e] is also attested. All five patterns are attested in free variation, resulting in six allomorphic constructions: *i=haŋ ~ e=haŋ ~ e=heŋ ~ heŋ ~ eŋ ~ jʰ=aŋ ~ jʰ=eŋ.* This is illustrated in (19).

(19)   1. i > e / __ + =haŋ

      i=haŋ [e.hãŋ]   PROX=LOC1

      2. a > e / e + =h __ ŋ

      i=haŋ [e.hẽŋ]   PROX=LOC1

      3. e > Ø / __ + =heŋ

      i=haŋ [hẽŋ]   PROX=LOC1

      4. =h > Ø / + __ eŋ

      i=haŋ [ẽŋ]   PROX=LOC1

      5. i [i ~ e] > j / __ + =haŋ ~ =heŋ

      i=haŋ [jʰãŋ ~ jʰẽŋ]   PROX=LOC1

2.5.2.2.8.   Vowel lowering and vowel harmony (sporadic)

Vowel lowering is attested on the copula *mu-* when followed by the imperfective morpheme =*na*. The lowering of /u/ is possibly triggered by the vowel quality of the clitic =*na*. The vowel /a/ is open or low and central making the high back vowel /u/ become lower and unrounded. This results in two allomorphic constructions: *mu=na ~ mʌ=na.* This is illustrated in (20).

(20)   u > ʌ / m __ + =na

mu=na   [mʌ.na]   COP=NPST

The high back vowel /u/ of the copula *mu-* can undergo regressive vowel harmony and become [o] when encliticized with the perfect morpheme *=o*. This results in two allomorphic constructions: *mu=o ~ mo=o*. This is illustrated in (21).

(21)   u > o / m __ + =o

mu=o   [moo]   COP=NMZ:REL/PERF

2.5.2.2.9.   Regressive vowel harmony (sporadic)

The mid central vowel /ʌ/ of the interrogative morpheme *gʌ-* can undergo regressive vowel harmony, resulting in copying the vowel of the suffixing morpheme. This is illustrated in (22).

(22)   ʌ > V_i / g __ + =C V_i (C)

gʌ-tɕuk   [gu.tɕuk]   INT=much

gʌ-haŋ   [ga.haŋ]   INT=LOC1

gʌ-tʌ   [gʌ.tʌ]   INT=MAN

2.5.2.2.10.   Loss of nasalization (regular)

Roots featuring a velar nasal at coda undergo nasalization of the vowel (§ 2.5.2.2.10). When the root is suffixed or encliticized with a morpheme starting with a vowel the resyllabification that occurs entails the loss of nasalization. This is also the case with all glide vowel or two-vowel sequences. Nasalization is preserved when the root is encliticized or suffixed by a morpheme featuring a consonant at onset. This is illustrated by the following morphophonemic rule and examples from (23) to (26):

(23) Ṽ > V / (onset) ＿＿ C[velar nasal] + V

      jaŋ=o   [ja.ŋo]   taste=NMZ:REL/PERF

      kʰaŋ=u   [kʰa.ŋu]   cook=2SG.IMP.TR

      waŋ=a   [wa.ŋa]   come=PST

      baŋ=i   [ba.ŋi]   stone=INST

(24) Ṽ > V / (onset) ＿＿ C[velar nasal] + V

      djaŋ=a   [dea.ŋa]   be.near=PST

      ljuŋ=u   [lju.ŋu]   light.up.fire=2SG.IMP.TR

      ɾuiŋ=i   [ɾu.ŋi]   bamboo=INST

      sʌjŋ=a   [sʌ.ŋja]   be.rotten_smell.bad=PST

(25)   *jaŋ=sa*   [jãŋ.sa]   taste=NMZ1

      kʰaŋ=sa   [kʰãŋ.sa]   cook=NMZ1

      waŋ=sa   [wãŋ.sa]   come=NMZ1

      baŋ=ko   [bãŋ.ko]   stone=GEN

(26)   *djaŋ=sa*   [deãŋ.sa]   be.near=NMZ1

      ljuŋ=sa   [ljũŋ.sa]   light.up.fire=NMZ1

      ɾuiŋ=ko   [ɾuĩŋ.ko]   bamboo=GEN

      sʌjŋ=sa   [sʌj̃ŋ.sa]   be.rotten_smell.bad=NMZ1

2.6. Glide vowel or two-vowel sequences

A diphthong is primarily defined as a sequence of two vowels or a vowel and a glide that forms the nucleus of a single syllable (Catford 1988: 110–111; Ladefoged & Johnson 2011: 306). Based on this definition, it is problematic to consider that diphthongs exist in Chepang.

As described in § 2.3.1.8 and 2.3.2.5, the bilabial and palatal approximants or glides /j/ and /w/ are consonant phonemes which can hold the position of syllable onset or coda and contrast in phonation in initial position (non-breathy vs. breathy). In final

position, they can be preceded by laryngeal phonemes (/h/ and /ʔ/). In addition, /w/ and /j/ hold other positions in the syllable structure: /w/ is a medial consonant following velar consonants, while /j/ can occur as an on-glide or off-glide within the nucleus, when preceded and/or followed by consonants (§ 2.2, § 2.3.5.5).

If one considers diphthongs in Chepang as having a different phonemic status than the combination of a vowel and glide, or than two vowels combined together to form the nucleus, it would entail that: diphthongs would exist in a specific environment, i.e. between consonants, in a C_1 (L,w) V_1 G/V_2 (Lg) C_2 type syllable; or that diphthongs would exist in closed syllables with no on-glide in the nucleus, in a (C_1) (L,w) $V_{1[\text{-glide}]}$ G/V_2 (Lg) C_2 type syllable.

While these distributional configurations exist in some cases involving a palatal glide or approximant /j/, they are not attested for all sequences, and hence cannot constitute an argument for treating a diphthong as a specific phonemic unit. For instance, the bilabial approximant or glide /w/ is never found followed by a consonant, or the sequence /ui/ formed with two vowels realized with a synizesis is not restricted to appear in the nucleus between consonants but is also found in open syllables. In addition, the palatal approximant or glide /j/ in position of off-glide in a nucleus followed by a coda entails metathesis when the root is suffixed or encliticized by a morpheme whose onset starts with a vowel. This morphophonological change further underlies the independent character of the phoneme /j/ when present in a sequence. Lastly, if diphthongs exist, one has to posit that syllables formed with off-glides would be open syllables when not followed by a consonant. This would result in the existence of approximant and glides consonants only in initial position.

Recognizing the existence of diphthongs entails more complexity regarding the syllable structure. In fact, analyzing sequences formed with a glide and a vowel or two vowels without using the notion of diphthong better serves understanding their phonetic realizations and morphophonological changes, their distribution, as well as the origin of their formation. Hence, it is more accurate to say that a nucleus can be formed by a vowel and a glide sequence or by a two-vowel sequence, and that the former is diphthongized since it involves a glide, and that the latter is realized with a synizesis since it involves

two vowels. The sequences attested in the nucleus are as follows: /je/, /ja/, /jʌ/, /jo/, /ju/, /aj/, /ʌj/, /oj/, /ui/, /ue/.

In the following sub-sections, I describe the distribution of glide vowel or two vowels sequences, through the examination of the type of segments that can possibly combine, and of the type of phonemic environment within which these sequences appear or are restricted to appear. All the observations made in this section regarding the distribution and realization of glide vowel or two-vowel sequences in Chepang are summarized in Figure 21.

Figure 21. Glide vowel or two-vowel sequences

(C₁)	(L, w)	V₁	(G/V₂)	(Lg)	(C₂)	| nucleus
all	ɾ l w			h ʔ	all	
w/wʰ			i e a ʌ o	(Lg)	(C₂)	V
C₁	(L)		e a ʌ o j+V	(Lg)	w	V, j+V
C₁[+velar]	w < *kʷ~kw		i e a o	(Lg)	(C₂)	V
j/jʰ			u a o~ʌ	(Lg)	(C₂)	V
(C₁)	(L, w)		e a ʌ o j+V	(Lg)	j	V, j+V
(C₁)	(L)	u	i e	(Lg)	(C₂)	u + i / e
C₁	(L)	j	u o ʌ e	(Lg)	C₂	j + V
C₁[+cor.sibil][dorsal]		j	a	(Lg)	C₂	j + V
C₁[+coronal][labial]	(L[+coronal])	e	a	(Lg)	C₂	e + a
(C₁)	(L)		e a ʌ o j		C₂[+velar]	V + j

2.6.1. /j/ and /w/ at syllable onset and coda

Both /j/ and /w/ are consonants that can occur at syllable onset and coda. At onset, they contrast phonemically with their breathy counterparts, as shown in Table 63.

Table 63. Palatal and bilabial glides at syllable onset

j	*jak-*	'catch, stop, prevent from falling'
w	*wak-*	'break into pieces, split'
jʰ	*jʰak-*	'be bitter (like bitter gourd)'
wʰ	*wʰan-*	'be sharp, pointy'

Breathiness is not contrastive at syllable coda, where only /j/ and /w/ can occur, as shown in Table 64.

Table 64. Palatal and bilabial glides at syllable coda

j	*ŋaj-*	'find, meet'
w	*ŋaw-*	'roast'

At coda, /j/ and /w/ can be preceded by the glottal consonants /h/ and /ʔ/, as illustrated in Table 65.

Table 65. Palatal and bilabial glides in complex coda

hj	*ohj-*	'prick, pierce, scratch'
ʔj	*oʔj-*	'be friable (texture), floppy, not rigid'
hw	*tʌhw-*	'dig with mattock (*tʌk*)'
ʔw	*tʌʔw-*	'be bent back and up, curled up'

When /j/ and /w/ are at coda, all but high vowels are found in the nucleus, in addition to glide vowel sequences. Sets of illustrative roots are given in Table 66 and Table 67 respectively for the /j/ and /w/ at coda.

Table 66. Possible nucleus with palatal glide at syllable coda

ej	*gwej- ~ go-*	'yam, taro (gen.)'
aj	*haj-*	'do'
ʌj	*mʌj- ~ mi-*	'be small'
oj	*goj-*	'man, boy, son'
jaj	*njaj*	'flour (corn, millet, buckwheat, wheat)'
joj	*ljoj*	'hernia (umbilical), lump under skin'

Table 67. Possible nucleus with bilabial glide at syllable coda

ew	*ɾew-*	'miss, not be as good without'
aw	*dʰaw-*	'wash (clothes)'
ʌw	*dʰʌw-si*	*Cleistocalyx.operculatus*-tree
ow	*sɾow-*	'throw a jet of water (like with water gun)'
jaw	*sjaw-*	'become'
jʌw	*bljʌw-*	'be nauseated'

2.6.2. Sequences /ui/ and /ue/

As seen in § 2.3.5.5, the two vowels sequences /ui/ and /ue/ contrast with the combination of a bilabial approximant /w/ and a vowel. They primarily contrast in their pronunciation since /ui/ [ɥi] and /ue/ [ɥe] are realized with a synizesis. In addition, while /ui/ and /ue/ belong to the nucleus, /w/ is a consonant that can occupy the initial, medial and final position in the syllable structure (§ 2.2, § 2.3.5.5). Some examples of roots featuring /ui/ and /ue/ are given in Table 68.

Table 68. Sequences /ui/ and /ue/

	LOTHAR		MANAHARI		meaning
	RAP-13	RAK-6	MAN-4	RAK-8[32]	
ui	*ɾuiŋ* [ɾʉiŋ] *ɾiŋ* [ɾiŋ]	*ɾuiŋ* [ɾʉiŋ]	*ɾuiŋ* [ɾʉiŋ]	*ruyŋ*	'bamboo'
	ɾam-suiʔ ~ -siʔ [ɾam.sʉiʔ~siʔ]	id.	*ɾam-suiʔ* [ɾam.sʉiʔ]	*ram-suyʔ*	'sickle'
ue	*keʔ* [keʔ]	*kueʔ* [kʉeʔ]	*keʔ* [keʔ]	*keʔ*	'fish hook'

When following the glide /j/ at syllable onset, the sequences /ui/ and /ue/ and the glide may result in a labialized palatal approximant /ɥ/ or bilabial approximant /w/ followed by the last vocalic segment of the sequence. This raises the question of the nature of the consonant at syllable onset. It nevertheless seems more accurate to consider these roots as formed with a palatal glide /j/ at syllable onset. In fact, the variation attested in MAN-4 or RAK-6 with certain roots that entail the deletion of one of the vocalic segments leads to the presence of the palatal glide /j/ at onset followed by a monophthong vowel /u/ or /i/ in the case of /ui/, or /e/ in the case of /ue/. This is illustrated in Table 69.

Table 69. Palatal glide /j/ onset with sequences /ui/ and /ue/

	LOTHAR		MANAHARI		meaning
	RAP-13	RAK-6	MAN-4	RAK-8[33]	
ui	*juin* [jʉ.in] [ɥin] [win] id.		*jun* [jʉn]	*winʔ*	'bat'
ue	*jueɾ* [jʉ.eɾ] [ɥeɾ] [weɾ] id.		*jeɾ* [jeɾ]	*weɾ*	'hail'
	jues [jʉ.es] [ɥes] [wes]	*jes* [jes] ~ *jʌs* [jʌs]	*jes* [jes]	*wes*	'red jungle fowl'
	juel- [jʉ.el] [ɥel] [wel] id.		-	*wel*	'inspect'

---

[32] Caughley 2000; Caughley 2016

[33] Caughley 2000; Caughley 2016

164

2.6.3. On-glide /j/ sequences

When following a consonant or consonant cluster at syllable onset, on-glide /j/ sequences are found with all vowels but the high vowel /i/, as follows: /je/, /ja/, /jʌ/, /jo/, /ju/. This is illustrated in Table 70.

Table 70. On-glide /j/ sequences

	LOTHAR		MANAHARI		meaning
	RAP-13	RAK-6	MAN-4	RAK-8[34]	
je	*mje-ko ɾʰus* [mje]	id.	-	-	'coccyx'
ja	*gjap* [gjap]	id.	*gjap* [gjap]	*gyap*	'needle'
jʌ	*tjʌj-* [tjʌj]	id.	-	-	'get caught on (clothe, branch)'
jo	*tɕjok-* [tɕjok]	id.	*tɕjok-* [tɕjok]	*cyok-*	'wake up, get up (after sleeping)'
ju	*njup-* [njup ~ njʉp]	id.	*njup-* [njup]	*nyup-*	'be soft, smooth'

2.6.4. Off-glide /j/ sequences

Sequences with an off-glide /j/ can be formed in combination with any vowel except the high vowels /i/ and /u/, as follows: /ej/, /aj/, /ʌj/, /oj/. This is illustrated in Table 71. These sequences are only followed by velar consonant codas. This differentiates them from the sequence /ui/ or /ue/, which can be followed by a larger set of consonant codas: velar and glottal consonants, or alveolar nasals and lateral approximant /l/. However, similarly to /ui/, off-glide /j/ sequences can entail metathesis when the coda of the root is a velar stop /k/ or velar nasal /ŋ/ followed by a suffix or clitic starting with a vowel (§ 2.6.7.2.1). Metathesis is not attested with /ue/.

Table 71. Off-glide /j/ sequences

aj	*blajk-*	[blajk]	'feel lazy, tired, annoyed'
ʌj	*sʌjŋ-*	[sʌjŋ]	'be rotten, smell bad'
oj	*ojk-*	[ojk]	'flour made of roasted corn'

---

[34] Caughley 2000; Caughley 2016

2.6.5. Variation on roots formed with the palatal approximant or glide

Roots presenting a palatal glide sequence, or a palatal approximant or glide consonant /j/ at coda show more variation than roots formed with the bilabial approximant or glide /w/. The variation attested with /j/ is free and sporadic in that it does not affect all the roots featuring /j/ in the same environment. Some examples of these roots are presented in Table 72. This variation often results in the deletion of one of the segments (either the vowel or the glide) with or without an intermediate stage of change in the quality of the vowel, leaving a monophthong in the nucleus. This change is also attested with the vowel sequences /ui/ and /ue/. Morphophonological changes attested in glide vowel or two-vowel sequences are described in § 2.6.7.

Table 72. Variation on roots formed with the palatal approximant or glide

	LOTHAR		MANAHARI		meaning
	RAP-13	RAK-6	MAN-4	RAK-8[35]	
ʌjC	kʌjk ~ kek	kʌjk	kʌjk	kəyk	'neck, nape'
	gʌjŋ ~ geŋ	id.	-	gəyŋ	'trap'
	mʰʌjk ~ mʰek	id.	-	mʰʌjk	'have a lucid dream'
	srʌjk ~ srek	srʌjk	srek ~ rʰek	srʌjk	'louse'
	sʌjk ~ sek	id.	sejk	sʌjk	'tooth'
ʌj#	mʌj- ~ mi-	id.	mʌj-	məj- ~ mi-	'be small'
	mrʌj- ~ mri-	id.	-	-	'mix millet and corn (porridge)'
jaC	grjak- ~ grek-	grek-	grek-	grjah	'sticks, wigs, small pieces of wood'
	pjan- ~ pen-	pan-	pen-	pehn- ~ panʔn-	'tie to'
	ljak- ~ lek-	id.	-	lek-	'smear (cow-dung and water)'
jʌC	sjʌmbʌr ~ simbʌr	id.	-	-	'ginger (dry)'
jeC	mjen ~ min	min	-	min	'flying termite (swarmer)'
juC	lʰjuʔn- ~ lʰiʔn-	id.	-	-	'lift up (clothe, curtain, net)'

---

[35] Caughley 2000; Caughley 2016

2.6.6. Origin of the glides

The palatal glide /j/ is not found preceded or followed by the high front vowel /i/ and the bilabial glide /w/ does not occur preceded of followed by the high back vowel /u/. This shows that /j/ and /w/ are phonemically indistinguishable from the vowel segments /i/ and /u/ respectively, since their realization is predictable, i.e. in complementary distribution in specific environments: diphthongization would occur when the vowel /i/ or /u/ precedes a different vowel at syllable onset or follows a different vowel at coda. Hence, it is possible that /j/ and /w/ historically arose from /i/ and /u/, respectively, such as:

(27)　u > w / #__ V (coda), (onset) (V/G) V (Lg) ___ #

　　　 i > j 　/ #__ V (coda),

　　　　　　 / (onset) ___ V (Lg) (C) #,

　　　　　　 / (onset) V ___ C[+velar] #,

　　　　　　 / (onset) (V/G) V (Lg) ___ #

However, the sequences /ui/ [ʉi] and /ue/ [ʉe] do not fit this hypothesis, given that in these sequences, /i/ and /u/ are not diphthongized. That is, /ui/ is not pronounced /wi/ or /uj/, and /ue/ is not pronounced /we/, i.e., both undergo synizesis. In addition, as seen in § 2.3.5.5 and 2.6.2, /ui/ contrasts with /wi/ in initial position, and as a medial consonant, /w/ preceding a vowel only follows velar consonants. This suggests that the sequence /ui/ [ʉi] is more likely to be a different realization of /uj/ than of /wi/. In fact, /uj/ is not attested in the nucleus, while /oj/, /ʌj/, and /aj/ are. That is, if the glide /j/ in /oj/, /ʌj/, and /aj/ developed from diphthongization of the vowel /i/, diphthongization simply did not apply for /ui/ because of the fact that both /u/ and /i/ are high vowels.

There are nevertheless two counterarguments to this. First, the fact that /oj/, /ʌj/ and /aj/ as nuclei are only found preceding velar coda consonants, while /ui/ can precede velar consonants as well as the glottals /h/ and /ʔ/, and the alveolar nasal /n/. Second, the fact that it does not explain the form of /ue/ since both /ui/ and /ue/ share the same distribution and possible synizesis realization. Indeed, there could be a complementary

distribution between the two, result of a lowering of the vowel /i/ in the case of /ue/ in front of the consonants /l/, /s/, and /ɾ/ which are not found following /ui/. The consonants /l/, /s/, and /ɾ/ can be described as continuants considering that the flap /ɾ/ is realized more like a soft trill in final position. However, /ue/ is also attested in one root whose final consonant is not a continuant but a glottal stop /ʔ/. What is more, by contrast with /ui/, each root featuring the sequence /ue/ shows a variation that consists in either the deletion of /u/ in the case of /kueʔ/ ~ /keʔ/ 'fishhook,' or a change at the onset for the other roots starting with the palatal approximant /j/, resulting in a labialized palatal approximant /ɥ/ or bilabial approximant /w/.

While it is likely that approximant or glide consonants /j/ and /w/ originate from the high vowels /i/ and /u/ respectively, there remain questions regarding the forms of the sequences /ui/ and /ue/ in light of this hypothesis.

2.6.7. Morphophonology of glide vowel or two vowels sequences

As explained in § 1.9, morphophonological changes can be characterized as different types: regular, in free variation, or sporadic. In this section, I describe the morphophonological changes that affect vowels in combination with a glide and specify the type of change.

2.6.7.1.   Within roots

Within the root, attested morphophonological changes for glide vowel sequences are as follows:

- Fronting and raising (regular)
- Fronting or fronting and deletion (free variation)
- Raising (sporadic)
- Fronting (and raising), or fronting (and raising) and deletion (sporadic)
- Fronting, raising, and deletion (sporadic)
- Synizesis (regular)

- Synizesis (free variation)

2.6.7.1.1.　Fronting and raising (regular)

The central low vowel /a/ is fronted and raised to [æ] when followed by a palatal glide /j/ at syllable coda. This is illustrated in (28).

(28)　a > æ / (onset) __ j (coda)

aj　　[æj]　　'mother-in-law; siblings' mother-in-law'

maj　　[mæj] 'flesh, meat'

saj　　[sæj]　　'fruit, seed'

blajk-　[blæjk] 'feel lazy, tired, bored'

2.6.7.1.2.　Fronting or fronting and deletion (free variation)

The high back vowel /u/ can be fronted and pronounced like a high central rounded vowel [ʉ] when following a palatal glide /j/. Fronting occurs specifically when consonants are present at syllable onset and coda. This is illustrated in (29).

(29)　u > ʉ / C (Lg) j __ C

njum-　[njum ~ njʉm]　　　'be good, tasty'

kljut-　[kljut ~ kljʉt]　　　'peel, shell'

bjuk-　[bjuk ~ bjʉk]　　　'suck'

This change can be followed by a second change entailing the deletion of the palatal glide, as in (30).

(30)　1. u > ʉ / C (C) j __ C

　　　2. j > ∅ / C (C) __ ʉ C

njum-　[njum ~ njʉm ~ nʉm] 'be good, tasty'

kljut- [kljut ~ kljʉt ~ klʉt]   'peel, shell'

bjuk- [bjuk ~ bjʉk ~ bʉk]   'suck'

2.6.7.1.3.   Raising (sporadic)

The low central vowel /a/ followed by a palatal glide /j/ at syllable coda shows fronting and raising in its pronunciation, as described in § 2.6.7. When the glide vowel sequence /aj/ forms the nucleus and coda of the noun-phrasal clitics /=kaj/ DAT or /=paj/ DIS, the vowel /a/ does not necessarily show fronting and raising: it can also be found pronounced like a close-mid central vowel [ɐ] or even a schwa [ə], rather than a fronted vowel [æ]. This is illustrated in (31).

(31)   a > ɐ, ə / = C ___ j

=*kaj*   [kæj ~ kɐj ~ kəj]   DAT

=*paj*   [pæj ~ pɐj ~ pəj]   DIS

This change can be followed by a second change entailing the deletion of the palatal glide. Deletion most likely occurs when the syllable is unstressed. This is illustrated in (32).

(32)   1. a > ɐ, ə / = C ___ j

2. j > Ø / = C ɐ ~ ə ___ [unstressed]

=*kaj*   [kæj ~ kɐj ~ kəj ~ kɐ ~ kə]   DAT

=*paj*   [pæj ~ pɐj ~ pəj ~ pɐ ~ kə]   DIS

2.6.7.1.4.   Fronting (and raising), or fronting (and raising) and deletion (sporadic)

When the glide vowel sequence /ʌj/ forms the nucleus and coda of the noun-phrasal enclitic /=sʌj/ ABL, the back open-mid vowel /ʌ/ can be pronounced like a central

open-mid vowel [ɜ], a front open-mid vowel [ɛ], or a front close-mid vowel [e]. This is illustrated in (33).

(33)   ʌ > ɜ, ɛ, e / = s ___ j

   =sʌj   [sʌj ~ sɜj ~ sɛj ~ sej]   ABL

This change can be followed by a second change entailing the deletion of the palatal glide. Deletion most likely occurs when the syllable is unstressed. This is illustrated in (34).

(34)   1. ʌ > ɜ, ɛ, e / = s ___ j

   2. j > ∅ / = sɜ ~ sɛ ~ se ___ [unstressed]

   =sʌj   [sʌj ~ sɜj ~ sɛj ~ sej ~ sʌ ~ sɜ ~ sɛ ~ se]   ABL

The morphophonological change attested with /=sʌj/ is not attested with other roots or morphemes having the same syllabic structure, i.e., consisting of a palatal glide in coda position. Some of these roots are reported in (35).

(35)   bʌj-   [bʌj]   'give'

   glʌj   [glʌj]   'langur (grey)'

   hʌj-   [hʌj]   'peep, look inside'

   lʌj   [lʌj]   'own'

   mʌj-   [mʌj]   'banana'

   sʌj   [sʌj]   'porcupine'

2.6.7.1.5.   Fronting, raising, and deletion (sporadic)

There are roots or morphemes that show a possible change from /ʌj/ to [e] in close syllables. This morphophonological change, entailing fronting, raising, and deletion, does not present any transitional allophonic stage as those attested with /=sʌj/ ABL (see § 2.6.7.1.4). The roots which feature a change from /ʌj/ to [e] share the same

environment: a close syllable with a velar consonant at coda, which is the specific coda that precedes a vowel and palatal glide /j/ sequence in closed syllables (§ 2.6.4). In this environment, the nucleus formed by the monophthong [e] is only present when the suffixing or encliticizing morpheme that follows the root starts with a consonant. With suffixes or enclitics starting with a vowel, the glide vowel sequence [ʌj] occurs. This change is illustrated in (36).

(36)   ʌj > e / (onset) ____ C[velar] # + C rhyme

        dʑʌjk-   [dʑʌjk ~ dʑek] 'bite'

        gʌjŋ   [gʌ̃jŋ ~ gẽŋ]   'trap'

        kʌjk   [kʌjk ~ kek]   'neck, nape'

        sɾʌjk   [sɾʌjk ~ sɾek] 'louse'

        sʌjŋ-   [sʌ̃jŋ ~ sẽŋ]   'be rotten, smell bad'

        ʌjŋ-   [ʌ̃jŋ ~ ẽŋ]   'hurt (heart) after a shock in the back'

However, not all the roots presenting the same environment undergo this change. As with /=sʌj/, this change can be characterized as sporadic and in free variation. The roots in (37) do not feature a change from /ʌj/ to [e].

(37)   *tʌjk-*   [tʌjk]   'burn a field (slash and burn)'

        ɾʌjŋ ti [ɾʌ̃jŋ.ti]   'pus (of wound)'

        ʌjk-   [ʌjk]   'suffocate, be in pain'

2.6.7.1.6.   Synizesis (regular)

The sequence /ja/, can be either pronounced [ja], or undergo synizesis and be pronounced [ea]. This change is illustrated by the morphophonemic rule and example in (38). Additional examples are given in Table 73.

(38)   ja > ea / (onset) [labial][coronal-sibil] ___ (coda)

        djah- [deah] 'now'

These two surface realizations are in complementary distribution. [ja] is observed when the sequence follows alveolo-palatal affricates, velars, and the alveolar fricative /s/. Note that the alveolar fricative /s/ presents an alveolo-palatal allophone [ɕ] in front of /i/ and /j/. Hence, the sequence /ja/ can be described as pronounced [ja] when following dorsal consonants (/k, g/) and coronal sibilant consonants (s ~ ɕ, tɕ, tɕʰ, dʑ, dʑʰ). Elsewhere, following labial or coronal consonants that are not sibilants, /ja/ is pronounced [ea]. Illustrative roots for these two surface realizations are given in Table 73.

The sequence /ja/ is not attested following the glottal fricative /h/. However, it is possible that this type of structure be the source for the breathy palatal approximant /jʰ/.

Table 73. /ja/ allophones [ja] and [ea]

[ja]	sjan ~ ɕjan	[sjan ~ ɕjan]	'insect (gen.)'
	tɕjaʔm-	[tɕjaʔm]	'fall on, crush'
	tɕʰjak-	[tɕʰjak]	'mow, cut (grass, weed)'
	dʑjaɾ	[dʑjaɾ]	'yam (*Dioscoreaceae*)'
	dʑʰja-	[dʑʰja]	'heal (for the shaman)'
	kjah-	[kjah]	'weave'
	gjap	[gjap]	'needle'
[ea]	djah	[deah]	'now'
	tjaw-	[teaw]	'up'
	ɾjak	[ɾeak]	'tick'
	ljam	[leam]	'path, road'
	njam	[neam]	'sun'
	mjaŋ	[meaŋ]	'hair (of head)'
	kɾjap-	[kɾeap]	'cry, weep'
	kljaj-	[kleaj]	'be crooked, crossed (leg)'
	gɾjahm-	[gɾeahm]	'act upon two things at once'

2.6.7.1.7.   Synizesis (free variation)

When a root has a vowel /a/ followed by a palatal approximant or glide /j/ in coda position, the sequence can undergo synizesis. This is specifically attested with nominal roots lacking inflectional morphology. This change may be triggered by stress. This is illustrated in (39).

(39)   j > i / (onset) a ___ #[+stressed]

aj	[æj ~ æi]	'mother-in-law; siblings' mother-in-law'
maj	[mæj ~ mæi]	'flesh, meat'
naj	[næj ~ næi]	'clothe'
saj	[sæj ~ sæi]	'fruit, seed'

2.6.7.2.   At morpheme boundaries within words (internal sandhi)

The following morphophonological changes are attested at morpheme boundaries for glide vowel sequences:

- Metathesis (regular)

2.6.7.2.1.   Metathesis (regular)

Metathesis is attested at morpheme boundary for the glide vowel sequence formed with the off-glide /j/ and the two vowels sequence /ui/ (but not with /ue/). This morphophonological change is attested when the root is followed by a suffix or enclitic presenting a vowel in initial position. Thus, metathesis occurs with a limited set of vowel initial suffixes or enclitics which correspond to inflectional or derivational morphology associated with the part of speech of the root. The examples in (40) are verbal roots encliticized with the norminalizer or perfect morpheme =*o*, and nominal roots encliticized with the instrumental morpheme =*i*. This change is illustrated by the morphophonemic rule and examples in (40). More examples are given in Table 74.

(40)   (onset) V j, ui C > (onset) V C =/- j V, i / (C) V __ C =/-V

    ŋajk=o   [ŋa.kjo]   stir.flour.water=NMZ:REL/PERF

    ɾuiŋ=i   [ɾu.ŋi]   bamboo=INST

As a side note, when metathesis happens with a homorganic vowel as the initial element of the suffix or enclitic, such as the instrumental or ergative =*i*, no vowel lengthening, nor additional diphthongization is attested.

Table 74. Metathesis with off-glide vowel and /ui/ sequences

aj	*hajk-*	[hajk]	*hajk=o*	[ha.kjo]	carry_slinging-NMZ:REL/PERF
	kɾajŋ-	[kɾãjŋ]	*kɾajŋ=o*	[kɾa.ŋjo]	bump_together-NMZ:REL/PERF
ʌj	*ʥʌjk-*	[ʥʌjk]	*ʥʌjk=o*	[ʥʌ.kjo]	bite-NMZ:REL/PERF
	sʌjŋ-	[sʌ̃jŋ]	*sʌjŋ=o*	[sʌ.ŋjo]	smell_bad-NMZ:REL/PERF
ui	*kɾuik-*	[kɾʉik]	*kɾuik=o*	[kɾu.kjo]	bend_sharply-NMZ:REL/PERF
	ɾuiŋ	[ɾʉ̃iŋ]	*ɾuiŋ=i*	[ɾu.ŋi]	bamboo-INST
	luiʔ	[lʉiʔ]	*luiʔ=i*	[lu.ʔi]	bow-INST
ue	*juel*	[jʉel]	*juel=o*	[jʉe.lo]	examine_inspect-NMZ:REL/PERF

2.7.   Native rhymes

In this section, I present all the native rhymes attested in Chepang. The rhymes formed with a simple coda are reported in Table 75, and those with a complex coda in Table 76. All monophthong vowels can occur in a rhyme formed with any simple coda. The vowel /i/ is not found followed by the flap /ɾ/ or the vowel /e/ by the bilabial nasal /m/. In addition, as described in § 2.6, the high vowels /i/ and /u/ are never followed by the palatal or bilabial approximants /j/ and /w/, nor is the high-mid vowel /e/ followed by /j/.

Table 75. Native rhymes with simple coda

	p	t	k	ʔ	h	s	m	n	ŋ	l	ɾ	j	w
i	ip	it	ik	iʔ	ih	is	im	in	iŋ	il	-	-	-
e	ep	et	ek	eʔ	eh	es	em	en	eŋ	el	eɾ	ej	ew
a	ap	at	ak	aʔ	ah	as	am	an	aŋ	al	aɾ	aj	aw
ʌ	ʌp	ʌt	ʌk	ʌʔ	ʌh	ʌs	ʌm	ʌn	ʌŋ	ʌl	ʌɾ	ʌj	ʌw
o	op	ot	ok	oʔ	oh	os	om	on	oŋ	ol	oɾ	oj	ow
u	up	ut	uk	uʔ	uh	us	um	un	uŋ	ul	uɾ	-	-
je	-	-	jek	jeʔ	-	-	-	jen	-	-	-	-	jew
ja	jap	jat	jak	-	jah	jas	jam	jan	jaŋ	jal	jaɾ	jaj	jaw
jʌ	-	-	-	-	-	-	jʌm	jʌn	jʌŋ	-	-	jʌj	jʌw
jo	jop	jot	jok	joʔ	joh	jos	jom	jon	joŋ	-	joɾ	joj	jow
ju	jup	jut	juk	juʔ	-	jus	jum	jun	juŋ	-	juɾ	-	-
aj	-	-	ajk	-	-	-	-	-	ajŋ	-	-	-	-
ʌj	-	-	ʌjk	-	-	-	-	-	ʌjŋ	-	-	-	-
oj	-	-	ojk	-	-	-	-	-	-	-	-	-	-
ui	-	-	uik	uiʔ	uih	-	-	uin	uiŋ	-	-	-	-
ue	-	-	-	ueʔ	-	ues	-	-	-	uel	ueɾ	-	-
wi	-	-	-	-	-	-	-	-	-	-	-	-	-
we	-	-	-	-	-	-	-	-	-	-	-	wej	-
wa	-	-	-	-	-	-	-	-	waŋ	wal	waɾ	-	-
wo	-	-	-	-	-	-	-	-	-	-	woɾ	-	-

Table 76. Rhymes with complex coda

	hm	ʔm	hn	ʔn	hŋ	ʔŋ	hl	ʔl	hɾ	ʔɾ	hj	ʔj	hw	ʔw
i	-	-	-	iʔn	ihŋ	iʔŋ	-	-	-	-	-	-	-	-
e	ehm	-	ehn	eʔn	ehŋ	eʔŋ	ehl	-	-	-	-	-	ehw	eʔw
a	ahm	aʔm	ahn	aʔn	ahŋ	aʔŋ	ahl	aʔl	-	aʔɾ	ahj	aʔj	ahw	aʔw
ʌ	ʌhm	-	ʌhn	ʌʔn	ʌhŋ	ʌʔŋ	-	-	ʌhɾ	ʌʔɾ	ʌhj	ʌʔj	ʌhw	ʌʔw
o	-	oʔm	ohn	oʔn	ohŋ	oʔŋ	ohl	oʔl	ohɾ	-	ohj	oʔj	-	-
u	uhm	uʔm	uhn	uʔn	uhŋ	uʔŋ	uhl	uʔl	uhɾ	uʔɾ	uhj	-	-	-
je	-	-	-	-	-	-	-	-	-	-	-	-	-	jeʔw
ja	jahm	jaʔm	jahn	jaʔn	jahŋ	jaʔŋ	-	-	jahɾ	-	jahj	jaʔj	jahw	jaʔw
jʌ	-	-	-	-	jʌhŋ	-	-	-	-	-	-	-	-	-
jo	-	-	-	joʔn	-	-	-	-	johɾ	joʔɾ	-	-	-	-
ju	juhm	juʔm	juhn	juʔn	juhŋ	juʔŋ	-	juʔl	juhɾ	-	-	-	-	-
aj	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ʌj	-	-	-	-	ʌjhŋ	ʌjʔŋ	-	-	-	-	-	-	-	-
oj	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ui	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ue	-	-	-	-	-	-	-	-	-	-	-	-	-	-
wi	-	-	-	-	-	-	-	-	-	-	-	-	-	-
we	-	-	-	-	-	-	-	-	-	-	wehj	-	-	-
wa	-	-	-	-	-	-	-	-	-	-	-	-	-	-
wo	-	-	-	-	-	-	-	-	-	-	-	-	-	-

2.8. Glottal consonants realization and tonal contrast

In this section, I describe the realizations of the glottal stop /ʔ/ and glottal fricative /h/ in final and pre-final sonorant position. While there is no noticeable variation attested with the realization of the glottal fricative /h/, the realization of the glottal stop /ʔ/ varies amongst speakers. I start by giving a summary of the findings of the present study of laryngeal feature realizations (§ 2.8.1). I describe the main glottal stop realizations observed in final (§ 2.8.2) and pre-final sonorant positions (§ 2.8.3), and glottal fricative realizations in final and pre-final sonorant positions (§ 2.8.4). Finally, I show that some speakers (MAN-4) have completely lost the glottal stop in pre-final sonorant positions,

leaving no acoustic trace, not having replaced the original contrast by another, while having preserved it in final position on part of the lexicon (§ 2.8.6).

2.8.1. Summary of the findings

Chepang shows a three-way tonal contrast that correlates with the realization or absence of two laryngeal features in final and pre-sonorant final positions: the glottal fricative /h/ and glottal stop /ʔ/, as follows, where C=consonant, V=vowel, S=sonorant:

CV(C)	level / mid (slight fall mid to low) - modal
CVʔ(S)	high / mid-raising (mid to high) [36] - creaky
CVh(S)	raising (mid to high)-falling (high to mid) - breathy

This is schematized in Figure 22.

Figure 22. Three-way tonal contrast and voice quality

CV (C) level /mid (mid to low) modal   CV ʔ (S) high / raising creaky   CV h (S) raising-falling breathy
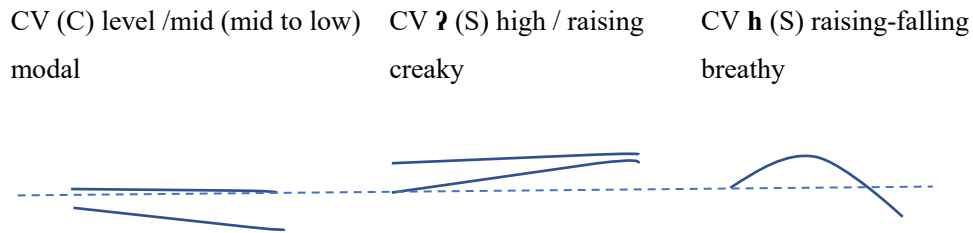
Three types of glottal stop /ʔ/ realizations are observed in final position while two types are observed in pre-sonorant position, and a single type of realization is observed for the glottal fricative /h/ in final and re-final sonorant position, as follows:

---

[36] Caughley (1982) analyzes the pitch associated with the glottal stop as high falling.

1- Glottal stop realization at coda

Type 1: glottal closure and release, creaky voice, no change in pitch

Type 2: glottal closure and release, creaky voice, high / mid-raising pitch

Type 3: high / mid-raising pitch, no glottal closure, no creaky voice

2-Glottal stop realization in pre-final sonorant position

Type 1: Glottal closure, creaky voice, high pitch, sonorant devoicing (+/-)

Type 2: High / mid-raising pitch

3-Glottal fricative realization at coda

Type 1: glottal fricative, raising-falling pitch

2.8.2. Glottal stop realizations at coda

Three main types of glottal stop realization are observed in final position:

- Type 1: Closure and release, creaky voice, no change in pitch

- Type 2: Closure and release, creaky voice, high pitch

- Type 3: High pitch

2.8.2.1. Type 1: Closure and release, creaky voice, no change in pitch

Figure 23 presents waveforms associated with a Type 1 realization of the glottal stop at coda in the root *laʔ* 'arrow' (b) against its absence in the root *la* 'rope' (a). Both roots were pronounced in isolation by a male speaker (RAK-8, 30-40 yo). The main visible acoustic cue for the glottal stop in the waveform is the glottal closure and release. The burst, or release of the airflow, may occur close to the end of the vowel segment. In other words, the closure of the glottal stop may present different length. Here the long closure is a possible result of the roots pronounced in isolation. This is also observed with other voiceless stops under similar conditions, such as with /k/ in the word *tʌk* 'small mattock' (Figure 24).

In addition to a visible glottal closure and release on the waveform, the shape of the final segment of the nucleus appears to be different, result of the manifestation of creaky voice. In Figure 25, the spectrograms of the same roots show differences in voice quality. Creaky voice is present at the end of the vowel of the root *la?* 'arrow' (b) where the voicing pulses appears further apart than in the vowel of the root *la* 'rope' (a).

Finally, Type 1 realization shows no difference in F0 (red) or intensity (yellow) on the vocalic segment in presence or absence of the glottal stop. The only difference is a visible rise in intensity (yellow) accompanying the glottal burst. While F0 is similar for both roots (mean 147 Hz), the F0 pitch tracker (red) fails to be detected as soon as creaky voice starts, i.e., on the final segment of the nucleus. This is observed in other languages, such as Mandarin Chinese (Yu 2010).

Figure 23. Waveform - Type 1 Glottal stop realization

(a) *la* 'rope'

MAI_la

Time (s)
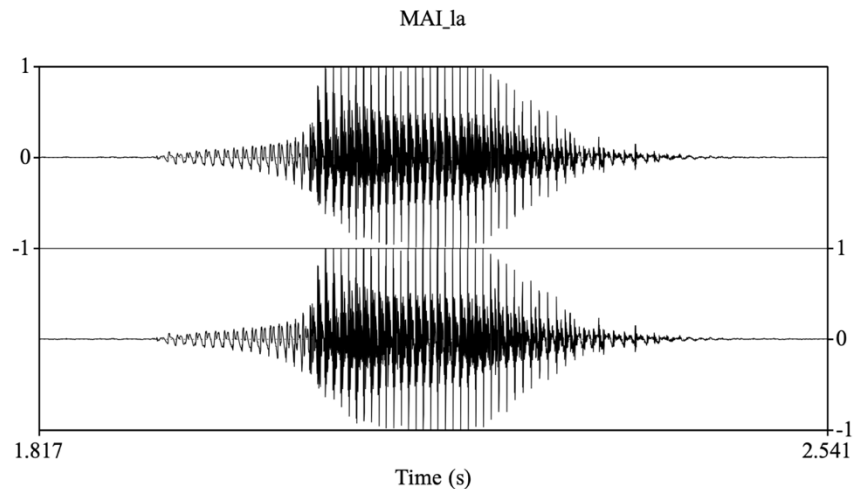
(b) *laʔ* 'arrow'
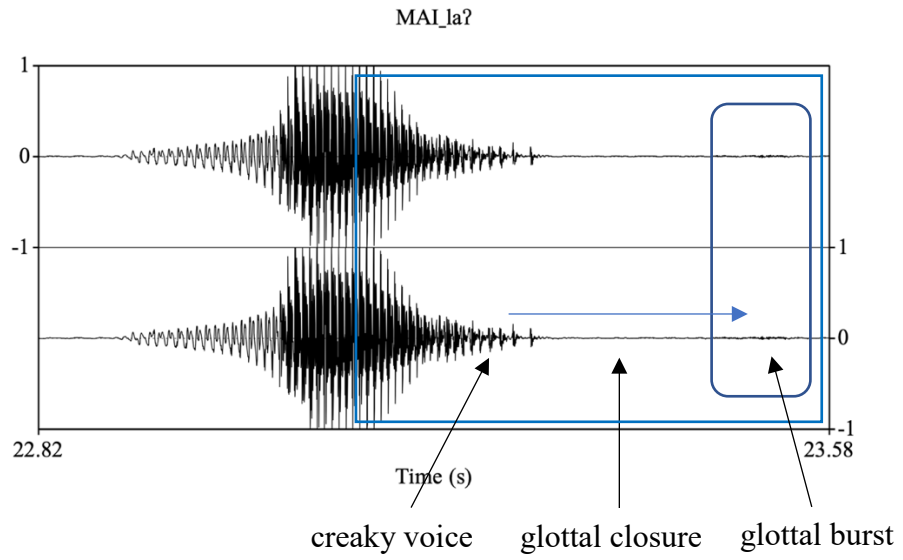
Figure 24. Waveform - Realization of /k/ at coda in isolation: *tʌk* 'small mattock'
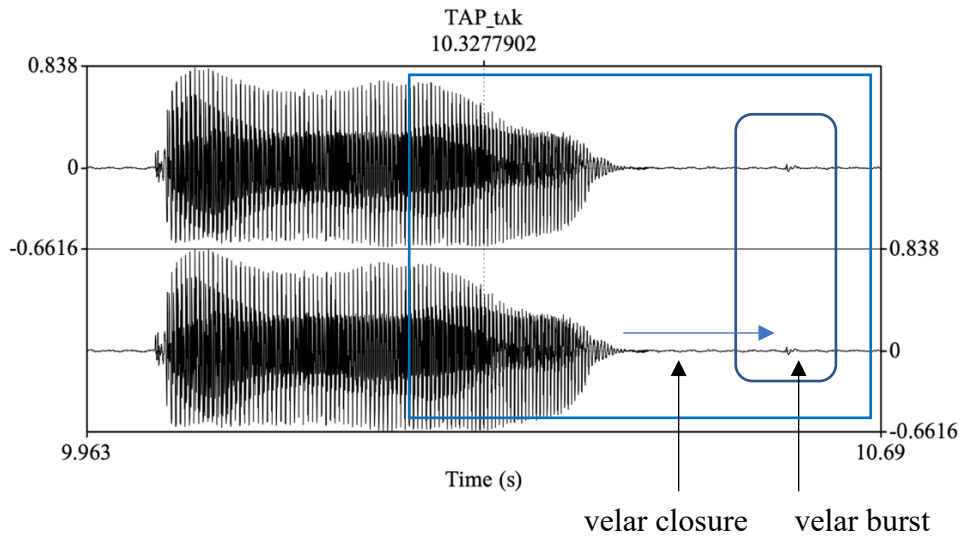
Figure 25. Spectrogram, pitch, and intensity track - Type 1 Glottal stop realization

(a) *la* 'rope'

(b) *laʔ* 'rope'
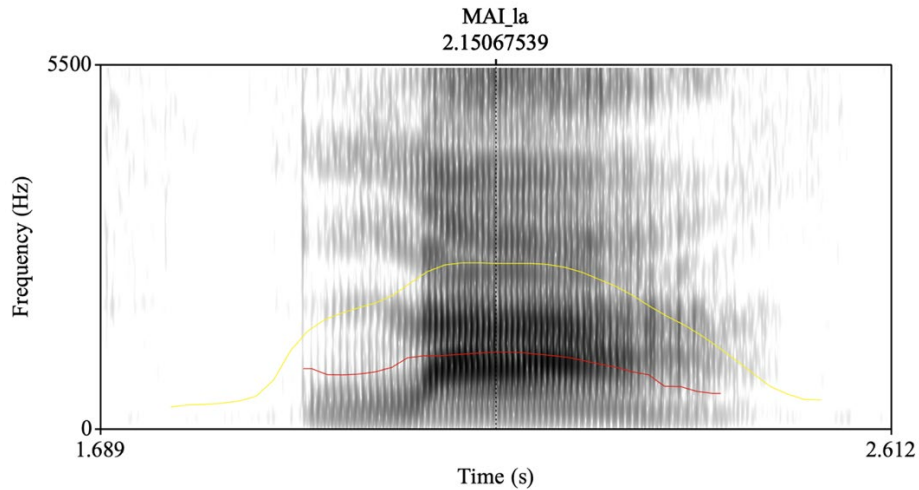
creaky voice   glottal burst
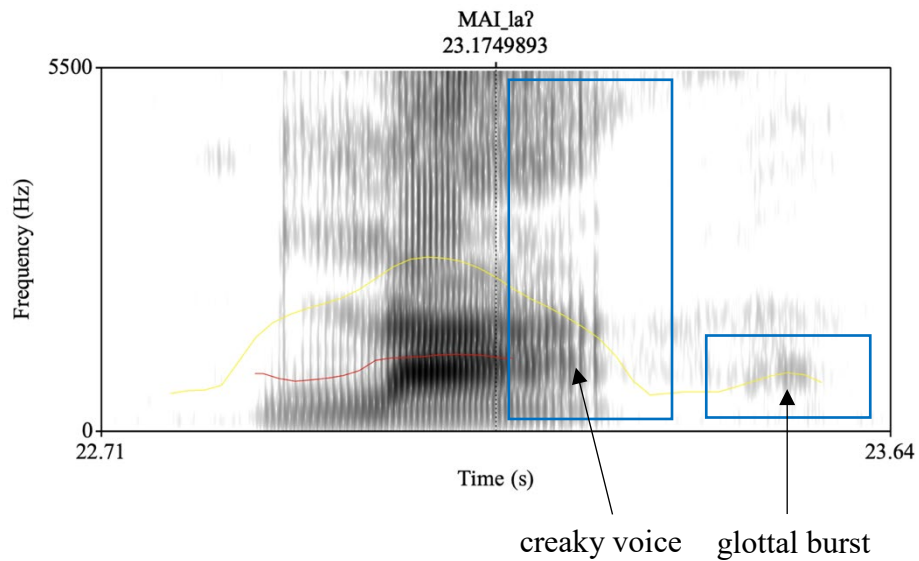
2.8.2.2.   Type 2: Closure and release, creaky voice, high pitch

The comparison of the waveforms and spectrograms of a Type 2 realization of the glottal stop in final position of the root *laʔ* 'arrow' (b) against its absence in the root *la* 'rope' (a), is presented in Figure 26 and Figure 27. Both roots were pronounced in isolation by a female speaker (RAP-13, 20-30 yo). Similarly to Type 1 realization, the glottal stop shows

closure and release as well as creaky voice. These acoustic cues are visible on both the waveform (Figure 26) and the spectrogram (Figure 27). Still, intensity (yellow) shows a rise at the burst or release of the glottal stop (Figure 27).

By contrast with Type 1 realization, high pitch is attested as an additional acoustic cue (

Figure 27). The two roots show a difference in F0 of around 20 Hz, with 251 Hz (mean) for *la?* 'arrow' and 226 Hz (mean) for *la* 'rope.' Here as well, the pitch tracker (red) fails to be detected when creaky voice starts.

Figure 26. Waveform - Type 2 Glottal stop realization

(a) *la* 'rope'

(b) *la?* 'rope'

creaky voice   glottal closure   glottal burst
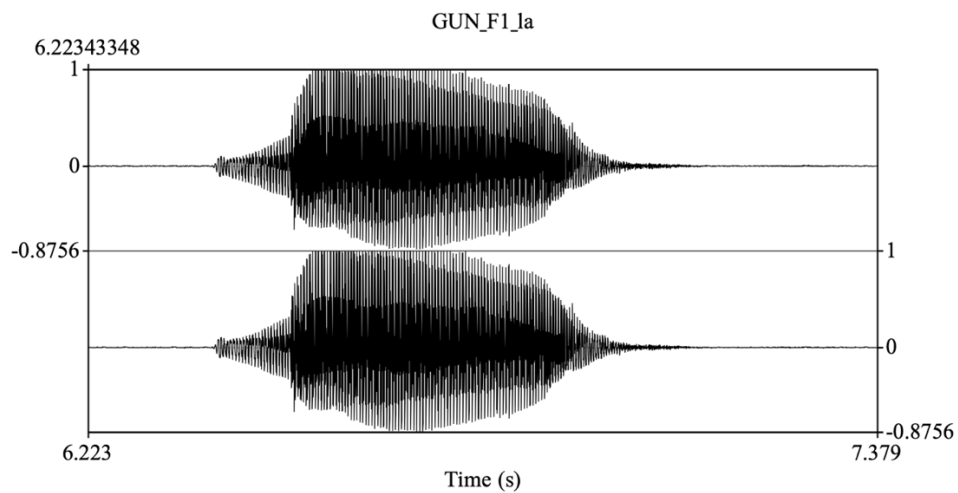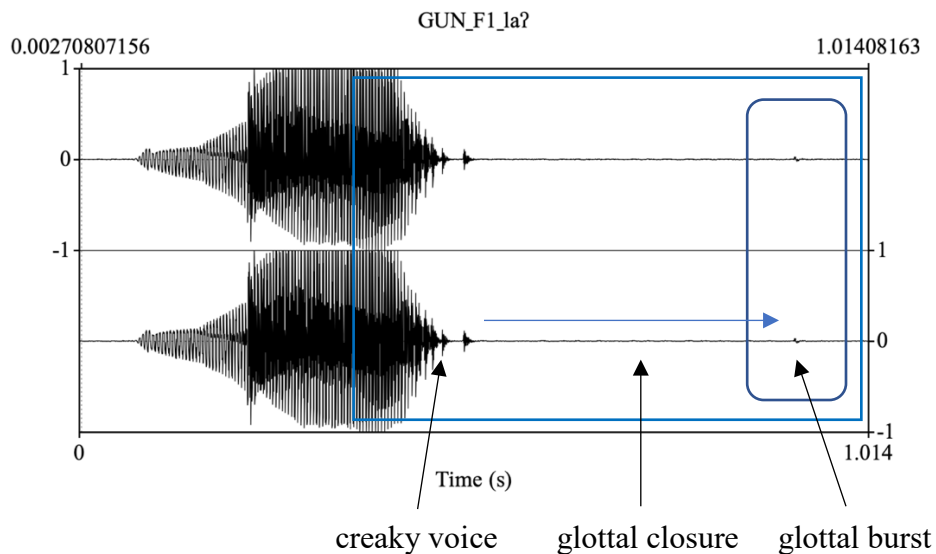
Figure 27. Spectrogram - Type 2 Glottal stop realization

(a) *la* 'rope'

GUN_F1_la

6.22343348

(b) *la?* 'arrow'

GUN_F1_la?

0.00270807156   1.01408163

high pitch   creaky voice   glottal burst
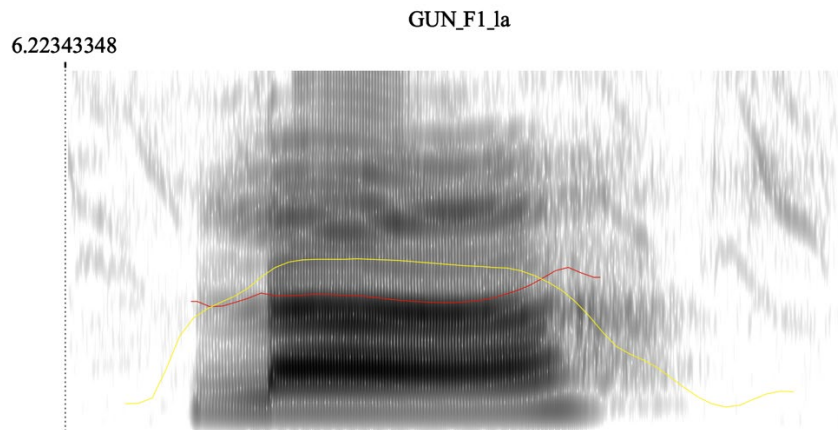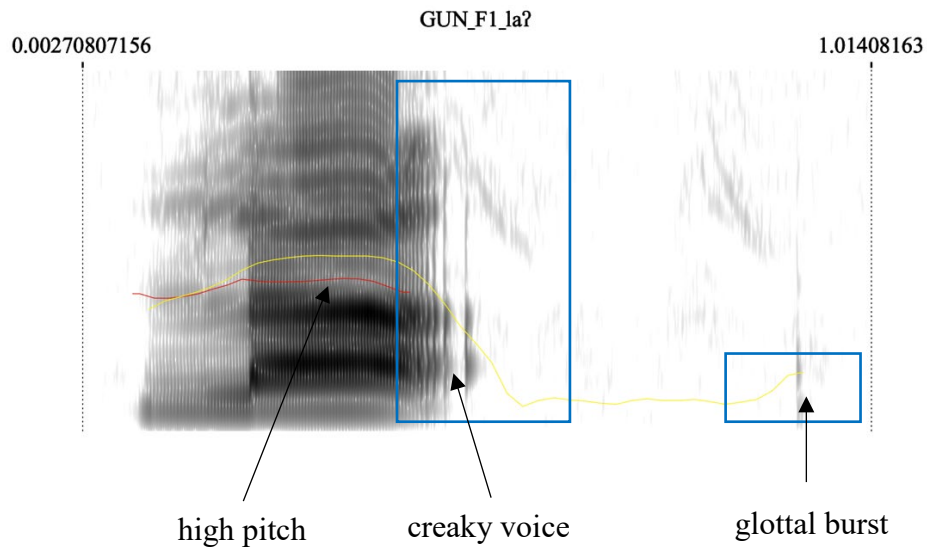
2.8.2.3.   Type 3: High pitch

Figure 28 and Figure 29 show the waveforms and spectrograms of a Type 3
realization of the glottal stop at coda of the root *la?* 'arrow' (b) against its absence in the
root *la* 'rope' (a). Both roots were pronounced in isolation by a female speaker (RAK-6,
30-40 yo). By contrast with Type 1 and 2, there is no trace of closure of the glottal stop in
Type 3, nor presence of creaky voice, as illustrated by the waveform (Figure 28) and the

184

spectrogram (Figure 29). The only acoustic cue left in Type 3 realization is the presence of a high pitch (Figure 29). The two roots show a difference in F0 of around 20 Hz, with 254 Hz (mean) for *laʔ* 'arrow' and 233 Hz (mean) for *la* 'rope.' Here as well, the pitch tracker (red) fails to be detected when creaky voice starts.

Figure 28. Waveform - Type 3 Glottal stop realization
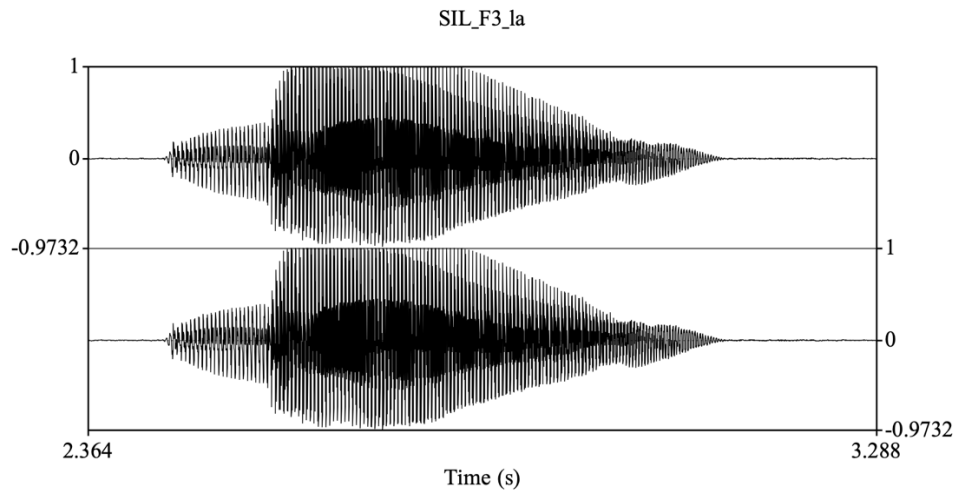
(a) *la* 'rope'

(b) *laʔ* 'arrow'
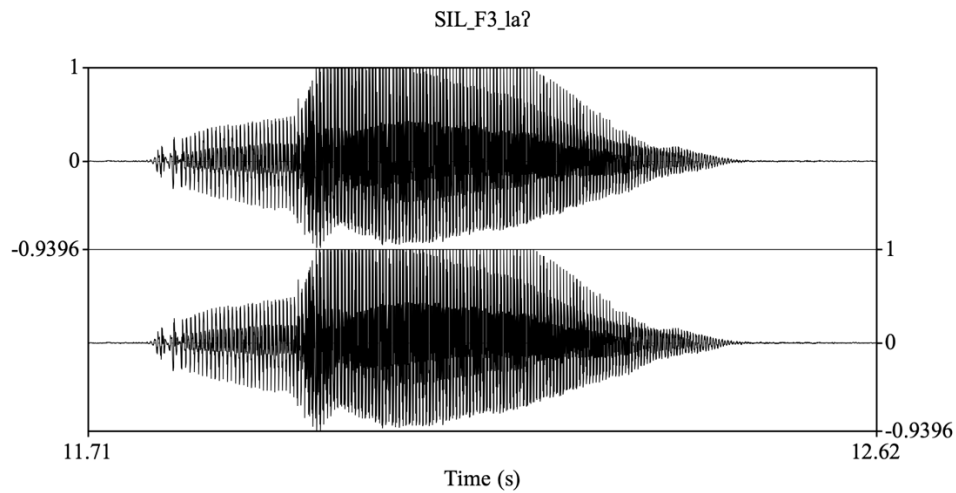
Figure 29. Spectrogram - Type 3 Glottal stop realization

(a) *la* 'rope'

SIL_F3_la

2.36376909   3.28776686

(b) *laʔ* 'arrow'

SIL_F3_laʔ

11.7142747   12.6177382
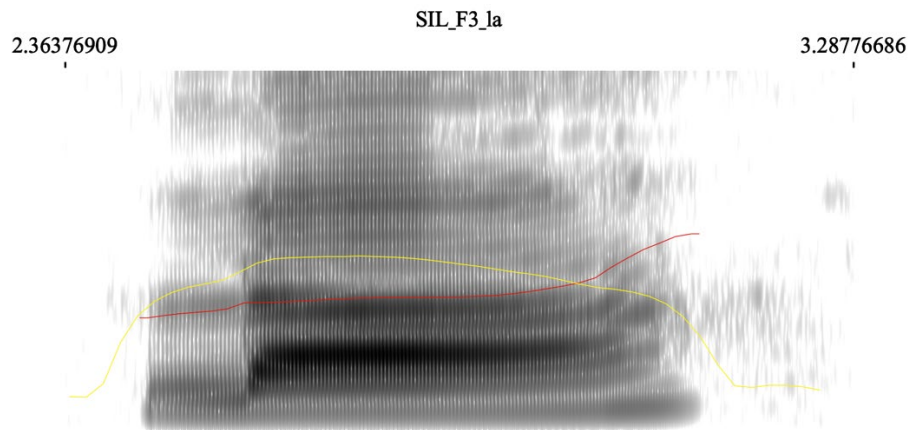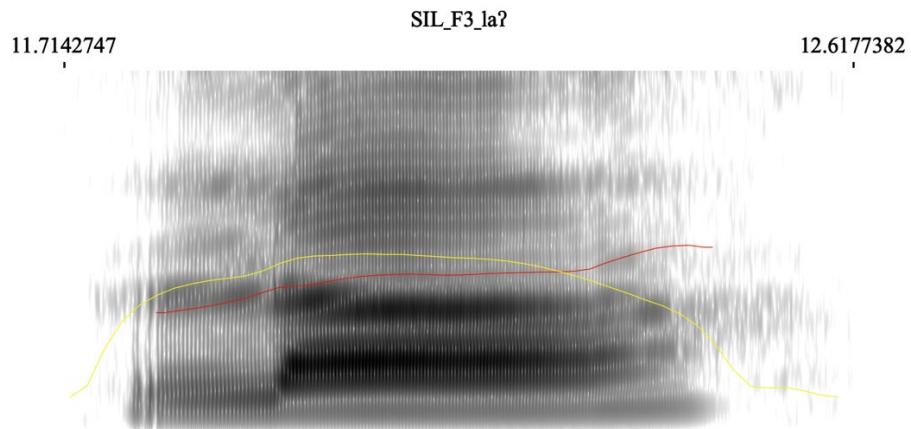
2.8.2.4.   Glottal stop in final position: closure, creaky voice, and F0

These three types of realization of the glottal stop in final position may be thought as a continuum, in that the three acoustic features, glottal closure in combination with creaky voice, and difference in F0, are interdependent in determining the presence or quality of one another. To sum up the observations made on the variation of the glottal stop realization in final position, we can say that the presence of a glottal closure and release is accompanied by creaky voice, regardless of the presence or absence of a

186

contrast in F0. However, there seems to be a difference in F0 as well regarding the absence or presence of creaky voice.

As we can see comparing Type 2 and 3 realizations, while F0 is high in absence of creaky voice (Figure 29), it is slightly falling, with a high falling pitch contour, in presence of creaky voice (Type 2,

Figure 27).

This suggests that, while glottal closure and creaky voice seem to be indissociable acoustic cues in the realization of the glottal stop at coda, their absence or presence may determine a change in pitch quality (high vs. high-falling contour). However, the correlation between absence or presence of glottal closure / creaky voice and difference in F0 remains to be explored quantitatively.

2.8.3. Glottal stop realization in pre-final sonorant position

Two types of realization are attested in pre-final sonorant position:
Type 1: Closure, creaky voice, high pitch
Type 2: High pitch

2.8.3.1.   Type 1: Closure, creaky voice, high pitch

Figure 30 presents the waveforms and spectrograms of a Type 1 realization of the glottal stop in pre-final sonorant position. The roots *dal* 'lentil' (a) and *daʔl* 'earthworm' (b) were pronounced in isolation by a male speaker (TAP-13, 20-30 yo). The glottal stop occurs before the sonorant. Its presence is visible through the glottal closure at the end of the nucleus. In addition to glottal closure, pitch is higher in presence of the glottal, with height of 185 Hz (mean) against 161 Hz (mean) in absence of glottal. A difference of 24 Hz in F0 on the nucleus contrasts both roots.

Figure 30. Waveform and spectrogram – Type 1 Glottal stop realization in pre-final sonorant position
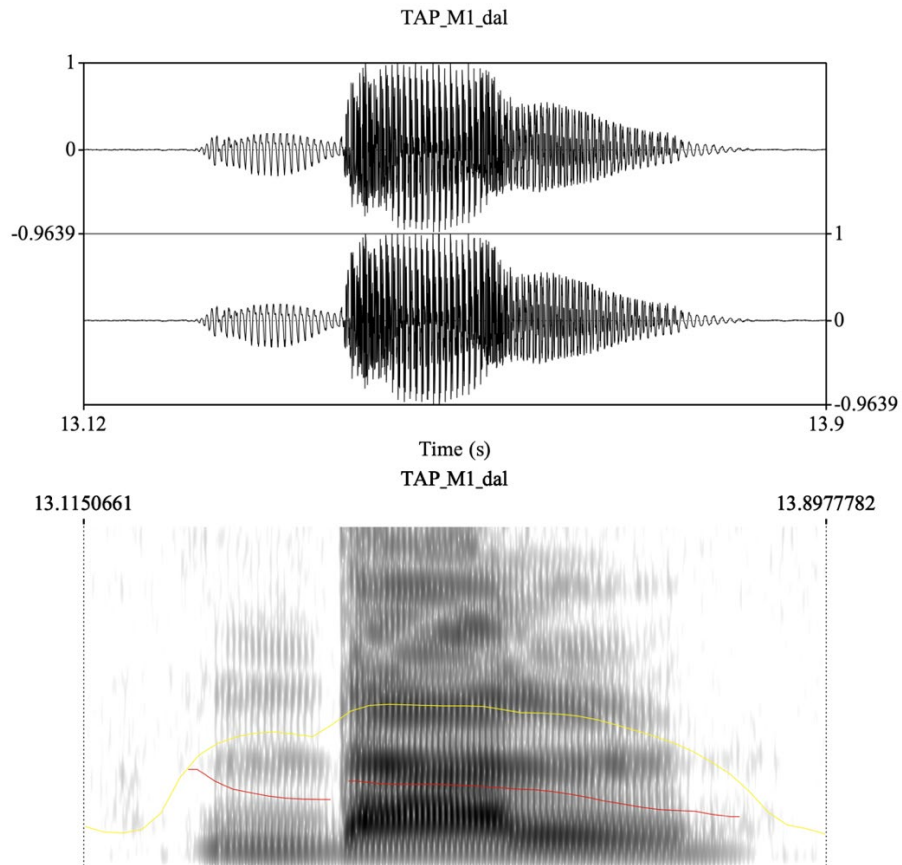
(a) *dal* 'lentil'

Figure 30, continued

(b) *daʔl* 'earthworm'

creaky voice   glottal closure
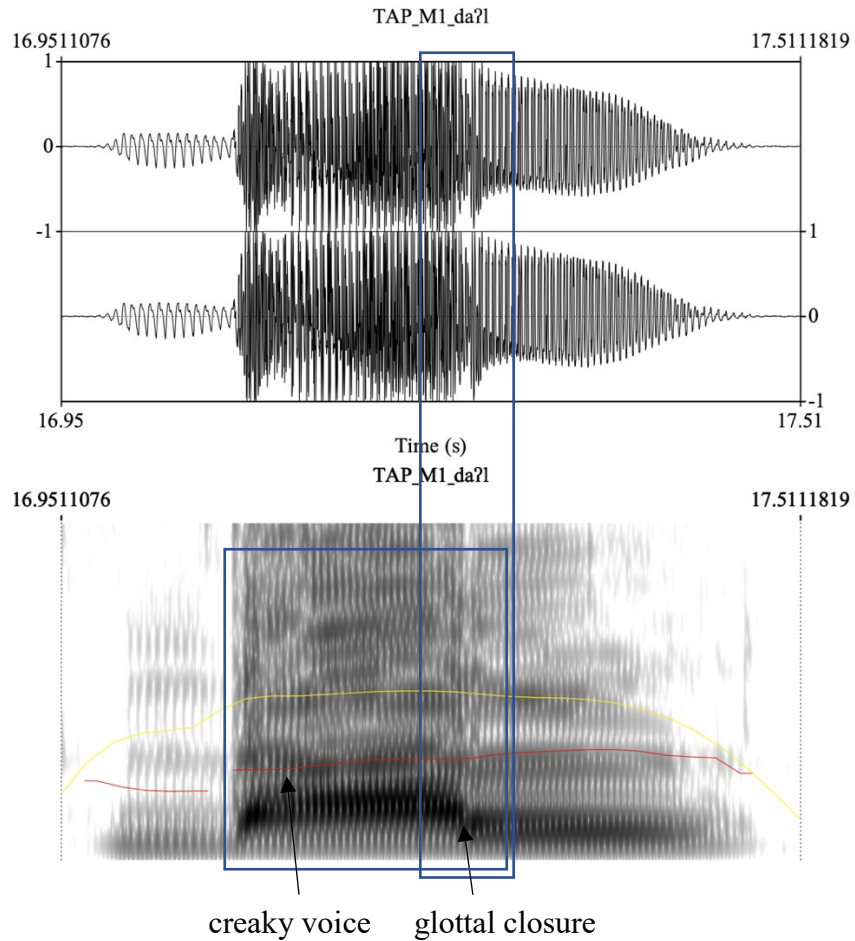
2.8.3.2.   Type 2: High pitch

Type 2 realization is illustrated in Figure 31 with the roots *dal* 'lentil' and *daʔl* 'earthworm' pronounced by a female speaker (RAP-13, 20-30 yo). The contrast resides in pitch height. The roots *dal* 'lentil' (234 Hz mean) and *daʔl* 'earthworm' (246 Hz mean) are contrasted with a F0 difference of 12 Hz measured over the nucleus. This difference can also be higher, as shown in (c) with *daʔl* 'earthworm.'

Figure 31. Waveform and spectrogram - Type 2 Glottal stop realization in pre-final sonorant position

(a) *dal* 'lentil'
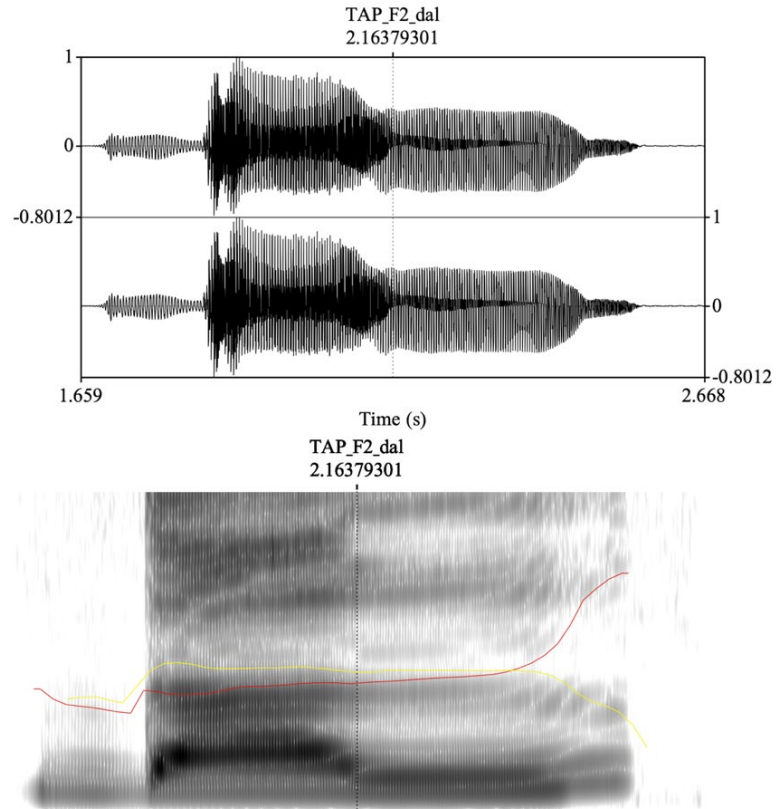
Figure 31, continued

(b) *daʔl* 'earthworm'

TAP_F2_daʔl
5.82449882
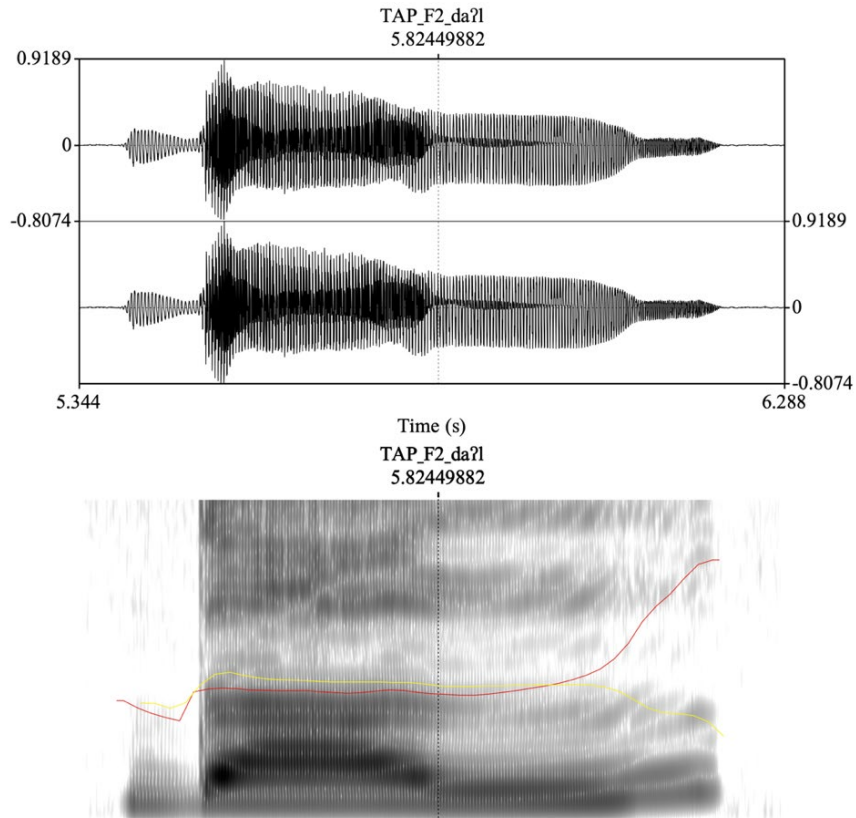
2.8.4. Glottal fricative realization in final position

The glottal fricative in final position is fully realized, and the pitch is raising and sharly falling, as one can see in

Figure 32. The attested pitch height is 271 Hz mean. The roots *lah* 'moon,' *laʔ* 'arrow' and *la* 'rope' were pronounced by the same female speaker (RAK-6 SIL, 34 yo) in the following carrier-phrase: *didi, ___ toalaŋ*. 'Elder sister, I said ____.' each phrase repeated three times.

Figure 32. Spectrogram - Glottal fricative realization in final position
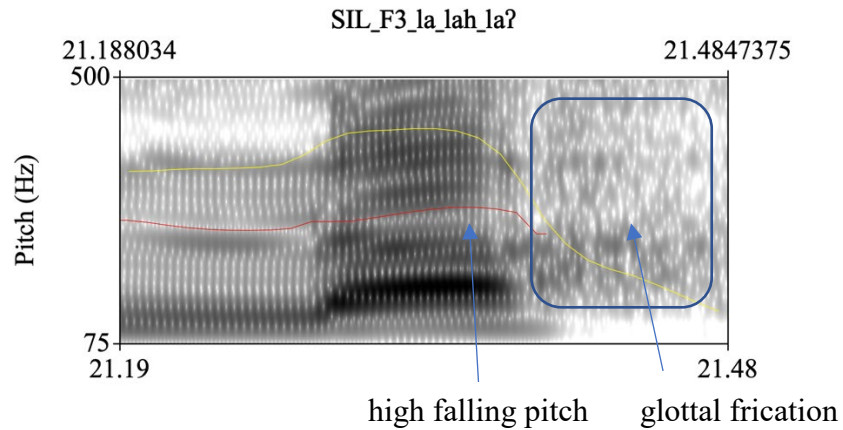
When comparing the realizations of (a) *la* 'rope,' (b) *laʔ* 'arrow,' and (c) *lah* 'moon,' their main difference reside in pitch height, pitch contour, and the presence of frication or breathiness with *lah* 'moon.' This is shown in Figure 33. The pitch attested for the root *la* 'rope' is 213 Hz mean, for *lah* 'moon' 271 Hz mean and for *laʔ* 'arrow' 274 Hz mean. While the F0 difference between *lah* 'moon' and *laʔ* 'arrow' is not great, a significant contrast is observed in pitch contour: a sharper pitch fall towards the end accompanies the realization of the glottal fricative; the raise of the pitch occurs earlier and less sharply with the glottal stop while that of the glottal fricative occurs later and more sharply. In addition, the presence of frication or breathy voice on the last part of the segment distinguish both consonants as well.

Figure 33. Spectrogram - Glottal fricative realization in final position – in comparison with final vowel and final glottal stop

(a) *la* 'rope'

(b) *laʔ* 'arrow'

raise start   slightly falling

(c) *lah* 'moon'

sharp raise start   sharply falling
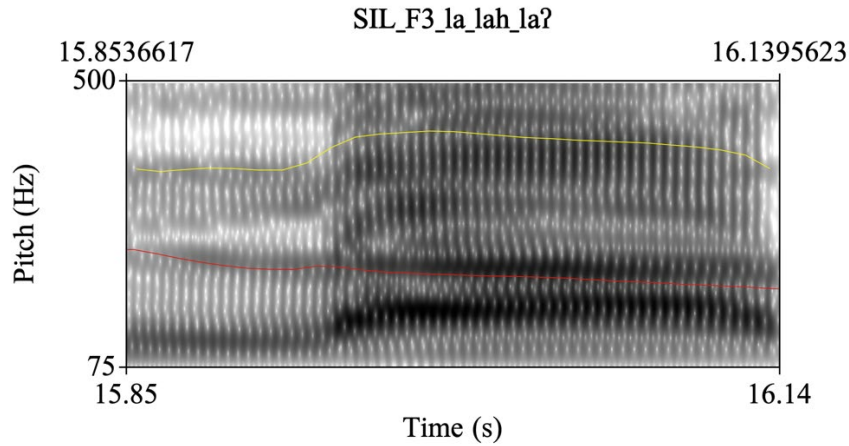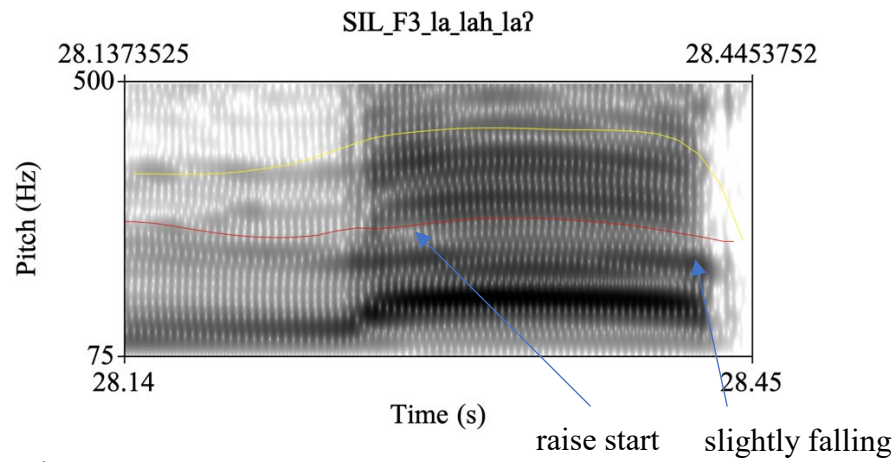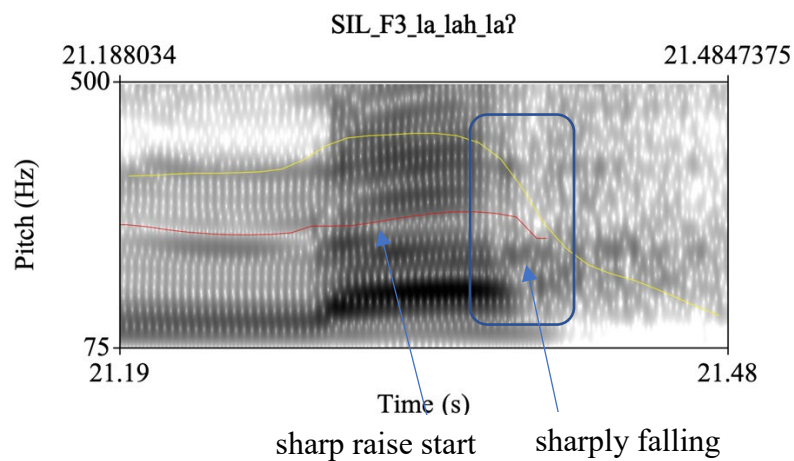
2.8.5. Glottal fricative realization in pre-final sonorant position

The glottal fricative in pre-final sonorant position is fully realized, and the pitch raises and sharly falls, as one can see in Figure 34. The attested pitch height is 264 Hz mean. The roots *tuŋ-* 'drink,' *tuʔŋ* 'foot (of tree)' and *tuhŋ-* 'be drained out' were pronounced by the same female speaker (RAK-6 SIL, 34 yo) in the following carrier-phrase: *didi, ___ toalaŋ.* 'Elder sister, I said ____.' all pronounced in second position of a three times repetition.

Figure 34. Spectrogram - Glottal fricative realization in pre-final sonorant position
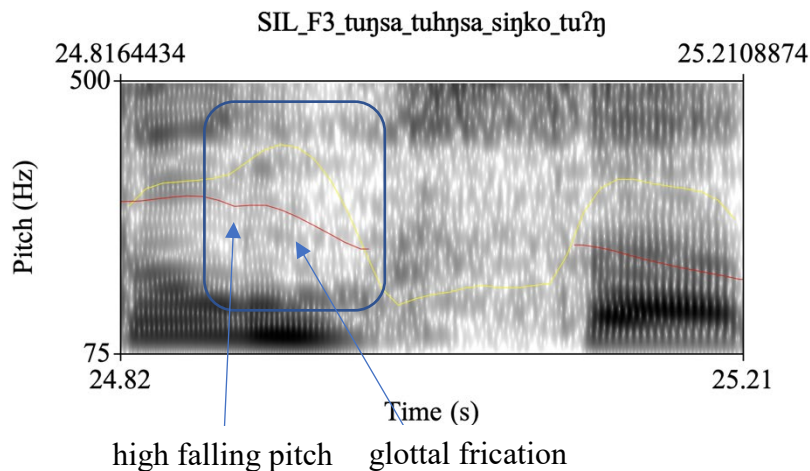
When comparing the realizations of (a) *tuŋ-* 'drink,' (b) *tuʔŋ* 'foot (of tree),' and (c) *tuhŋ-* 'be drained out,' their main differences reside in pitch height, pitch contour, and the presence of frication with *tuhŋ-* 'be drained out.' This is shown in Figure 35. The pitch attested for the root *tuŋ-* 'drink' is 218 Hz mean, for *tuhŋ-* 'be drained out' 264 Hz mean and for *tuʔŋ* 'foot (of tree)' 259 Hz mean. While the F0 difference between *tuhŋ* 'be drained out' and *tuʔŋ-* 'foot (of tree)' is not great, a significant contrast is observed in pitch contour: a sharper pitch fall towards the end accompanies the realization of the glottal fricative; the raise of the pitch occurs less sharply with the glottal stop than with that of the glottal fricative. In addition, the presence of frication or breathy voice on the last part of the segment distinguishes both consonants as well.

Figure 35. Spectrogram - Glottal fricative realization in pre-final sonorant position – in comparison with final sonorant and pre-final sonorant glottal stop

(a) *tuŋ-* 'drink'

SIL_F3_tuŋsa_tuhŋsa_siŋko_tuʔŋ

(b) *tuʔŋ* 'foot (of tree)'

SIL_F3_tuŋsa_tuhŋsa_siŋko_tuʔŋ

smooth raise start

slightly falling

(c) *tuhŋ-* 'be drained out'

SIL_F3_tuŋsa_tuhŋsa_siŋko_tuʔŋ

sharp raise start   sharply falling
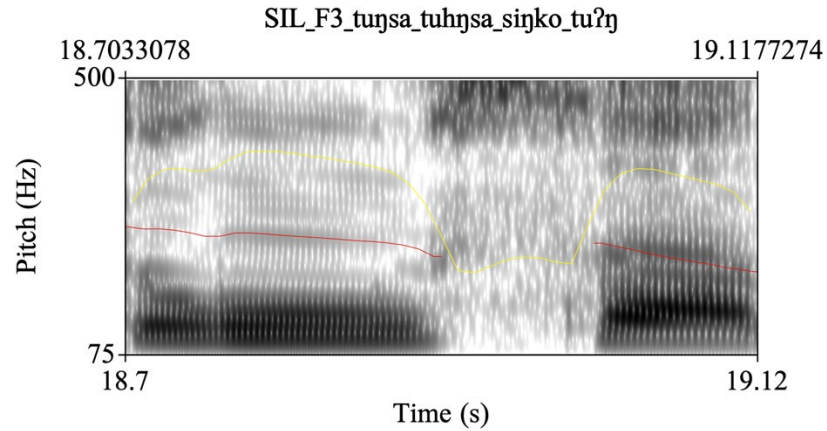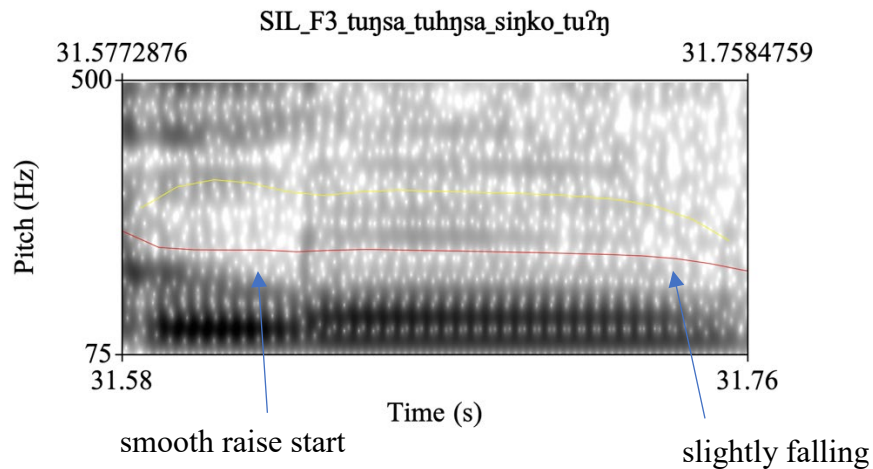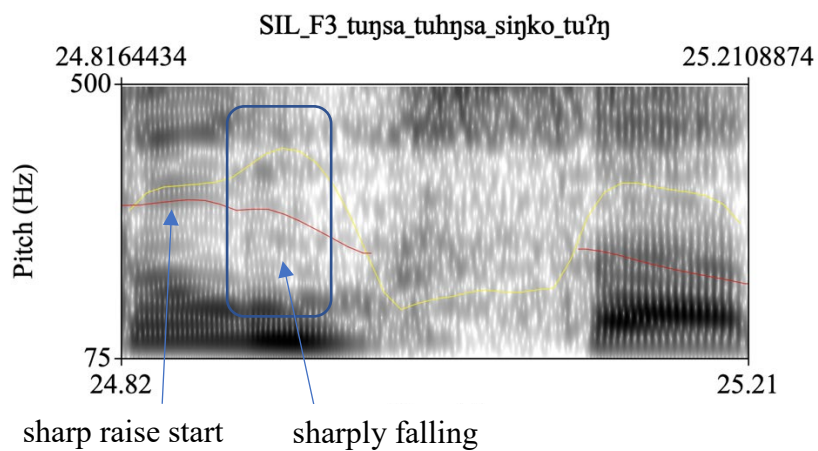
2.8.6. Loss of laryngeal contrasts

Some speakers (MAN-4) show a complete loss of glottal stop in pre-final sonorant position. Amongst these speakers, the status of the glottal stop in final position remains unclear and needs further study, since its presence can vary within a single individual.

Figure 36 illustrates the loss of the glottal stop in pre-final sonorant position with the root *daʔl* 'earthworm' followed by the root *sjaʔn* 'insect, bug.' It is pronounced in isolation by a female speaker (BAN, 50-60 yo). As we can see, no trace of glottal stop closure or difference in pitch is attested.

Figure 36. Waveform and spectrogram – Loss of glottal stop in pre-final sonorant position

daʔl sjaʔn 'earthworm'
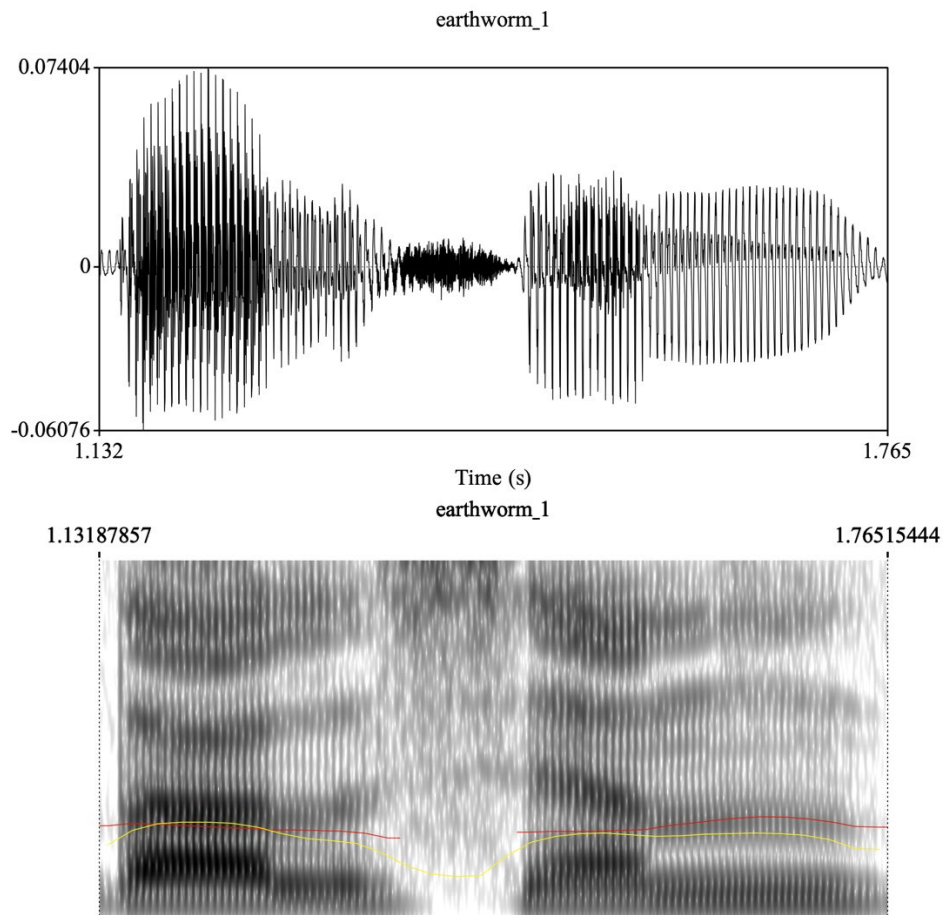
2.9.   Tonogenesis in Chepang

Tonogenesis is the process by which an atonal language historically develops tonal contrasts. In Chepang, tonogenesis is an on-going process whose analysis can shed light on how tonal contrasts emerge in language. In this section, I describe how the emergence of tonal contrasts in Chepang has taken place and the challenges that this innovative system brings to its description. Tonogenesis in Chepang finds its source in the realization of laryngeal features in final and pre-final sonorant positions. The development of tones in Chepang primarily correlates with the gradual loss of laryngeal contrasts on pre-final and final consonants.

Glottal stop realizations, which include a high and mid-raising pitch (Type 2 and 3), have led to the development of a two-way tonal contrast (§ 2.9.1). While the realization of the glottal fricative remains segmental in all studied varieties, it also correlates with raising-falling pitch contour. Both glottal consonants have developed a similar pitch height but distinct pitch contours in their realizations and voice quality. This raises the question of whether the Chepang tonal system should be considered to feature three tones rather than two, since one of them combines with the presence of a glottal fricative segment in its realization, i.e., breathiness (§ 2.9.2). Beyond the laryngeal features of final and pre-final sonorant consonants /ʔ/ and /h/, I briefly investigate the pitch contours observed with other types of phonation, specifically, voicing and aspiration or breathiness of initial consonants (§ 2.9.3).

Finally, I discuss the Chepang innovative tonal system with regard to other cases of tonogenesis in TH languages and TH tone typology (§ 2.9.4). For the present analysis, I used a Praat (Boersma & Weenink 2021) script (Reetz 2021) to extract Fundamental Frequency (F0) measurements (Hz). The measurements were taken at 10 equidistant points (0% to 100%) over the length of syllabic segments. F0 values over time were plotted in Excel to extract F0 contours.

2.9.1. Two-way tonal contrast

The realizations of the glottal stop in final (Type 2 and 3) and pre-final sonorant (Type 1 and 2) positions, which involves a high or mid-raising pitch contour, are the source of the emergence of a two-way tonal contrast, such as:

- Tone 1: a level or mid tone with words that do not feature any glottal stop (¯)

- Tone 2: a high or mid-raising (mid to high) tone with words which originally featured a glottal stop (˘)

This two-way tonal contrast is illustrated in Figure 37 and Figure 38 with the pitch contours of two minimal pairs featuring a low central vowel /a/: *la* 'rope,' *laʔ* 'arrow,' and *dal* 'lentil,' *daʔl* 'earthworm.'

These words were pronounced in isolation and in the following carrier-phrase: *didi, ___ toalaŋ.* 'Elder sister, I said ___.'

The pitch contours in Figure 37 and Figure 38 account for the F0 mean values of the words repeated three times in isolation and in the carrier-phrase by 3 female speakers of the Lothar varieties (GUN, TAP, SIL), i.e., a total of 72 tokens, that is 18 tokens per word (Figure 37), or 36 tokens per tone (Figure 38).

Figure 37. Two-way tonal contrast: mean pitch contours of two minimal pairs
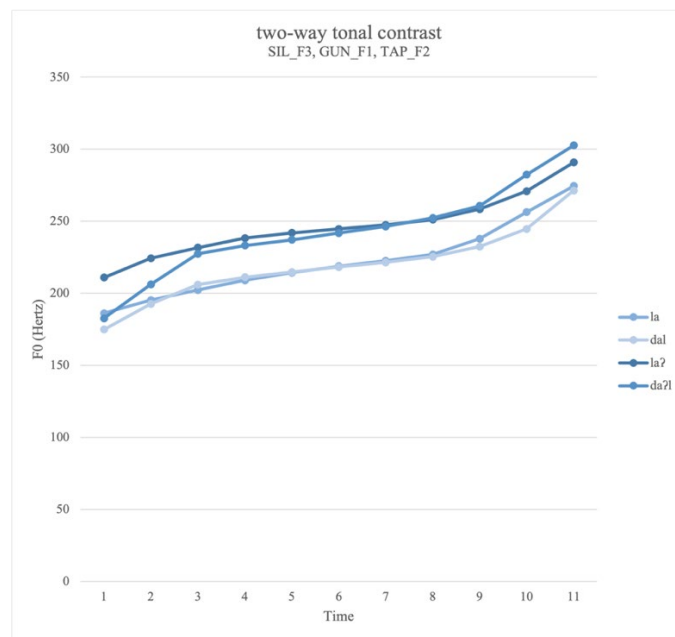
Figure 38. (A) Two-way tonal contrast: mean pitch contours of Tone 1 and Tone 2
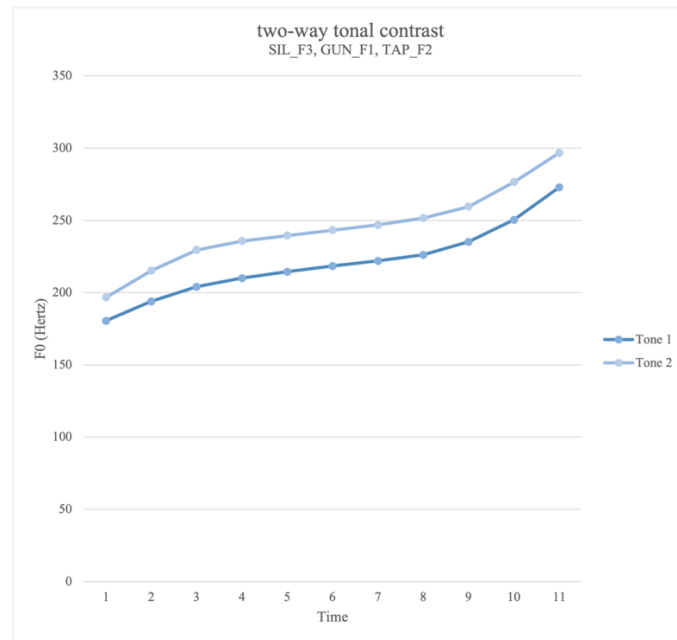
In Figure 37 and Figure 38, one can see that Tone 2 is higher than Tone 1, but that Tone 1 is raising as much as Tone 2. This is an effect of intonation that can be observed when speakers are repeating words or clauses in this type of elicitation session. The same minimal pairs were recorded in different carrier-phrases with another female speaker of the Lothar varieties (RAK-6). The words *la* 'rope' and *laʔ* 'arrow' were repeated four times in the following carrier-phrase: *ipaj ___ kʰelʌ.* 'This is not ___.' The words *dal* 'lentil' and *daʔl* 'earthworm' were also repeated four times in two different carrier phrases: *ŋai dal ʥelaŋ* 'I eat lentil.' and *ŋai daʔl ʥeŋʌlʌ* 'I don't eat earthworm.'

The pitch contours of Tone 1 and Tone 2 in Figure 39 account for the F0 mean values of these 16 tokens, that is, 8 tokens per tone. The speaker did not repeat the carrier-phrases with the intonation characteristic to repetition. Rather, the speaker "reset" between each clause, as if they were individual units and not part of a set. In a more natural setting, with a lower influence of intonation, Tone 1 is level. It contrasts with Tone 2 which has the mid-raising contour shape commonly observed.

Figure 39. (B) Two-way tonal contrast: mean pitch contours of Tone 1 and Tone 2
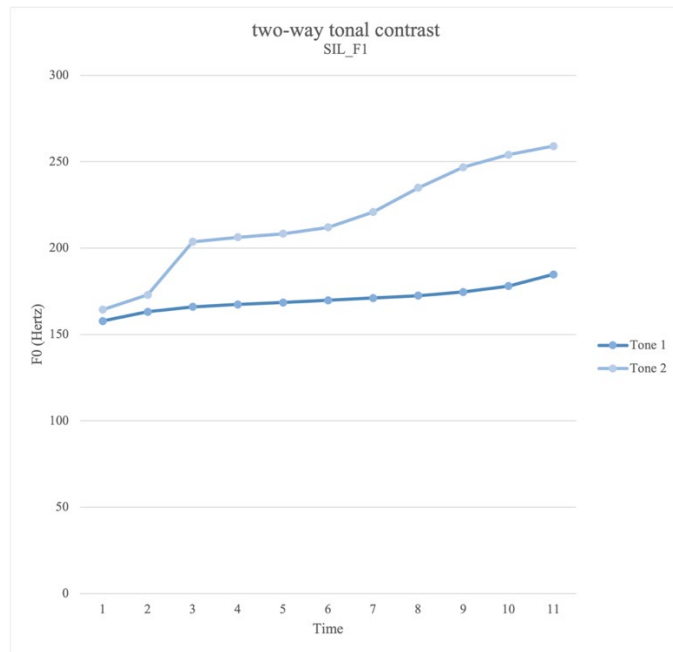
The tone bearing unit is the word. In disyllabic words, both first and second syllables show the same F0 differences. This is illustrated in Figure 40 with the words *alsa* 'go' and *aʔlsa* 'take away,' where the F0 contour of the first syllable is preserved on the second syllable. Three tokens of each word were pronounced in isolation by a female speaker of the Lothar varieties (RAP-13).

Figure 40. Tone bearing unit: word
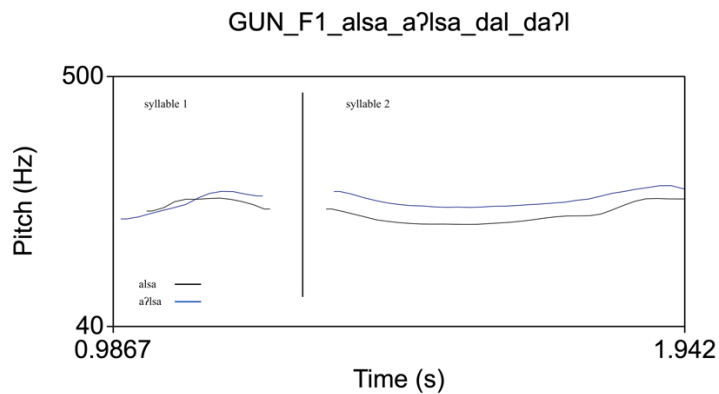
If one considers that Chepang has developed a two-way tonal contrast, this system implies that the realization of Tone 2 either solely relies on its high or mid-raising F0 contour (Type 3 glottal realization) or correlates with phonation, specifically creaky voice, and the presence of a glottal stop constriction (Type 2 glottal realization).

In Table 78, I summarize the two-way tonal contrast that developed from the realization of the phonemic laryngeal features of the glottal stop /ʔ/.

Table 77. Chepang tonogenesis: a two-way tonal contrast

Two-way tonal contrast	Tone 1	Tone 2
syllable structure	CV(C)	CVʔ(S)
F0 contour	level / mid	raising (mid to high)
phonation	modal	+/- creaky
phonemic segment		+/- glottal stop constriction
tone notation	ˉ	ˊ
pitch schematic representation		

2.9.2. Three-way tonal contrast

The realization of the glottal fricative in final and pre-final sonorant position also correlates with a difference in F0. The pitch raises as well with the glottal fricative but, by contrast with the glottal stop, the contour falls more sharply. By contrast with the

glottal stop, the realization of the glottal fricative combines both the presence of a glottal fricative segment and a high or mid-raising pitch contour.

If the F0 contour that correlates with the presence of the glottal fricative is part of the tonal system, it can be considered a three-way tonal contrast, as follows:

- Tone 1: a level or mid tone with words that do not feature any glottal stop (ˉ)

- Tone 2: a high or mid-raising (mid to high) tone with words which originally featured a glottal stop (˘)

- Tone 3: a raising-falling tone in presence of a glottal fricative segment (ˆ)

Under the analysis that the tonal system of Chepang features three tones means that Tone 2 and 3 are produced with the same F0 contour in terms of height while their contour and phonation are different: while Tone 2 can correlate with creaky voice, Tone 3 combines F0 raising-falling contour and the presence of the glottal fricative, i.e., breathiness. This three-way tonal contrast is illustrated in Figure 41, for the following triplet: *la* 'rope,' *laʔ* 'arrow,' *lah* 'moon.' These words were pronounced in isolation and in the following carrier-phrase: *didi, ____ toalaŋ.* 'Elder sister, I said ___.' The pitch contours account for the F0 mean values of the words repeated three times in isolation and in the carrier-phrase by three female speakers of the Lothar varieties (RAK-6, RAP-13), i.e., a total of 54 tokens, that is, 18 tokens per tone.

Figure 41. Three-way tonal contrast: mean pitch contours of a triplet
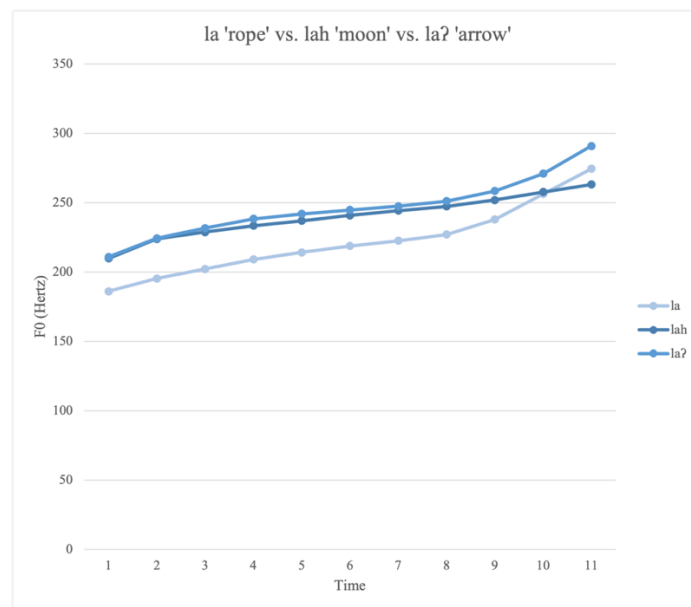
In order to have a better picture of the pitch contours of Tone 1, 2, and 3, since vowels have their own intrinsic F0, I have added to the analysis of the triplet in Figure 41, the F0 values of two other triplets that feature different vowels: a high front vowel /i/ and a close-mid vowel /e/. These two additional triplets are: *le* 'tongue,' *leh* 'spleen,' *leʔ-* 'take, buy,' and *li-* 'Chebulic myrobalan,' *lih-* 'worry,' *liʔ-* 'be heavy.' The words of these three triplets were pronounced in isolation and in the same carrier-phrase by three female speakers of the Lothar varieties (GUN, TAP, SIL), i.e., a total of 172 tokens, that is, 54 tokens per tone. The pitch contours of Tone 1, 2, and 3 are presented in Figure 42.

Figure 42. Three-way tonal contrast: mean pitch contours of Tone 1, Tone 2, and Tone 3
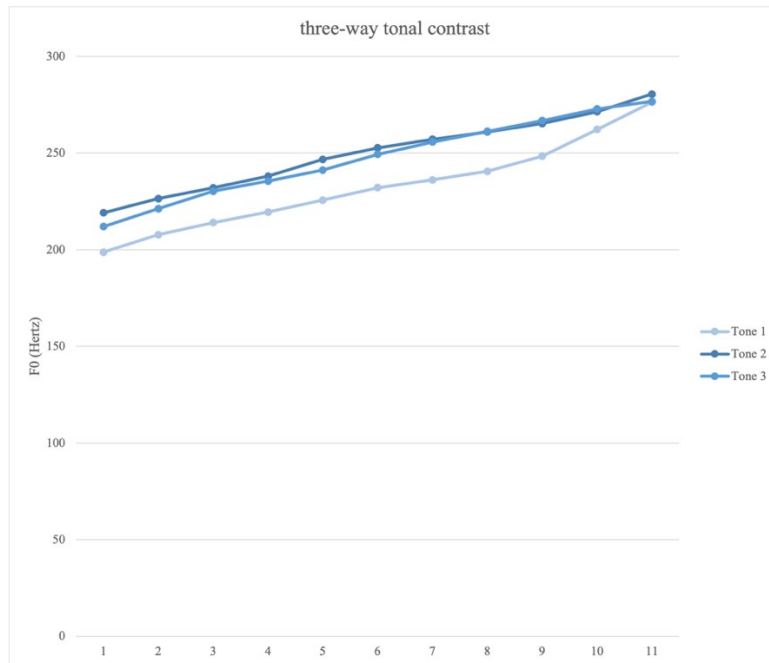
The question of the number of tones in Chepang remains open. While the F0 contour of Tone 2 can correlate with the presence of a glottal stop constriction and/or creaky voice, it can also be the sole acoustic cue speakers rely on to understand the word. By contrast, Tone 3 combines with the presence of the glottal fricative segment This difference in the degree of independence of the suprasegmental features questions the soundness of considering that a three-way tonal contrast exists phonologically. In Table 78, I summarize the three-way tonal contrast that developed from the phonemic laryngeal features of the glottal stop /ʔ/ and the glottal fricative /h/.

203

Table 78. Chepang tonogenesis: towards a three-way tonal contrast

Two-way tonal contrast	Tone 1	Tone 2	Tone 2
syllable structure	CV(C)	CVʔ(S)	CVh(S)
F0 contour	level / mid	raising (mid to high)	raising-falling
phonation	modal	+/- creaky	+/- breathy
phonemic segment		+/- glottal stop constriction	+ glottal fricative constriction
tone notation	ˉ	ˊ	ˆ
pitch schematic representation			

2.9.3. Other phonation types and F0 contour

There is no noticeable pitch difference between words that contrast through the voicing of the initial consonant. This is illustrated in Figure 43 with the words *paŋ* 'husband's younger brother' and *baŋ* 'stone.' Both words were pronounced by the same female speaker in isolation (left) and in the following carrier phrase (right): *didi, ___ toalaŋ.* 'Elder sister, I said ___.' In isolation, the pitch changes because of the differences in intonation that apply to the different repetition (R1, R2, R3).

Figure 43. Voiced final consonants: no pitch difference
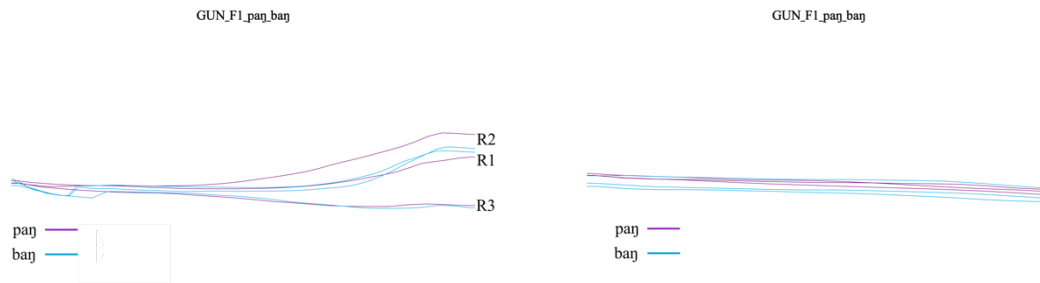
Words starting with an aspirated or breathy consonant show additional pitch differences that contrast with words that start with unaspirated or non-breathy consonants. Words with a voiceless aspirated or voiced breathy initial consonant have a higher F0 contour, respectively falling and raising. This difference in F0 is also present on the second syllable where the F0 difference remains through a similar high flat tone. This is illustrated in Figure 44 with the words *pe-sa* 'be nice' and *be-sa* 'be thin (things)' which contrast with their aspirated and breathy counterparts *pʰe-sa* 'leave behind' and *bʰe-sa* 'separate within the family.' Both words were pronounced by the same female speaker in isolation (left) and in the following carrier phrase (right): *didi, ___ toalaŋ.* 'Elder sister, I said ___.' In isolation, the pitch changes because of the differences in intonation that apply to the different repetition (R1, R2, R3). Differences in F0 have yet not replaced the segmental differences between the consonants. The differences in F0 associated with the presence of breathiness and aspiration could eventually participate in the development of a four-way tonal contrast.

Figure 44. Voiced initial consonants: noticeable pitch difference
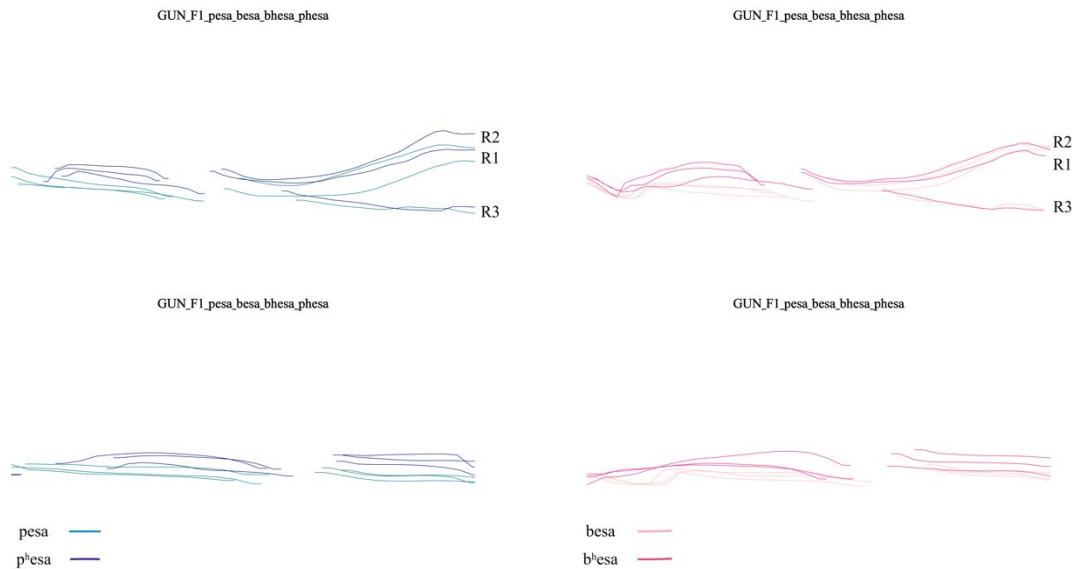
2.9.4. Tonogenesis and TH tone typology

The loss of final laryngeal consonants giving birth to tonal contrasts is a diachronic development attested in languages of East Asia (Haudricourt 1954; Matisoff 1973; Thurgood 2007; Michaud & Sands 2020). A final glottal stop can give rise to either a high/raising or low/falling tone. The final glottal stop in Old Chinese gave rise to a high or raising tone in Middle Chinese (Haudricourt 1954). This is also the case for Vietnamese (Haudricourt 1954).

In Chepang, we have seen that the glottal stop has developed a high or mid-raising pitch contour while the glottal fricative a raising-falling pitch contour. However, both realizations show differences in voice quality (creaky vs. breathy). As shown in § 2.8.2, a better understanding of the correlation between glottal closure / creaky voice and pitch quality (High to High-Falling contour) would help determine the typological tonal profile towards which Chepang is possibly evolving.

In languages of Southeast Asia, tonal contrast is often typologically characterized by the presence of a mere contrast in F0, or an amalgam of both F0 and voice quality or register (Hildebrandt 2007; Kirby & Brunelle 2017). In Nepal, Manange and Tamang (Bodic) exhibit both pitch and breathiness as acoustic correlates of a four way tonal

contrast (Hildebrandt 2007; Mazaudon & Michaud 2008). Another example is Takale Kham, a Himalayan language (like Chepang), which relatively recently developed a four way tonal contrast (Watters 2002; Wilde 2017). This tonal contrast distinguishes two tones, each of which shows modal and lax voice register differences (Watters 2002: 36–39). Wilde (2017) compares Takale Kham with three other Kham varieties (Sheram, Ghusbang and Gamale) and finds that Takale Kham developed its mid-falling tone - lax voice register (*T-2 Lax* in Watters (2002)) from the loss of final glottal stop in roots with breathy nuclei which also presents vowel lengthening. In this case, Wilde (2017) shows that the final glottal stop in Takale Kham comes from final alveolar and velar stops in Proto-Kham.

Whether some varieties of Chepang qualify as tonal remains an open question that requires an in-depth study of the variation in glottal realizations across more speakers from different gender, age, and places.

2.10. Word structure and stress

In this section, I describe the structures of roots and stems that form words (§ 2.10.1) and the placement of stress (§ 2.10.2). Stress is a supra-segmental feature characterized by the prominence of certain syllable(s) in a word. Roots cannot be further analyzed into constituent morphemes. In basic terms, a single root, or a combination of roots or root(s) and derivational morphemes (inflected for number or not in the case of a nominal root) can form a stem (§ 3.1). A stem can be inflected with case markers for instance when nominal (§ 3.4.5) or with verbal morphology when verbal (§ 5.7).

While I do not intend to define what a word is in a developed manner, a word in its minimal form could be a single root or non-inflected stem when nominal (equivalent of a single morpheme), or an inflected stem when verbal (since verbal roots are bound, i.e., they do not occur without verbal morphology), that has a phonological and syllabic structure, that minimally bears one stress and that conveys a meaning that is either devoid of contextual ground when pronounced in isolation, or that has a contextual significance in a process that happens, happened, or would happen in the world when pronounced in a carrier phrase or connected speech in semi-natural or natural discourse. In other words, a

word has a phonological shape, a syllabic structure, bears stress, is formed by a single or several morphemes, and its meaning is significant in an expression translated by the speaker to the addressee who would understand it.

This description mainly focuses on nominal words, but I discuss some examples of inflected verbal stems as well. I will show that word-level stress when a word is pronounced in isolation is consistent with the stress placement found when the word is pronounced in a carrier phrase or in a declarative sentence that does not carry any specific pragmatic intonation that would cover stress pattern.

2.10.1. Root and stem, or word structure

Most of Chepang native lexical roots (nominal or verbal) are monosyllabic. Dissyllabic or trisyllabic nominal native roots are rare, and trisyllabic verbal native roots do not exist. Dissyllabic verbal roots constitute a small set that features a second syllable whose rhyme ends in a back open-mid vowel /ʌ/ (§ 5.3). Dissyllabic and trisyllabic native nominal stems (non-inflected for number) are often the result of complex nominal formation, such as compounding (§ 3.3) Dissyllabic and trisyllabic verbal stems are formed through derivational processes (§ 5.6). Finally, dissyllabic and trisyllabic structures are typical of non-native stems borrowed from Nepali, or English via Nepali in the case of nouns since English verbs have not yet made their way to the Chepang lexicon.

In Table 79, I show some examples of monosyllabic, dissyllabic, and trisyllabic verbal and nominal roots, and in Table 80 dissyllabic and trisyllabic non-inflected nominal stems.

Table 79. Mono-, di-, tri- syllabic roots or simple stems

monosyllabic roots/stems		dissyllabic roots/stems		trisyllabic roots/stems	
ahm	'porridge'	*asaŋ*	'first floor (house)'	*amdani*	'income' (<N.)
baŋ	'stone'	*beka~beŋka*	'fisherman'	*basuɾi*	'flute (transverse)' (<N.)
daʔl	'earthworm'	*dwila*	'dragonfly'	*babaɾi*	'wild basil' (<N.)
gwej~goj	'Cocoyam, taro'	*gjawliŋ~gjawlaŋ*	'spider'	*gʰoŋkoɾok*	'Adam's apple'
hʌw	'younger brother'	*guluŋ~newliŋ*	'mongoose'	*hamali*	'tomato'
keʔ-	'fish'	*holjoŋ-*	'be loose'	*kʰʌɾbudʑa*	'watermelon' (<N.)
kim	'house'	*husu*	'fog, haze'	*lilami*	'secondhand clothes'
kwi	'dog'	*hʌŋsʌ*	'soul, spirit'	*mesagʌ*	'bamboo shoot'
loʔ	'leaf'	*kʌdʌw*	'millet' (<N.)	*mʌndʌli*	'assembly' (<N.)
njam	'sun'	*manta*	'person, people'	*nʌɾiwal*	'coconut' (<N.)
njum-	'be good, tasty'	*meɾu (arch.)*	'wife'	*paŋuli*	'firefly'
pɾat-	'faint'	*ʌm-tak-*	'heal-CAUS'	*pokʰʌɾi*	'pond' (<N.)
ɾap-	'love, like'	*mʰotoŋ*	'mouth'	*poɾʌli*	'sponge gourd'
ɾo	'flower'	*nelaw*	'nettle'	*puɾano*	'old' (<N.)
suk-	'plant'	*nukʌ-*	'hide self'	*ɾedijo*	'radio' (<N. <E.)
te-	'beg'	*pʌtɕʌw-*	'digest' (<N.)	*suntala*	'mandarin' (<N.)
ti	'water'	*sjʌmbʌɾ~simbʌɾ*	'ginger'	*tagaɾaw*	'nickerbean'
wa	'bird'	*tebul*	'table' (<N. <E.)	*tibiliŋ*	'butterfly'

Table 80. Mono-, di-, tri- syllabic complex nominal stems

dissyllabic complex nominal stems		trisyllabic complex nominal stems	
glaw-si	'red silk cotton tree'	*to-bɾok-sjaʔn*	'caterpillar *sp.*'
la-ge	'fishing rod'	*daŋ-saɾ-wa*	'heron, common hoopoe'
hi-diʔŋ	'day, afternoon'	*siŋ-doɾ-wa*	'woodpecker'
hʌw-djaŋ	'younger sister'	*liŋ~niŋ-saɾ-wa*	'bulbul (red-vented), nightingale'
ɾuiŋ-sja	'Hog deer, Sambar'	*umpuɾ-saj*	'timur'
wiʔ-la	'weasel, ferret, civet (masked)'	*metaŋ-saj*	'Himalayan blue pine tree fruit'

2.10.2. Stress realization and general observations

Monosyllabic, dissyllabic or trisyllabic roots or stems can form a word that in all cases, except for determinative or descriptive compound type 2, bears stress on the first syllable. Stress consistently applies at the word level, whether the word is pronounced in

isolation, in a carrier phrase, or in a phrase in connected speech or natural discourse in absence of specific pragmatic intonation.

In Chepang, the main acoustic cue for stress is pitch (F0). While intensity (or amplitude) often correlates with stress, it remains marginal, pitch being the most prominent feature. This is illustrated in Figure 45 with the dissyllabic root, stem or word /nemet/ [ˈne.met] 'ant' pronounced in isolation. Pitch (plain line) is higher on the first syllable and lower on the second, while intensity (spotted line) is high on the first syllable and decreasing towards the end of the second syllable.

In the following sub-sections, I describe the stress patterns observed with dissyllabic (§ 2.10.3) and trisyllabic (§ 2.10.4) words. The present analysis is not complete and only consists in an overview of the observed patterns. Further research should include the analysis of stress across more Chepang language varieties.

Figure 45. Stress, pitch (F0) vs. intensity in the word /nemet/ 'ant'
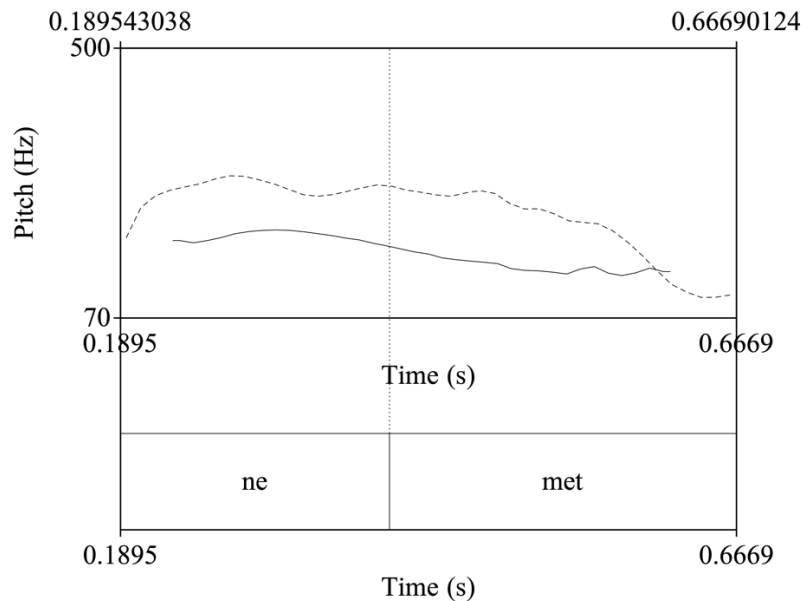
2.10.3. Stress on dissyllabic nominal root, stem, or word

Two types of stress patterns are observed with dissyllabic words: trochaic stress, described in § 2.10.3.1 and § 2.10.3.3, and spondaic stress, described in § 2.10.3.2. Some possible effects of elicited pronunciation on stress are discussed in § 2.10.3.4.

2.10.3.1. Trochaic stress on dissyllabic roots, determinative or descriptive compounds Type 1, synthetic compounds, and monosyllabic roots suffixed with the non-productive derivational morpheme -*si*

In dissyllabic words formed with a single root or non-inflected stem, stress is trochaic: the first syllable is stressed and the second syllable unstressed. The same trochaic stress pattern is observed on dissyllabic words formed with descriptive or determinative compounds (§ 3.3.1.1), regardless of the types of roots that occur in such compounds, synthetic compounds (§ 3.3.1.1.5), or with roots suffixed with the non-productive derivational morpheme -*si* which creates nominal stems that express species of trees (§ 3.3.4.1).

The dissyllabic root /tokɾak/ [ˈto.kɾak] 'toad' pronounced in isolation is illustrated in Figure 46, and in the carrier phrase /didi, _____ toalaŋ/ 'Elder sister, I said _____' in Figure 47. In both cases we observe a trochee (¯ �‿), i.e., first syllable stressed, second syllable unstressed. The examples of trochaic stress on the dissyllabic roots /nemet/ [ˈne.met] 'ant' in Figure 45 and /tokɾak/ [ˈto.kɾak] 'toad' in Figure 46 suggest that syllable weight does not matter in Chepang. In both roots, the first syllable is light (CV), and the second syllable is heavy (CVC and CCVC). The syllable can thus be light and stressed, and if heavy, does not attract stress.

Figure 46. Trochaic stress in the dissyllabic word /tokɾak/ 'toad' in isolation

GUN_F1_tik_ŋol_teksa_tʌk_taksa_tokɾak_tuk
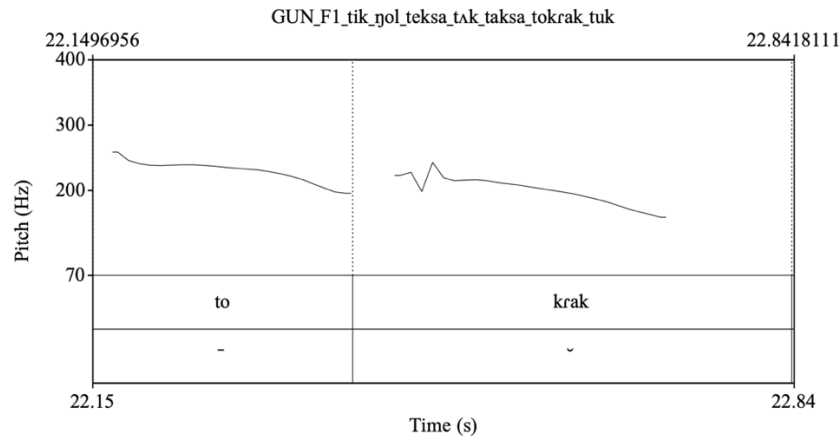
Figure 47. Trochaic stress in the dissyllabic word /tokɾak/ 'toad' in carrier phrase
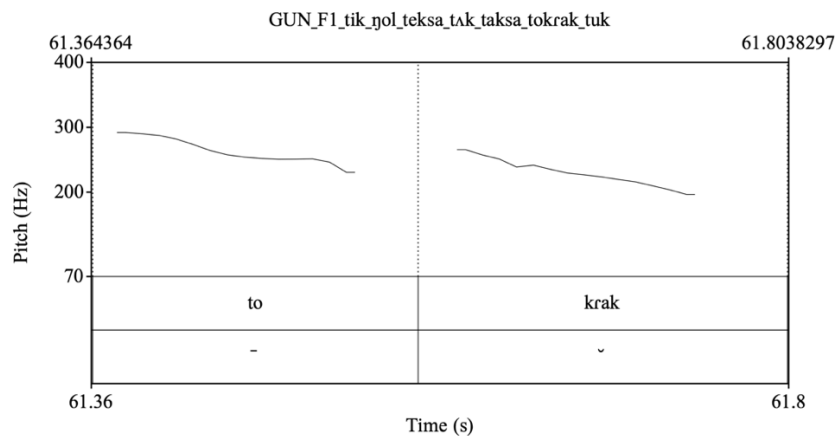
GUN_F1_tik_ŋol_teksa_tʌk_taksa_tokɾak_tuk

Trochaic stress in descriptive or determinative compounds formed with two nominal roots is illustrated with the dissyllabic word /li-saj/ [ˈli.saj] 'seed or fruit of the chebulic myrobalan tree' in Figure 48, uttered in the carrier phrase /didi, ____ toalaŋ/ 'Elder sister, I said ____', /mak-saj/ [ˈmak.saj] 'seed or fruit of the Bauhinia vahlii tree' in Figure 49, uttered in a sentence from a conversation, and with /go-sja/ [ˈgo.sja] 'wild jungle cat' in Figure 50, uttered in a sentence from a conversation. Trochaic stress in descriptive or determinative compounds formed with a verbal root preceding the nominal root head of the compound is illustrated with the dissyllabic word /dah-njam/ [ˈdáh.njam]

'sunrise,' literally 'reaching/arriving sun,' in Figure 51, uttered in a sentence from an expository text.

Figure 48. Trochaic stress in the dissyllabic word /li-saj/ 'fruit of the chebulic myrobalan tree' in a career phrase
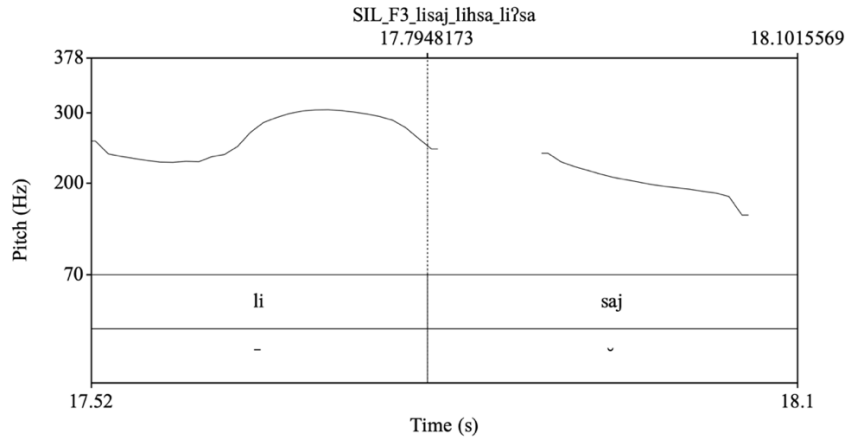
Figure 49. Trochaic stress in the dissyllabic word /mak-saj/ 'seed or fruit of the Bauhinia vahlii tree' in a sentence from a conversation
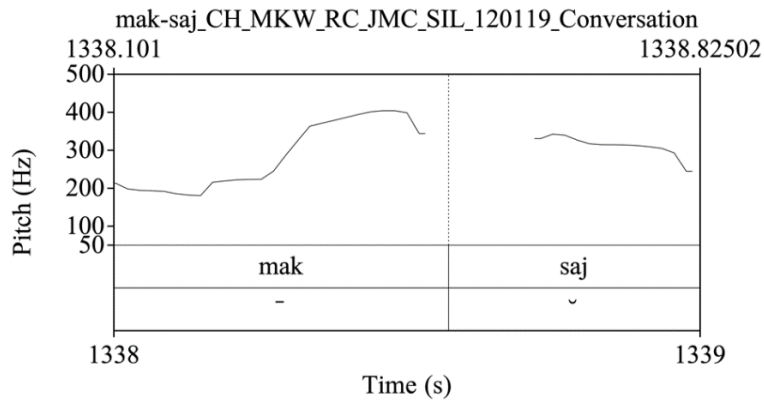
Figure 50. Trochaic stress in the dissyllabic word /go-sja/ 'wild jungle cat' in a sentence from a conversation

416.752254   417.08033

```
500
400
300
200

 70
        |      go      |      sja      |
        |      –       |      ˘        |
```

Pitch (Hz)

416.8   417.1

Time (s)
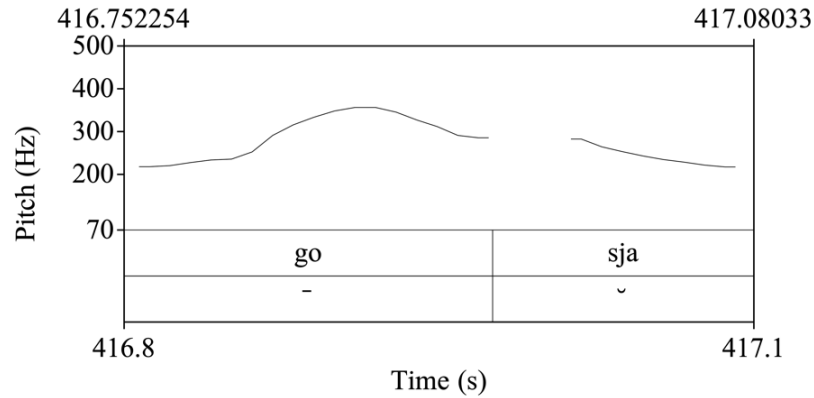
Figure 51. Trochaic stress in the dissyllabic word /dah-njam/ 'sunrise' in a sentence from an expository text

dah-njam_CH_MKW_DBC_MAI_1_020220_Local_History

0.6856494   1.0436072

```
500
400
300
200
100
 50
        |      dah     |      njam     |
        |      –       |      ˘        |
```

Pitch (Hz)

0.6856   1.044

Time (s)
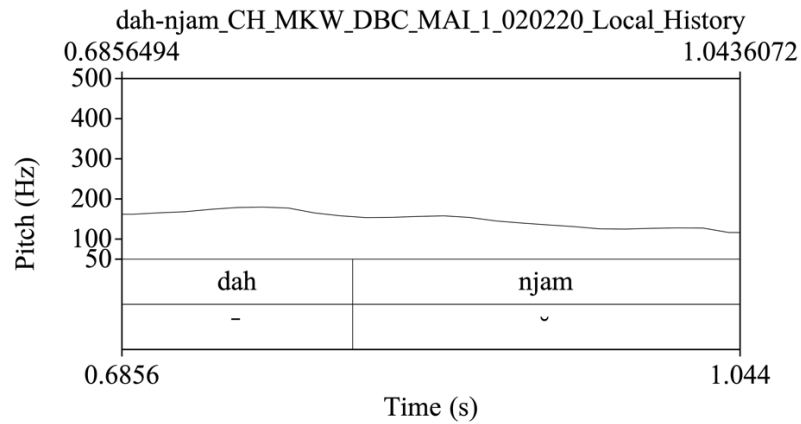
Synthetic compounds (§ 3.3.1.3) feature trochaic stress too, as illustrated with /juin-ɾaj/ [ˈjuin.ɾaj] in Figure 52, uttered in a sentence from a conversation, and /mʰeʔ-ku/ [mʰeʔ.ku] 'fire smoke' in Figure 53, uttered in the carrier phrase /____ muna/ 'There is ____.'

214

Figure 52. Trochaic stress in the dissyllabic word /juin-ɾaj/ 'story, folktale' in a sentence from a conversation

juin-ɾaj_CH_MKW_SC_SIL_010220_3_Life
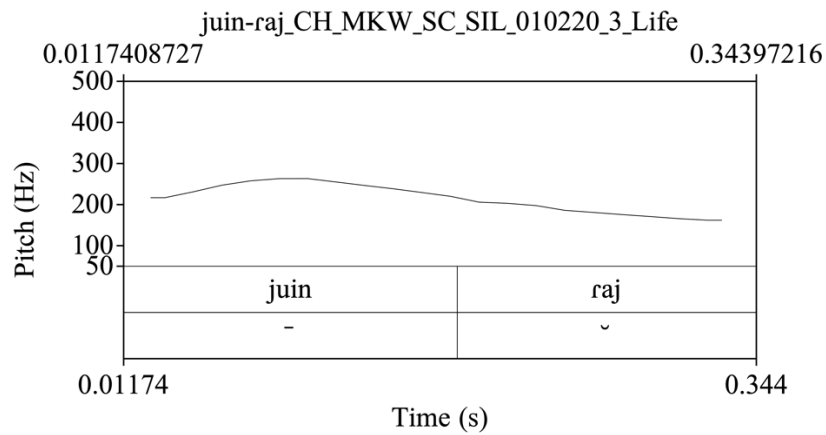
Figure 53. Trochaic stress in the dissyllabic word /mʰeʔ-ku/ 'fire smoke' in a carrier phrase

mʰeʔ-ku_SIL_F3_mʰeʔ-ku_carrier_phrase
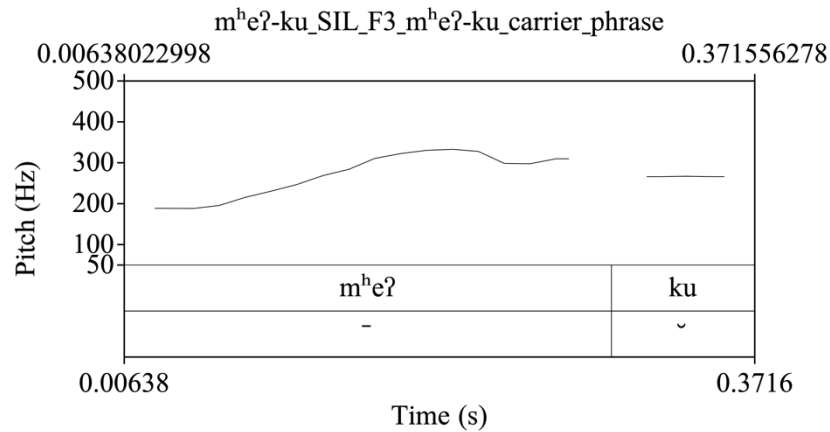
Trochaic stress is also attested in dissyllabic stems formed with monosyllabic roots suffixed with the non-productive derivational morpheme -*si*. This is illustrated in Figure 54 with the word /jo~jʌ-si/ [ˈjo~jʌ.si] 'Indian butter tree,' uttered in a sentence from a conversation.

215

Figure 54. Trochaic stress in the dissyllabic word /jo-si/ 'Indian butter tree' in a sentence from a conversation

jo-si_CH_MKW_RC_JMC_SIL_120119_Conversation

0.0153271984   0.395863413

0.01533   0.3959

Time (s)
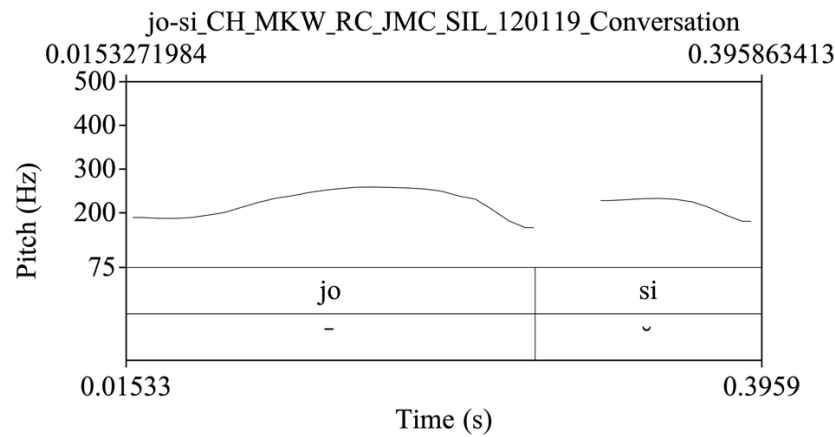
Finally, note that when the second syllable of a dissyllabic stem displays a glottal feature, either a pre-final sonorant or final glottal fricative or stop, the raising tone associated with it primes over the lexical stress expected pattern. This is illustrated in Figure 55, with the dissyllabic descriptive compound /jam-ahm/ [ˈjam.áhm] 'rice porridge' uttered in a sentence from a conversation, and in Figure 56, with the dissyllabic descriptive compound /tɕʰak-dʑaʔ/ [ˈtɕʰak-dʑáʔ] 'tiger spirit of the dead,' uttered in a sentence from an expository text.

Figure 55. Tone in the dissyllabic word /jam-ahm/ 'rice porridge' in a sentence from a conversation

jam-ahm_CH_MKW_RC_JMC_SIL_120119_Conversation

0.0127454005   0.437422262

0.01275   0.4374

Time (s)

216
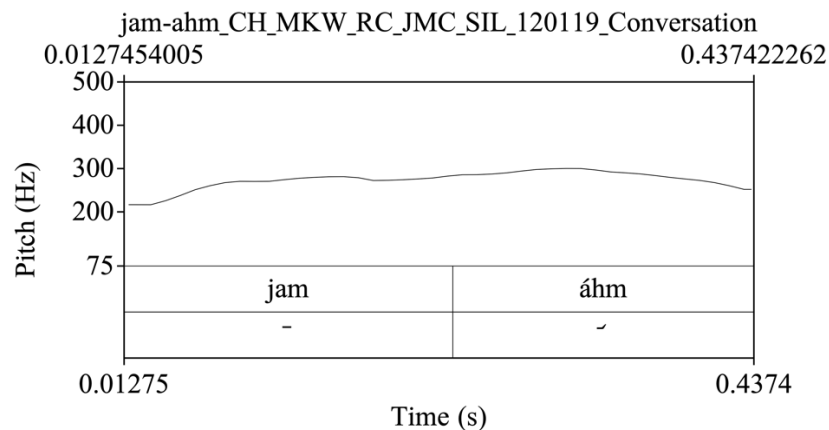
Figure 56. Tone in the dissyllabic word /tɕʰak-dʑaʔ/ 'tiger spirit of the dead' in a sentence from an expository text

tɕʰak-dʑaʔ_CH_MKW_SCBKC_SIL_081918_4_Chak_Ja_
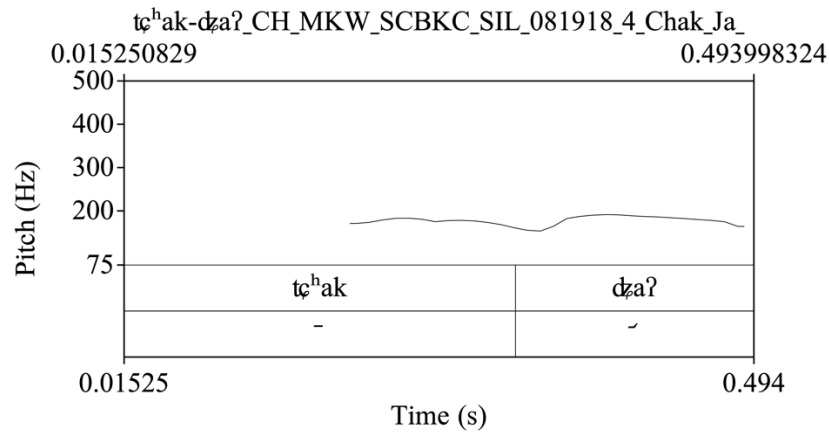
2.10.3.2. Spondaic stress on dissyllabic coordinative compounds, echowords, monosyllabic roots suffixed with the non-productive derivational morphemes -*tɕa* and -*tɕoŋ* or the productive prefixes *ma-* and *tɕoʔ-*

By contrast with determinative or descriptive compounds, coordinative compounds (§ 3.3.1), in addition to echowords (§ 3.3.2) and nominal stems formed with roots suffixed with the non-productive derivational morphemes -*tɕa* and -*tɕoŋ* to express a pair and group of kin relation (§ 3.3.4.2) or prefixed with the productive morphemes *ma-* and *tɕoʔ-* (§ 3.3.1.1.2) respectively meaning 'big' and 'small,' show a pitch prominence on each of the two syllables, resulting in a spondee (¯¯).

Spondaic stress in the dissyllabic coordinative compounds /la-ge/ [ˈla-ˈge] 'fishing rode' uttered in the carrier phrase /___ muna/ 'There is a fishing rod' and /laʔ-luiʔ/ [ˈláʔ-ˈluíʔ] 'arrow and bow' uttered in a sentence from an expository text are illustrated in Figure 57 and Figure 58, respectively, the dissyllabic echoword /maj-saj/ [ˈmaj.ˈsaj] 'meat and stuff' uttered in the carrier phrase /___ dʑeŋʌlʌ/ 'I don't eat ___' is illustrated in Figure 59, -*tɕa* derivation in Figure 60 with the word /ba-tɕa/ [ˈba.ˈtɕa] 'father and son' uttered in a sentence from a conversation, and ma- derivation with the word /ma-ɾu/ [ˈma.ˈɾu] 'big snake' uttered in the carrier phrase /___ muna/ 'There is a big snake' in Figure 61.

217

Figure 57. Spondaic stress in the dissyllabic word /la-ge/ 'fishing rod' in a phrase

POL_M1_la-ge_carrier_phrase_2

Figure 58. Spondaic stress in the dissyllabic word /laʔ-luiʔ/ 'arrow and bow' in a sentence from an expository text

laʔ-luiʔ_CH_MKW_SCBKC_SIL_081918_4_Chak_Ja_

Figure 59. Spondaic stress in the dissyllabic word /maj-saj/ 'meat and stuff' in a carrier phrase

maj-saj_SIL_F3_maj-saj_carrier_phrase

Figure 60. Spondaic stress in the dissyllabic word /ba-tɕa/ 'father and son' in a sentence from a conversation
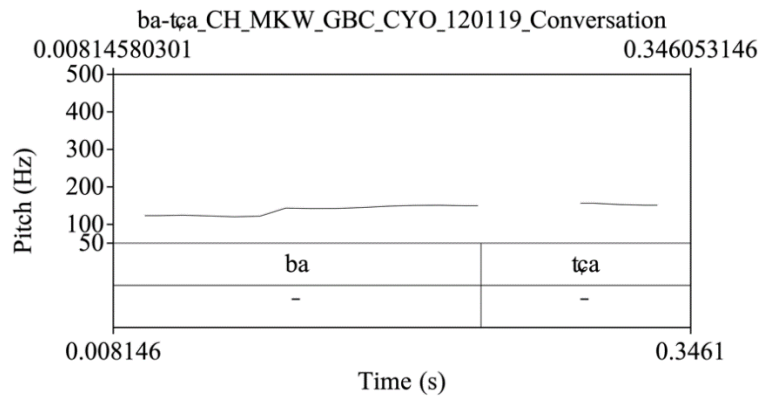
Figure 61. Spondaic stress in the dissyllabic word /ma-ɾu/ 'big snake' in a carrier phrase
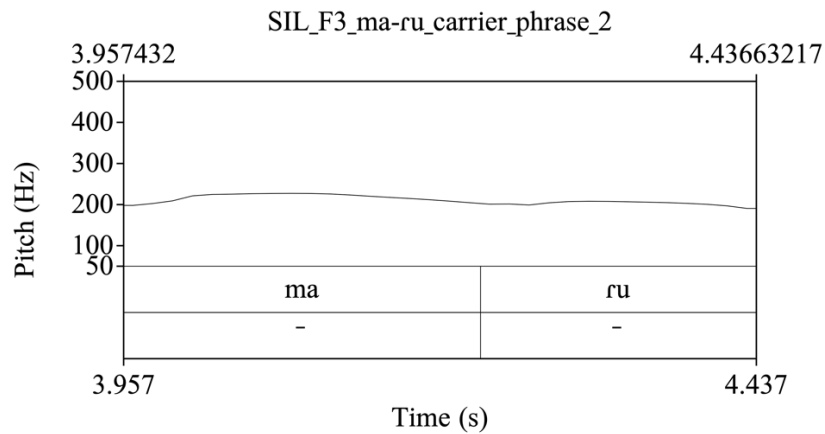
2.10.3.3.   Trochaic stress on dissyllabic inflected verbal stems

Dissyllabic words formed with inflected verbal stem show the same trochaic stress pattern as that found for dissyllabic roots and descriptive compounds. This is illustrated in Figure 62 with the word /lʰaknaŋ/ [ˈlʰak.nãŋ] 'I tell (a story)' uttered in the introduction of a narrative text, and in Figure 63 with the word /lʰakkan/ [ˈlʰak.kan] 'S/he told (a story)' uttered in isolation.

Figure 62. Trochaic stress in the dissyllabic word /lʰaknaŋ/ 'I tell (a story).' in the introduction of a narrative text.

lʰaknaŋ_CH_MKW_JMC_SIL_120619_2_Dung_Raja_ra_Dho_ng

0.014085074   0.504851476

Pitch (Hz)

| lʰak | naŋ |
| ‾ | ˘ |

0.01409   0.5049

Time (s)
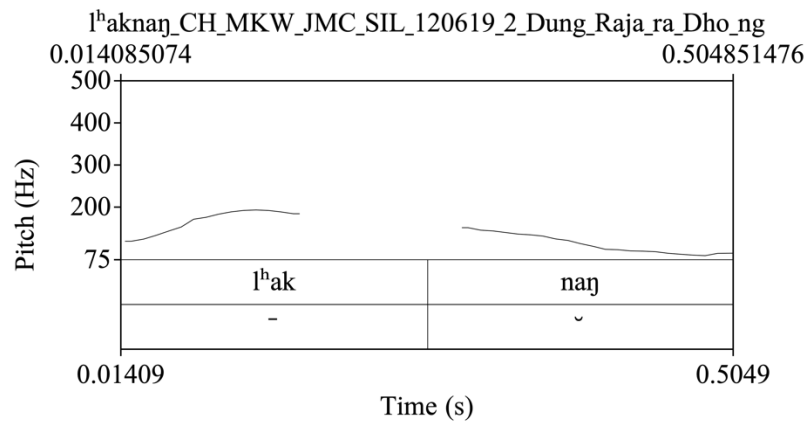
Figure 63. Trochaic stress in the dissyllabic word /lʰakkan/ 'S/he told (a story)' in isolation

SIL_F3_juinɾaj_lʰak-kan_isolation

0.676305304   1.44301839

Pitch (Hz)

| lʰak | kan |
| ‾ | ˘ |

0.6763   1.443

Time (s)
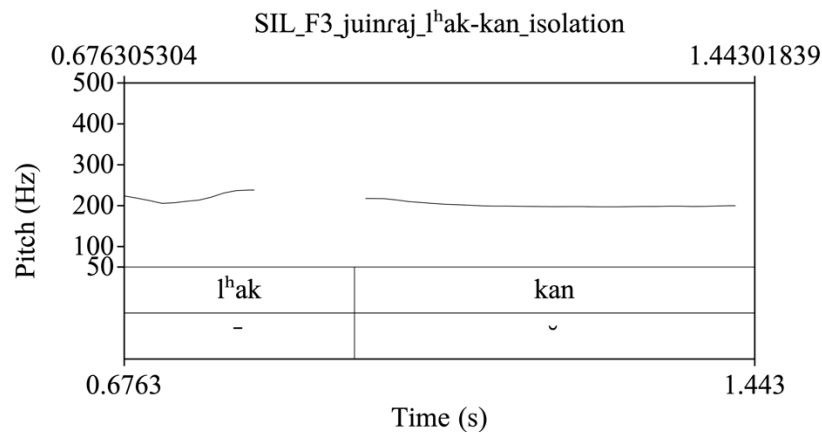
2.10.3.4.   Possible effects of elicited pronunciation on stress

When disyllabic words are uttered in isolation, and specifically in a word elicitation context, several types of pitch can be observed on the second syllable. These correlate with intonational stress that expresses different meanings that relate to pragmatics or discourse information structuring.

A fairly high pitch on the second syllable of a root resulting in a spondee (‾‾) can be triggered by the fact that the root or stem corresponds to the answer to the question

'How do you say *x*?' This is illustrated with the stem /ɾama/ [ˈɾa.ˌma] 'sickle' in Figure 64. This intonation conveys a meaning slightly different pragmatically, which relate to an epistemic coloration, which could be translated as 'obviously/certainly/for sure it is *x*,' or 'if you didn't get it yet, it's *x*.'

   When a dissyllabic word is pronounced in a three word repetition string, intonational stress can cover the second unstressed syllable in two different manners, tied to their position in the string: in first and second position, the final syllable of the stem may show a raising pitch which corresponds to an intonational stress that expresses the fact that more is about to be said, i.e., in this context, that the repetition string has not ended; in third position, the final syllable of the root may show a fairly falling pitch, which corresponds to an intonational stress that expresses the fact that nothing more is to be said, i.e., in this context, that the repetition string has come to an end. This is illustrated in Figure 64 with the nominalized verbal stem /ka.sa/ *ka=sa* put=NMZ1.

Figure 64. Intonational effect on words pronounced in isolation as an answer to 'How do you say *x*?'
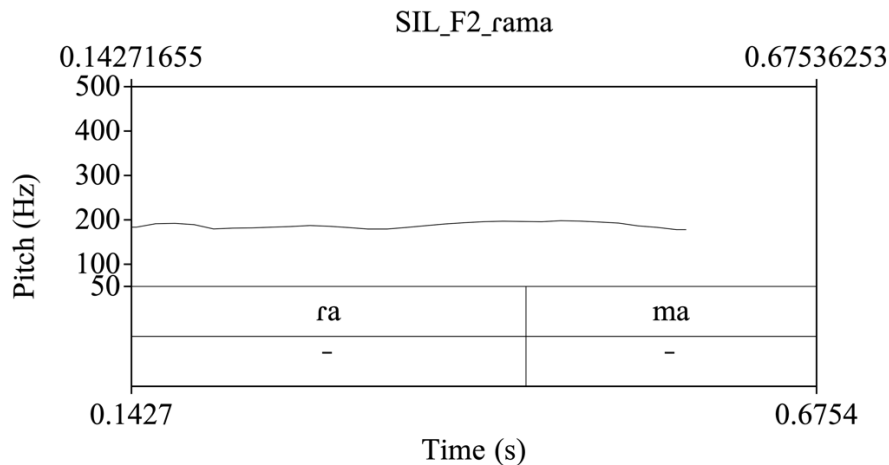
Figure 65. Intonational effect on words pronounced in isolation in a three repetitions string with the word
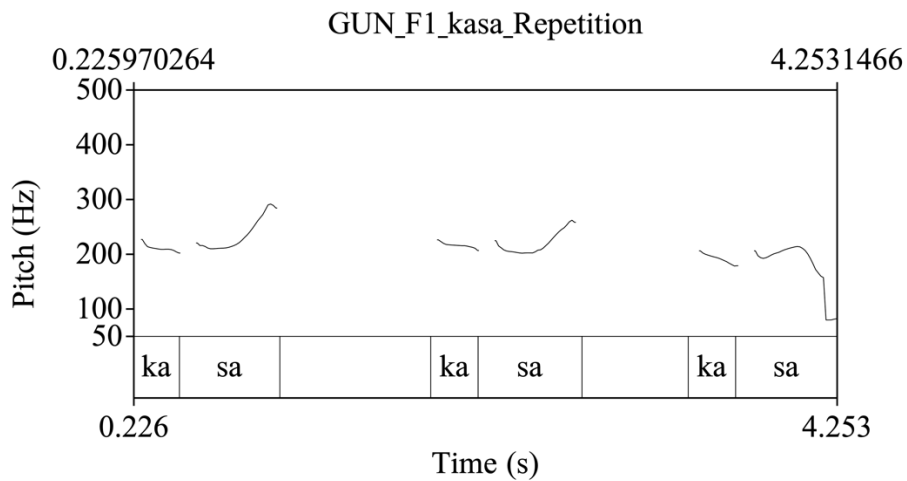
2.10.4. Stress on trisyllabic nominal root, stem, or word

Two types of stress are observed with trisyllabic words: antibacchius, described in § 2.10.4.1, and amphibrach is described in § 2.10.4.2.

2.10.4.1. Antibacchius on trisyllabic roots, descriptive compounds and roots prefixed with the productive morphemes *ma-* and *tɕoʔ-*

Trisyllabic roots have both the first and second syllable stressed, followed by an unstressed syllable. This stress pattern can be characterized as an antibacchius (¯¯˘). The trisyllabic root /paŋʌli/ [ˈpã.ˈŋu.li] 'firefly' uttered in the carrier phrase /___ tɕejewalaŋ/ 'I saw/found ___' is illustrated in Figure 66 and the trisyllabic root /tibiliŋ/ [ˈti.ˈbi.lĩŋ] 'butterfly' uttered in a sentence from a conversation is illustrated in Figure 67.

Figure 66. Antibacchius in the trisyllabic word /pãŋuli/ 'firefly' in a carrier phrase

paŋuli_SIL_F3_paŋuli_carrier_phrase

pa	ŋu	li
ˉ	ˉ	ˇ
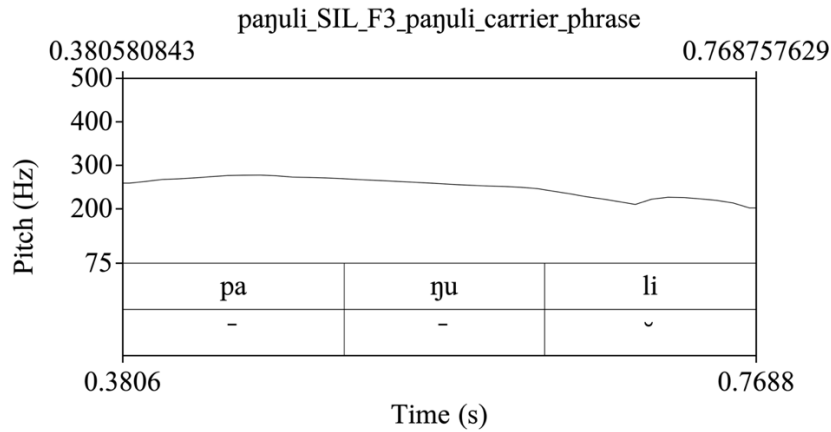
Figure 67. Antibacchius in the trisyllabic word /tibiliŋ/ 'butterfly' in conversation

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

ti	bi	liŋ
ˉ	ˉ	ˇ
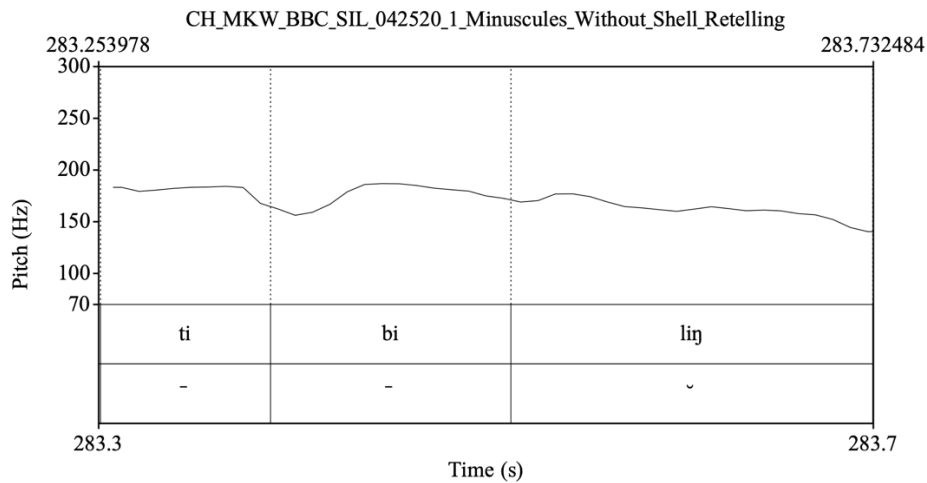
Antibacchius (ˉˉˇ) is also observed on descriptive compounds, as illustrated in Figure 68 with the word /umpuɾ-saj/ [ˈum.ˈpuɾ.saj] 'timur seed' uttered in the carrier phrase /___ muna/ 'There is some timur,' and in words formed with the productive prefixes *ma-* and *tɕoʔ-*, as illustrated in Figure 69 with the word /ma-kʌ-sja/ [ˈma.ˈkʌ.sja] 'big barking deer' uttered in isolation.

Figure 68. Antibacchius in the dissyllabic word /umpuɾ-saj/ 'timur seed' in a carrier phrase

SIL_F3_umpuɾ-saj_carrier_phrase_1

0.420337762   1.10862679

500
400
300
200
75

Pitch (Hz)

um	puɾ	saj
–	–	˘

0.4203   1.109

Time (s)
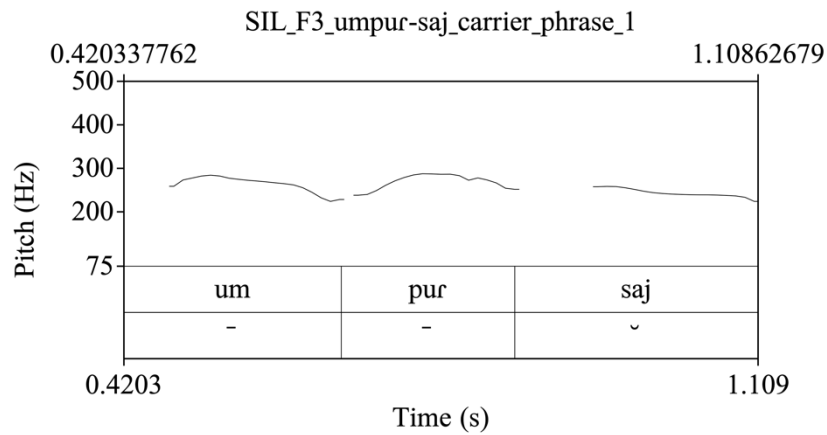
Figure 69. Antibacchius in the dissyllabic word /ma-kʌ-sja/ 'big barking deer' in isolation

SIL_F3_ma-kʌ-sja_isolation

3.07253427   3.65770951

500
400
300
200
75

Pitch (Hz)

ma	kʌ	sja
–	–	˘

3.073   3.658

Time (s)
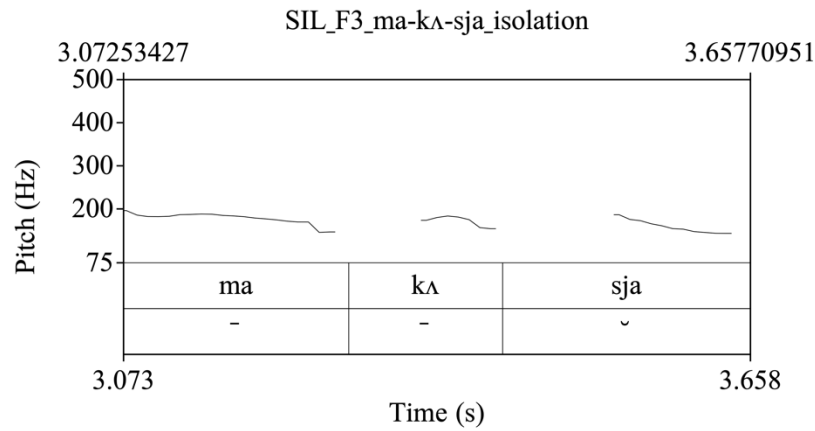
2.10.4.2.   Amphibrach on trisyllabic nouns

Amphibrach (˘ˉ˘) is also observed on trisyllabic nouns that may have originated from compounds, as illustrated in Figure 70 with the word /dikʌlak/ [ˈdi.ˈkʌ.lak] 'fish basket' uttered in an expository text.

Figure 70. Amphibrach in the trisyllabic word /dikʌlak/ 'fish basket' in a sentence from an expository text

dikʌlak_CH_MKW_MRNDC_SIL_081818_6_Dikalak
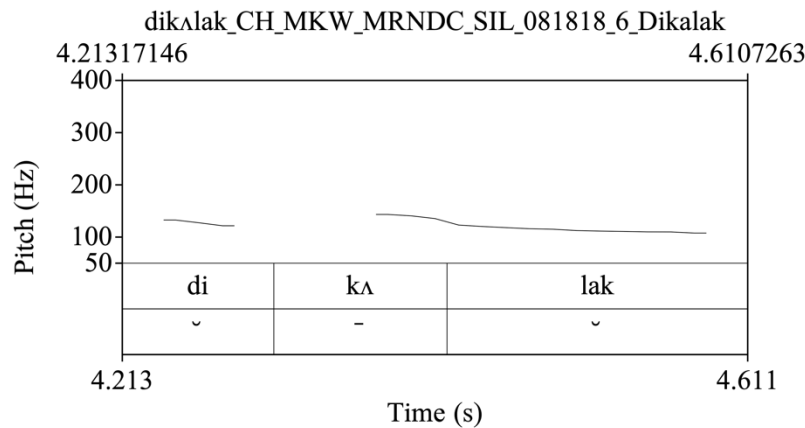
2.10.5. Stress induced glottal stop epenthesis in initial and intervocalic position

Caughley (1982) treats the glottal stop as a phonemic segment in initial position while I consider it phonemic only in final position. In his view, all syllables start with an underlying consonant and no vowel-initial syllable is allowed. This applies to native roots as well as roots borrowed from Nepali, which in fact itself does not feature any phonemic glottal stop in any position.

It seems that it is not always the case that a glottal stop is realized in initial position. The presence of a phonetic glottal stop in initial position can correlate with the environment surrounding the vowel-initial syllable word. When preceded by another vowel segment, that is, an open syllable, hiatus or resyllabification can be absent. If the vowel initial syllable is stressed, a glottal can be epenthesized, but when it is not stressed, it does not. This applies also at morpheme boundaries. Finally, the presence of a glottal stop in vowel-initial syllables will also depend on the "prosodic style" of the speaker. Some speakers will stress syllables in a way that shows hiatus or resyllabification entailing glottal stop epenthesis.

I suggest that glottal stop epenthesis in syllable initial position is articulatorily involved in the realization of stress. I observe that when a vowel-initial syllable is stressed, epenthesis of a glottal stop may or may not take place. When it does, stress is

perceived as much stronger than when it is absent. This equally applies for words in phrasal initial position and at morpheme boundary, whether the preceding syllable is open or closed. While this hypothesis requires more analysis through the comparison of stress realizations amongst speakers of different varieties, several studies have shown that prosodic (or suprasegmental) factors can indeed induce variation in word segmental realizations (Selkirk 1986; Cho, McQueen & Cox 2007; Kim & Cho 2013; Kim, Mitterer & Cho 2018).

2.11.   Chepang orthographies

I propose two orthographies that can be used to write in Chepang. In addition to representing the orthographic symbol proposed for each native phonemic segment (6 vowels, 6 vowel-glide and two-vowel sequences, and 33 consonants), I provide example words from Chepang along with their Nepali and English translations.

There is a growing interest in the Chepang community to have a writing system. The complexity of such endeavor in the case of Chepang has lied in the fact that Chepang features a glottal stop /ʔ/ which is not phonemic to Nepali or English, and thus has no available symbol in Devanāgarī or the Roman alphabet. Previous attempts (Caughley 1982; Caughley 2000; Caughley 2016) to develop an orthography have failed for the same reason, since the glottal stop was transcribed in Devanāgarī with the *visarga* diacritic symbol <ः> which does not correspond to the pronunciation of a glottal stop /ʔ/ but of a glottal fricative /h/; when read by Chepang speakers, this symbol leads to a mispronunciation and confuses the speakers. Therefore, it is better to add a new symbol specifically dedicated to the glottal stop, as proposed below.

The following tables only include the phonemes native to Chepang, and not borrowed consonants from Nepali. These latter are generally nativized. The borrowed retroflexes ट <ṭ> /ṭ/, ड <ḍ> /ḍ/, ठ <ṭh> /ṭʰ/, ढ <ḍh> /ḍʰ/ are often pronounced like alveolar stops, i.e., /t/, /d/, /tʰ/, and /dʰ/, respectively; these consonants can thus be transcribed as such, that is as alveolar stops त <t>, द <d>, थ <th>, ध <dh>. Another non-native phoneme

borrowed from Nepali is the post-alveolar fricative श <ś> /ʃ/ which is often pronounced like an alveolar fricative /s/. This also can be written as such, that is स <s>.

One proposed orthography is based on Devanāgarī and the other on the Roman alphabet. The Roman alphabet transcription is based on the International Alphabet of Sanskrit Transliteration (IAST), Velthuis, and Indian languages Transliteration (ITRANS), conventional systems of transliteration developed for Indo-Aryan (IA) languages written in Devanāgarī script. From IAST, another system called ISO-15919 was developed. One of the main differences between the two concerns the transcription of /e/ and /o/: with the macron <¯> in ISO-15919, such as <ē> and <ō>, and without it in IAST, such as <e> and <o>. A main difference between IAST and Velthuis lies in the distinction of the back mid-low vowel /ʌ/ from the central low vowel /a/. The vowel /ʌ/ is transcribed <a> in IAST and Velthuis, while /a/ is transcribed <ā> with the macron in IAST, and <aa> in Velthuis. The sound /ʌ/ may be written <a> and the sound /a/ either <ā> or <aa>.

The glottal stop /ʔ/, which is not phonological in IA languages, is transcribed <ˈ> (Unicode: U+02C8, UTF-8: CB 88) in the Roman alphabet, such as <jāˈ> or <jaaˈ> /dʑaʔ/ for 'tiger, leopard.' A new symbol has been integrated into Devanāgarī to represent the glottal stop in other languages that have it, such as Limbu, and a few other languages spoken in India. This symbol is ॽ (Unicode: U+097D, UTF-8: E0 A5 BD), such as जाॽ /dʑaʔ/ for 'tiger, leopard.'

Another difference between existing transliteration systems of Devanāgarī and the proposed Roman alphabet is the transcription of ङ as <ng> rather than <ṅ>.

The proposed Roman-based orthography was conceived to facilitate its usage with British and US English keyboards present on both basic mobile phones and smart phones. Since younger generations are texting in Chepang, proposing this roman orthography aims at supporting the use of the Chepang language in text messages.

Chepang vowels (6)

orthography			example words			translation	
roman	devanāgarī	IPA	roman	devanāgarī	IPA	Nepali	English
i	इ	/i/	ti	ति	/ti/	पानि	'water'
e	ए	/e/	le	ले	/le/	जिब्रो	'tongue'
a	अ	/ʌ/	bah	बह्र	/bʌh/	जेठाजु	'husband's elder brother'
aa / ā	आ	/a/	lā ~ laa	ला	/la/	डोरि	'rope'
o	ओ	/o/	ro	रो	/ɾo/	फुल्	'flower'
u	उ	/u/	gu	गु	/gu/	पिडालु	'taro'

Chepang vowel and glide and two-vowel sequences (6)

orthography			example words			translation	
roman	devanāgarī	IPA	roman	devanāgarī	IPA	Nepali	English
ui	उइ	/ui/	ruing	रउइङ्	/ɾuiŋ/	बाँस्	'bamboo'
ue	उए	/ue/	kue'	कउए?	/kueʔ/	बोलचि	'fishhook'
ai / ay	अय	/ʌj/	maing / mayng	मैङ्	/mʌjŋ/	नाम्	'name'
āi / āy aai / aay	आय	/aj/	ngāiksā / ngāyksā ngaaiksaa / ngaayksaa	ङाइक्सा	/ŋajksa/	ओड़ाल्नु	'stir'
yā / eā yaa / eaa	या	/ja/	dyāh / deāh dyaah / deaah	द्याह्र	/djah/	अब	'now, so, then'
oi	ओइ	/oi/	oik	औइक्	/oik/	पिठो	'roasted corn flour'

Chepang consonants (33)

orthography			example words			translation	
roman	devanāgarī	IPA	roman	devanāgarī	IPA	Nepali	English
p	प	/p/	pu'	पु?	/puʔ/	दाइ	'elder brother'
ph	फ	/pʰ/	phesā / phesaa	फेस्रा	/pʰesa/	छोड्नु	'let go, leave'
b	ब	/b/	bin	बिन्	/bin/	लसुन्	'garlic'
bh	भ	/bʰ/	bholsā / bholsaa	भोल्सा	/bʰolsa/	खन्नु	'dig up'
m	म	/m/	me'	मे?	/meʔ/	पुछर्	'tail'
mh	म्ह	/mʰ/	mhe'	म्हे?	/mʰeʔ/	आगो	'fire'
t	त	/t/	to	तो	/to/	ससुरा	'father-in-law'
th	थ	/tʰ/	thengsā / thengsaa	थेङ्	/tʰeŋ/	गन्नु	'count'
d	द	/d/	di'ng	दि?ङ्	/diʔŋ/	देउता	'deity'
dh	ध	/dʰ/	dhāhto / dhaahto	धाह्तो	/dʰahto/	तातो	'hot'

n	न	/n/	nāˑ / naaˑ	ना?	/naʔ/	दिदि	'elder sister'
nh	न्ह	/nʰ/	nhāpsā / nhaapsaa	न्हाप्सा	/nʰapsa/	ओछ्याउनु	'spread, put down'
k	क	/k/	kāh / kaah	काह्	/kah/	ढाड्	'back'
kh	ख	/kʰ/	khen	खेन्	/kʰen/	अनवार्	'face'
g	ग	/g/	guk	गुक्	/guk/	घाँटि	'throat, windpipe'
gh	घ	/gʰ/	ghon	घोन्	/gʰon/	हिले माछा	'barca snakehead fish'
ng	ङ	/ŋ/	ngā / ngaa	ङा	/ŋa/	म	'I, me'
ngh	ङ्ह	/ŋʰ/	nghʰur	ङ्हुर्सा	/ŋʰuɾ/	घुर्नु	'snore'
l	ल	/l/	lāh / laah	लाह्	/lah/	जुन्	'moon'
lh	ल्ह	/lʰ/	lhung	ल्हुङ्	/lʰuŋ/	मुटु	'heart'
r	र	/ɾ/	ru	रु	/ɾu/	सर्प	'snake'
rh	र्ह	/ɾʰ/	rhus	र्हुस्	/ɾʰus/	हाड्	'bone'
c	च	/tɕ/	cā / caa	चा	/tɕa/	घाउ	'wound'
ch	छ	/tɕʰ/	cheˑ	छे?	/tɕʰeʔ/	नुन्	'salt'
j	ज	/dʑ/	jāˑ / jaaˑ	जा?	/dʑaʔ/	बाघ्	'tiger, leopard'
jh	झ	/dʑʰ/	jhorsā / jhorsaa	झोर्सा	/dʑʰoɾ/	दोगनु	'greet'
y	य	/j/	yo	यो	/jo/	इन्द्रेणी	'rainbow'
yh	य्ह	/jʰ/	yhākto / yhaakto	य्हाक्तो	/jʰakto/	तितो	'bitter'
w	व	/w/	wiˑ	वी?	/wiʔ/	रगत्	'blood'
wh	व्ह	/wʰ/	whānto / whaanto	व्हान्तो	/wʰanto/	धारिलो	'sharp'
s	स	/s/	soˑ	सो?	/soʔ/	नसो	'vein'
h	ह	/h/	haw	हव	/hʌw/	भाइ	'younger brother'
'	?	/ʔ/	loˑ	लो?	/loʔ/	पात्	'leaf'

CHAPTER III

NOUNS AND NOMINAL MORPHOLOGY

In this chapter, I describe nouns and nominal morphology. I use the terms *nominal*, *nominal stem*, *noun stem* and *noun* interchangeably.

I briefly explain what I mean by *nominal root* and *nominal stem*, and what constitutes a noun-phrase (§ 3.1). I show morphological evidence for a distinction between a noun and a verb (§ 3.2). I then explore the processes attested in noun formation (§ 3.3).

Compounding is the most widespread nominal formation, and its structures are diverse (§ 3.3.1). Another type of nominal formation borrowed from Nepali is echo words (§ 3.3.2). A productive denonymic derivational morpheme -*mʌj* creates nouns that refer to the inhabitants of a locality; it shows additional functional developments and likely finds its source in a noun meaning 'people' (§ 3.3.3). Finally, two non-productive nominal formation processes are attested through the suffixation of bound morphemes: -*si* creates nouns which relate to species of trees, and -*tɕa* and -*tɕoŋ* express pairs and groups of kin referents, respectively (§ 3.3.4). The last two sub-sections dedicated to noun formation describe traditional first names attested in Chepang (§ 3.3.5), and kinship terminology (§ 3.3.6).

This chapter also describes nominal morphology, i.e., morphology associated with nouns that form noun-phrases to function as argument of a proposition (§ 3.4). I start by briefly describing the morphosyntactic shape of nominal derivational and inflectional morphology (§ 3.4.1) and give an overview of its paradigmatic and syntagmatic distribution (§ 3.4.2). I then describe singular, dual and plural number in addition to the nominalizing function of the plural morpheme (§ 3.4.3), derivational genitive and locational nominalizers (§ 3.4.4), case marking (§ 3.4.5), and relator nouns and postpositions (§ 3.4.6).

3.1. Nominal roots and stems

Nominal roots are mostly mono-, di-, or trisyllabic, such as *ro* 'flower,' *manta* 'person, people' or *tibiliŋ* 'butterfly' (§ 2.10).

A nominal stem may correspond to a bare nominal root or be the result of a complex formation and thus consist of more than one root, like in the case of a compound, such as *ba-ama* 'parents' ('father-mother'). Compounds may be formed through the combination of nominal roots, verbal and nominal roots, a compound and a nominal root, or else a noun and a deverbal noun. They can also combine root(s) and bound elements or derivational affixes.

All nominal inflectional morphemes (number and case marking) are enclitics attaching to the noun-phrase. Relator nouns and postpositions are not clitics (§ 3.4.1).

(41) *Chepang=lʌm=ko*   *lipi*   *na=lʌ.*
      Chepang=PL=GEN   writing   COP=NEG

'The Chepangs do not have a writing system.'
CH_MKW_SCBKC_SIL_081918_2_Chepang_king

3.2. Noun and verb distinction

Nouns and verbs are mainly distinguished cross-linguistically through the type of inflectional or derivational morphology they can take, and the function they can hold in a clause, such as argument or predicate. These criteria hold for Chepang.

In (42), the two noun-phrases formed with the noun *radza* 'king' are encliticized with the inflectional ergative morpheme *=i* and dative morpheme *=kaj*. These noun-phrases function as arguments of the predicate formed with the verb *kʰaj-* 'be able.'

By contrast, in order for a verbal root to form a noun-phrase argument of a predicate, it takes nominalizing morphology, such as the enclitics *=sa* and *=o* attached to the verbs *sat-* 'kill' and *noʔ-* 'speak,' respectively, in (42) and (43).

(42) *Dagu ɾadʑa=i   Ganamani   ɾadʑa=kaj*
      Dagu   king=ERG   Ganamani   king=DAT

      sat=sa=ma   kʰaj=u=lu.
      kill=NMZ1=ADD   be.able=3O/DIR=NEG

      'The king Dagu couldn't kill the king Ganamani.'
      CH_CTW_BBC_POL_102420_3_Chepang_Kings

(43) *Nepali   bʰasa   noʔ=o=kaj   sjoʔ*
      Nepali   language   speak=NMZ:REL=DAT foreigner

      to=na=ŋ=su.
      tell_say=NPST=1=1PL.EXCL

      'For those who speak Nepali, we say 'foreigner.'
      CH_MKW_MRNDC_SIL_081818_2_Chepang_Language_Culture

While nominal roots are free, verbal roots are bound. That is, nominal roots can form nominal words with no additional morphology, while verbal roots do not appear as words without additional morphology, either in a clause, or when expressed in their citation form. This is illustrated with the intransitive verb *al-* 'go' attached with the 2nd person intransitive imperative form =ʌ in (44).

(44) *"lʌw, lʌw,   tɕoʔ,   al=ʌ !"*   *dahj=ti=taŋ*
      well   well   child   go=2SG.IMP.INTR   say=SEQ1=ATT

      ama=i   to=a=tʰʌ=i.
      mother=ERG   tell_say=PST=INV=3>3SG

      '"Well, well, my child, go (there)!" the mother told her.'
      CH_MKW_DBC_MAI_2_020320_Newa_Dung

Some roots can be used as nouns or verbs with no morphological distinction, showing nominal and verbal properties, as shown in (45) to (47).

(45) *puʔ (arch.)* 'elder brother'
   puʔ-   'be elder'

(46) *naŋ=sʌj* *ŋa* *puʔ=na=ŋ.*
   2SG=ABL 1SG be.elder=NPST=1

  'Compared to you, I am the elder.'

  CH_CTW_SPC_POL_E

(47) *ŋa=ko*   *pʌɾiwaɾ=haŋ,* *ŋa* *puʔ=na=ŋ.*
   1SG=GEN family=LOC1   1SG be.elder=NPST=1

  'In my family, I am the elder.'

  CH_CTW_SPC_POL_E

3.3. Noun formation

In this section, I describe the four main morphological processes attested in noun formation: compounding (§ 3.3.1), echo words (§ 3.3.2), a productive denonymic derivational suffix *-mʌj* (§ 3.3.3), and two unproductive formation processes (§ 3.3.4): the suffix *-si* to create nouns that relate to species of trees, and the suffixation of two morphemes, i.e., *-tɕa* and *-tɕoŋ*, respectively refers to a pair or a group of human individuals that represent a kin relationship based on their resemblance. Finally, I describe the formation of traditional Chepang first names (§ 3.3.5) and kinship terminology (§ 3.3.6).

3.3.1. Compounding

In this section, I describe complex nominal stems formed through lexical compounding. Lexical compounding is a word formation process that derives a complex lexical word from the combination of two or more roots or stems (in the case of lexicalized compounds). While in most cases, the two elements of the compounds are free, independently used, there are many compounds formed with a bound element whose

meaning may be difficult to recover. In some cases, I propose historical analyses to reconstruct their possible original meaning.

Two main types of compounds are attested: determinative or descriptive compounds (§ 3.3.1.1) and coordinative compounds (§ 3.3.1.2).

The first type, determinative, or descriptive compounds, is the most attested and its formation more diverse. The second element is always nominal, functioning as the semantic head. Some are formed with two nominal roots (§ 3.3.1.1.1, § 3.3.1.1.2), or a verbal and a nominal root (§ 3.3.1.1.3); some show a nominal compound either modified by or modifying a nominal root (§ 3.3.1.1.4); some are formed with three elements, i.e., a noun, a deverbal noun and a noun (§ 3.3.1.1.5), and some constitute lexicalized noun-phrases (§ 3.3.1.1.6). The second type, coordinative compounds, is formed through the combination of two nominal roots (§ 3.3.1.2).

In addition to these two main types, there are also synthetic compounds formed with a noun and a deverbal noun or nominalized form of the verb, where the nominal root is an argument of the verb (§ 3.3.1.3). This type of compound is rare in our data: there is a set formed with the nominal root *mʰeʔ* 'fire' followed by a verb (§ 3.3.1.3.1) and another set formed with roots followed by a form *-ɾaj* that I analyze and reconstruct as a verb **ɾaj-* with the meaning of 'be perceived, felt' (§ 3.3.1.3.2).

More complex compounds are also attested, often formed with more than one syllable. In some cases, these may be analyzed as based on ideophones or onomatopoeia (§ 3.3.1.4).

3.3.1.1. Determinative or descriptive compounds

This type of compound is often referred to as determinative or descriptive[37], that is, one of the roots narrows down the semantic meaning of the other. The compounds described here are all endocentric: the meaning of the compound is a subset of the

---

[37] Referred to as *tatpuruṣa* and *karmadhāraya-tatpuruṣa* respectively in Sanskrit grammatical tradition.

meaning of the semantic head. The semantic head nominal root is preceded by a root that functions as a modifier, as the dependent.

The semantic head always refers to a generic semantic category or hypernym, such as *animal*, *bird*, *insect*, *plant*, or else *field* (often used in toponyms). This second element of the compound is always free, i.e., it can be used independently as a nominal. Table 81 gives examples of nominal roots that function as semantic heads.

Determinative or descriptive compounds can form kinship terms as well. Kinship terms are presented in § 3.3.6. An example of a kinship term formed through a determinative or descriptive compound is *tɕoʔ-bʰaw* 'son-in-law' where *tɕoʔ* means 'child' and *bʰaw* means 'husband,' i.e., literally 'child's husband,' expressing a possessive modifying relationship between the two nouns.

With most instances of this type of compounds, the modifying root is nominal (§ 3.3.1.1.1, § 3.3.1.1.2), but it can also be verbal (§ 3.3.1.1.3). Determinative or descriptive compounds can be formed with a compound modified by or modifying a nominal root (§ 3.3.1.1.4). Some are also formed combining a noun, a deverbal noun and another noun (§ 3.3.1.1.5). Finally, determinative, or descriptive compounds, can be lexicalized noun-phrases (§ 3.3.1.1.6). These are syntactic constructions, where the first stem, when nominal, is encliticized by the genitive morpheme *=ko*. A few examples are found with a verbal root as first element of the compound, in which case, it is attached with the nominalizer *=o*. I consider these constructions to be compounds, since the particular meaning they denote cannot be expressed otherwise in absence of the modifying possessive marker or nominalized construction.

Table 81. Widespread nominal roots used as semantic head in compounds

root	meaning	root	meaning
ahm	'porridge'	*musa*	'mushroom'
baŋ	'stone'	*njam*	'sun'
bʰʌjsi (<N.)	'buffalo'	*oʔn*	'grass'
bop	'snail'	*pas*	'yam *sp.*'
diʔŋ	'deity'	*ŋat*	'birdlime'
ʣaʔ	'tiger, leopard'	*ŋaʔ*	'fish'
gʰaŋ	'hole'	*ŋol*	'hornet'
gwej~goj	'taro'	*ɾaŋ*	'field'
jaŋ	'fly'	*ɾo*	'flower'
joʔm	'flea *sp.*'	*ɾʌ*	'creeper'
jam	'rice, millet'	*saj*	'fruit, seed'
juk	'monkey'	*siŋ*	'tree, wood'
kliʔ	'secretion, shit'	*sja*	'meat, flesh, cow'
lak	'yam *sp.*'	*sjaʔn*	'insect'
ljam	'path'	*ʨoʔ*	'child'
muh	'yam *sp.*'	*wa*	'bird'

3.3.1.1.1.   Noun – Noun

I present in Table 82 some examples of compounds formed with two nominal roots: the first root modifies the second root, which is the semantic head of the compound. The modifying relationship between the two nouns can be attributive or possessive. These compounds can be formed with native or borrowed roots from Nepali.

When the two roots of a compound are free, that is, used independently to form nominals, the meanings of the roots are often transparent regarding the meaning of the compound, like with *kʌdʌw-ahm* 'millet porridge' or *ɾiʔ-ahm* 'ancestors' porridge' (porridge offered to ancestors).

By contrast with *kʌdʌw-ahm* 'millet porridge,' which expresses an attributive relationship, *riʔ-ahm* 'ancestors' porridge' (porridge offered to ancestors) expresses a possessive relationship between the semantic head and the modifying possessor. This type of determinative compound is rarer and mostly found forming compounds relative to kinship terms (§ 3.3.6.1.2).

In some cases, either the current meaning of a root may not clearly match or reflect the meaning of the compound, or a root may not occur independently as a free nominal root, i.e., a bound root.

In the compound *moʔm-tɕoʔ* 'girl, woman' for instance, the root *tɕoʔ* means 'child' (extended to 'son'), and the root *moʔm* means 'granddaughter.' These elements do not carry the same meanings in the compound *moʔm-tɕoʔ* 'girl, woman.' It however reveals specific semantic concepts that are particular to the Chepangs: the root *moʔm* 'granddaughter' modifies the root *tɕoʔ* 'child,' which may suggest that all girls and women are socially considered granddaughters of someone.

The original meaning and the semantic layers of development of certain roots occurring in different compounds can be understood through comparison, as for example, with the compounds formed with the morpheme *aj*, as in: *aj-saj* 'cucumber,' *aj-gwej~goj* 'taro from previous season,' *aj-~ʌj-tɕʌn* 'crab *sp.*' The root *aj* is attested as a free root with the meaning of 'mother-in-law.' In the case of *aj-gwej~goj*, the compound is translated in Nepali as माउ तरूल <māu tarul>, which literally means 'mother taro.' In this case, the 'mother-in-law' associated with the figure of the 'mother' can explain the meaning of *aj-gwej~goj* 'taro from previous season,' since *aj* refers to an older root of yam. The compound *aj-~ʌj-tɕʌn* 'crab *sp.*' refers to a crab that sloughs its skin, exposing fresh, slippery skin like that of a newborn. The same can be said of *aj-saj* 'cucumber.' In these compounds, *aj-* likely carries the idea of 'before, first,' which probably was its original meaning, and the use of *aj-* in the compound *aj-saj* 'cucumber' may be the result of a metaphorical analogy with the texture of the *aj-~ʌj-tɕʌn* crab.

In some cases, the first root of the compound may be bound, or if free, may not occur in another compound. This is often the case with compounds that denote species of trees or plants, like with the first bound roots *mʌj-*, *jʌ-* and *a-* in compounds like *mʌj-saj* 'banana fruit/seed,' *jʌ-ro* 'Indian butter flower,' or *a-jam* 'foxtail millet.' In such

semantic domains, a few examples are found where the first root is attested independently, such as *njam-saj* 'Roxburghe fig fruit' formed with *njam* meaning 'sun,' or *tiŋ-saj* 'marking nut' formed with *tiŋ* meaning 'wrinkle.'

Other examples of such compounds are formed with the morpheme *sja*, which is free in some varieties and bound in others, and which may attach to free or bound first roots. The morpheme *sja* occurs as a free root which means 'cow' in some varieties (RAP-13, RAP-11, MAN-4) by contrast with others where the noun *tʰoɾ* is used instead to mean 'cow' (RAK-6). In others[38], the noun *sja* means 'meat, flesh,' which is likely its original meaning[39]. The morpheme *sja* in compounds clearly conveys the meaning of meat or flesh, since all refer to edible hunted preys, such as *kʌ-sja* 'barking deer,' *ɾuiŋ~ɾiŋ-sja* 'hog deer, sambar' or *gwa-~go-sja* 'small wild cat.' While the first roots *kʌ-* in *kʌ-sja* or *gwa-~go-* in *gwa-~go-sja* are bound roots, the root *ɾuiŋ~ɾiŋ* in *ɾuiŋ~ɾiŋ-sja* 'hog deer, sambar' is free and means 'bamboo,' used metaphorically suggesting the similarity in shape of the horns of both the hog deer and sambar with bamboo branches. The morpheme *sja* has evolved differently in all varieties and can be analyzed as a bound root in some varieties (RAK-6) where it only occurs in such compounds.

This type of compound recalls those denoting species of trees, where the first root, free or bound, is suffixed with the bound and non-productive morpheme *-si* 'tree' (§ 3.3.4.1). This non-productive morpheme is suffixed to native bound roots denoting species of trees and likely originates from *siŋ* 'tree, wood.' Note that when the first root expressing a species of tree is borrowed from Nepali, the bound form *-si* is absent, cf. *bel siŋ* 'Bengal quince tree' (<N. *bel*) vs. *got-si-siŋ* 'orchid tree' or else *tiŋ-si-siŋ* 'marking nut tree.' By contrast with the morpheme *-si*, which became a bound morpheme originating from a free root, the roots *saj* 'fruit, seed,' *ɾo* 'flower,' *jam* 'rice, millet' or *sja* 'meat, flesh' (in some varieties) for instance, remain free.

When only the meaning of the second root is recoverable, that of the first root is intrinsically reanalyzed as denoting the meaning of the referent expressed by the

---

[38] Reported by speakers of non-studied varieties spoken in Icchākāmanā and Kālikā Municipalities.

[39] The root *sja* is also cognate with the Magar root *sja* which means 'meat, flesh.'

compound, i.e., *a-* 'foxtail millet' in *a-jam* 'foxtail millet.' Whether the first root is bound or free, the function of the second root becomes merely that of denoting the type of semantic category to which the first root of the compound belongs. This is what functionally happens in compounds. But when the first root is bound, and the second root is also bound, like *-si* and *-sja* (in some varieties), they can be considered to having grammaticalized into derivational morphemes, although unproductive in the case of *-si*, or into some kind of class term or classifier. However, class terms and classifiers clearly do not have the same distributional and functional properties (Aikhenvald 2000: 1–4).

Finally, some noun-noun compounds involve metaphors based on physical resemblance between the signified of the compound and that of the first root. For instance, the compound *kwi-sjaʔn* 'dog caterpillar' is formed with the root *kwi* 'dog' modifying the root *sjaʔn* 'insect,' since this particular caterpillar has a small tail oriented upwards and a face that looks like that of a dog; in the compound *um-kliʔ* 'pus,' the root *um* meaning 'egg' refers to the secretion of a boil or furuncle that looks like an egg under the skin; the compound *nemet-saj* 'mulberry' is formed with *nemet* 'ant (red)' because the mulberry is thin, long and red, like the particular ant called *nemet*; or else in the compound *gwej~goj-ʨoʔ* 'boy, man,' the morpheme *gwej~goj* is attested as a root meaning 'taro,' or carrying the generic meaning of 'yam,' possibly used metaphorically to reflect a resemblance between a yam and male body parts.

Genitive constructions are also attested modifying *saj* 'fruit, seed' when the speaker needs to clarify that it is the seed and not the fruit that they are talking about, since *saj* means both when used independently. This is illustrated in (48). The speaker starts by using the genitive construction *mak-saj=ko saj* 'the seeds of the Bauhinia vahlii fruit,' and later uses simply *mak-saj* to refer to the same thing.

(48) *o=i,*   *o*   *mak-saj=ko*   *saj=taŋ,*
      DIST=ERG   DIST   Bauhinia.vahlii-fruit_seed=GEN   fruit_seed=ATT

      bʰitɾʌ=ko   *saj*   *nikal=ti, (…)*
      inside=GEN   seed   remove=SEQ1

      o=i   *mak-saj*   *pɾajk=ti*
      DIST=ERG   Bauhinia.vahlii-fruit_seed   remove=SEQ1

      waʔn=ti   *pʌjsa*   *bʌnaw=ti*   *gʌm=u=to.*
      take.away=SEQ1   money   make=SEQ1   keep=3O/DIR=REM.PST

      'Those seeds of the Bauhinia vahlii fruit, after he removed the seeds
      from inside, (…) having removed and taken away the seeds of the Bauhinia vahlii
      fruit, he made money and saved it.'
      CH_CTW_RC_KCR_101920_2_Sjaʔn

The types of compounds described here have in common that the second root refers to a generic category semantically narrowed down through the meaning of the first root. Further, whether the first or second root of the compound is free or bound, the second root, which is the semantic head of the compound, must be present to express the meaning denoted by the compound. This is also true when the first root is borrowed from Nepali, as with *gulap-ɾo* 'rose flower' where *gulap* (<N.) means 'rose' and *ɾo* 'flower.'

Table 82. Noun – Noun compounds

N-N	meaning	root 1	meaning	root 2	meaning
kʌdʌw-ahm	'millet porridge'	*kʌdʌw*	'millet' (<N.)	*ahm*	'porridge'
ɾiʔ-ahm	'porridge of ancestors'	*ɾiʔ*	'ancestor'	*ahm*	'porridge'
bel-siŋ	'Bengal quince tree'	*bel*	'Bengal quince' (<N.)	*siŋ*	'tree, wood'
got-si-siŋ	'orchid tree'	*got-si*	'orchid-tree'	*siŋ*	'tree, wood'
tiŋ-si-siŋ	'marking nut tree'	*tiŋ-si*	'marking nut-tree'	*siŋ*	'tree, wood'
tiŋ-saj	'marking nut fruit/seed'	*tiŋ*	'wrinkle'	*saj*	'fruit, seed'
mʌj-saj	'banana fruit/seed'	*mʌj-*	'banana'	*saj*	'fruit, seed'
bel-saj	'Bengal quince fruit'	*bel*	'Bengal quince' (<N.)	*saj*	'fruit, seed'
nemet-saj	'mulberry fruit/seed'	*nemet*	'ant'	*saj*	'fruit, seed'
aj-saj	'cucumber fruit/seed'	*aj*	'mother-in-law'	*saj*	'fruit, seed'
aj-goj~gwej	'taro (prev. season)'	*aj*	'mother-in-law'	*goj*	'yam'
kim-goj~gwej	'taro cultivated at home'	*kim*	'house'	*goj*	'yam'
kim-wa	'hen, bird raised at home'	*kim*	'house'	*wa*	'bird
ʥaʔ-wa	'coppersmith barbet bird'	*ʥaʔ*	'tiger, leopard'	*wa*	'bird'
moʔm-tɕoʔ	'girl, woman'	*moʔm*	'granddaughter'	*tɕoʔ*	'child, son'
goj~gwej-tɕoʔ	'boy, man'	*goj*	'taro'	*tɕoʔ*	'child, son'
a-jam	'foxtail millet'	*a-*	'foxtail millet'	*jam*	'rice, millet'
saŋa-jam	'barn millet'	*saŋa-*	'barn millet'	*jam*	'rice, millet'
kʌ-sja	'barking deer'	*kʌ-*	'barking deer'	*sja*	'meat, prey'
ɾuiŋ~ɾiŋ-sja	'hog deer, sambar'	*ɾuiŋ*	'bamboo'	*sja*	'meat, prey'
gwa-sja	'small wild cat'	*gwa-*	'small wild cat'	*sja*	'meat, prey'
kwi-sjaʔn	'dog caterpillar insect'	*kwi*	'dog'	*sjaʔn*	'insect'
tɕjuk-sjaʔn	'silkworm insect'	*tɕjuk-*	'silkworm'	*sjaʔn*	'insect'
mik-kliʔ	'tear, rheum'	*mik*	'eye'	*kliʔ*	'secretion, shit'
no-kliʔ	'earwax secretion'	*no*	'ear'	*kliʔ*	'secretion, shit'
um-kliʔ	'pus secretion'	*um*	'egg'	*kliʔ*	'secretion, shit'
ja-ŋaʔ	'Cyprinidae fish'	*ja-*	'Cyprinidae fish'	*ŋaʔ*	'fish'
gulap-ɾo	'rose flower'	*gulap*	'rose' (<N.)	*ɾo*	'flower'
koko-ɾo	'Proso millet flower'	*koko-*	'Proso millet'	*ɾo*	'flower'
jʌ-ɾo	'Indian butter flower'	*jʌ-*	'Indian butter'	*ɾo*	'flower'
lan-ɾaŋ	'spirit field (toponym)'	*lan*	'spirit, witch'	*ɾaŋ*	'field'
tes-ɾaŋ	'cemetery'	*tes-*	'cemetery'	*ɾaŋ*	'field'

3.3.1.1.2.   *ma-* and *tɕoʔ-* – Noun / Compound

A series of descriptive or determinative compounds is formed with either the noun *ma* 'mother' or *tɕoʔ* 'child, son' in initial position. These nouns modify the nominal semantic head of the compound with the meanings of 'big' and 'small,' respectively.

The semantic head consists of a closed class of nouns since all nouns are not attested in this type of compound. These nouns mainly relate to animals, but others like *kim* 'house,' *siŋ* 'tree,' or else *sʌjk* 'tooth,' are also attested. The semantic head can be a single nominal root or a compound.

These forms are considered compounds since they syntactically behave like nouns and not like a noun modified by an adjective. An adverb like *mah~maha* 'very' can modify an adjective like *pe=o* 'be nice=NMZ:REL,' but does not modify a noun like *ma-kim* big-house,' as illustrated in (49) to (51).

However, semantically, they contrast with compounds since they do not derive a particular type of referent, like *kʌ-sja* 'barking deer,' but just a referent qualified as big or small, such as *ma-kʌ-sja* 'big barking deer.' Some examples of compounds formed with *ma-* 'mother' and *tɕoʔ-* 'child' are given in Table 83.

(49)   *maha   pe=o   ma-kim*
        very   be.nice=NMZ:REL   big-house

        'very nice big house'

        CH_MKW_ PC_SIL_E

(50)   *maha   tahŋ=o   kim*
        very   be.huge_be.like=NMZ:REL   house

        'very huge house'

        CH_MKW_ PC_SIL_E

(51)   **maha   ma-kim*
        very   big-house

        CH_MKW_ PC_SIL_E

Table 83. *ma* and *ʨoʔ* – Noun / Compounds compounds

ma-N	meaning	*ʨoʔ*-N	meaning
ma-kim	'big-house'	*ʨoʔ-kim*	'small-house'
ma-wa	'big-bird'	*ʨoʔ-wa*	'small-bird'
ma-siŋ	'big-tree'	*ʨoʔ-siŋ*	'small-tree'
ma-ɾu	'big-snake'	*ʨoʔ-ɾu*	'small-snake'
ma-sja	'big-cow' (RAP-13)	*ʨoʔ-sja*	'small-cow' (RAP-13)
ma-kʌ-sja	'big-barking.deer'	*ʨoʔ-kʌ-sja*	'small-barking.deer'
ma-pjak	'big-pig_hog'	*ʨoʔ-ɾu*	'small-pig_hog'
ma-meʨʰja	'big-goat'	*ʨoʔ-meʨʰja*	'small-goat'
ma-ʣaʔ	'big-tiger'	*ʨoʔ-ʣaʔ*	'small-tiger'
ma-sʌjk	'big-tooth'	*ʨoʔ-sʌjk*	'small-tooth'

3.3.1.1.3. Verb – Noun

In verb-noun compounds, we can find the same type of nominal roots used as semantic heads in compounds formed with two nominal roots (Table 81). Examples of verb-noun compounds are given in Table 84.

In verb-noun compounds, the first root is often intransitive, and in most cases a stative intransitive verb; a few examples are found with a transitive verb as first element. In both cases, the first root modifies the second root, i.e., the semantic head, narrowing down its meaning. The nominal root is an underlying argument of the verb, either S with stative intransitive verbs, or either A or P with transitive verbs.

In Chepang, stative verbs take specific morphology when forming a non-verbal predicate, but they can also take regular intransitive morphology, form a verbal predicate and function as a dynamic verb (§ 5.4). They can also be used as adjectivals modifying nouns when encliticized by the nominalizer =*o* (§ 5.1). In verb-noun compounds, the verbal root occurs with no additional morphology and functions as a modifier of the

second root, similarly to what happens in noun-noun compounds (§ 3.3.1.1.1). Note that verbal roots in Chepang are all bound roots.

By contrast with compounds formed with two nominal roots, the meaning of the first root is often transparent. The meaning of the first root denotes specific characteristics that relate to color, taste, texture, or condition, in the case of stative intransitive verbs, or else that describe processes that are applied by or that apply to the referent of the second root in the case of transitive verbs.

Examples of compounds formed with stative intransitive verbs used as modifying elements are *du-ŋaʔ* 'balitoridae fish,' literally 'the fish that is red,' or 'red fish;' *jar-njam* 'sunset,' literally 'the sun that is yellow,' or 'yellow sun;' *brok-sjaʔn* 'caterpillar (dark skin, white hair),' literally 'the insect that has white hair;' or else *sjak-maj* 'raw meat,' literally 'the meat that is alive.' With stative intransitive verbs, the nominal root or semantic head holds the syntactic role of the verb's S argument.

Transitive verbs are also found in this type of compound. In such cases, the modified nominal root can either hold an underlying syntactic role of A or P.

The transitive verb *graj-* is used in both RAP-13 and RAK-6 varieties to mean 'grind to make flour (corn, millet, rice)' and semantically extended to mean 'grind (cardamom, cinnamon, black pepper),' in RAK-6. This verb occurs in compounds where the modified root can be either the A or P argument. In both RAP-13 and RAK-6 varieties, the compound *graj-baŋ* 'handmill stone' denotes a stone used to grind flour. The nominal root *baŋ* 'stone' holds an underlying syntactic role of A argument. In RAK-6, the compound *graj-sukmel* 'cardamom powder' is used as well, and in this case, the modified root *sukmel* 'cardamom' is the P argument of the verb. But this compound does not occur in RAP-13, where people use the lexicalized compound *grjak-~grek=o sukmel* 'cardamom powder,' literally 'the cardamom that got crushed,' with the transitive verb *grjak-~grek-* 'press on, pound, crush (cereals and condiments that have a chaff or outer dry layer).' This verb is also found in the compound *grjak-~grek-siŋ* 'sticks, twigs' used in both RAP-13 and RAK-6. In this compound, the root *siŋ* 'tree, wood' is the P argument, similarly to *graj-sukmel* or *grjak-~grek=o sukmel*. In RAP-13, the difference between *grjak-~grek=o sukmel* 'cardamom powder,' and *grjak-~grek-siŋ* 'sticks, twigs' lies in the fact that the cardamom got necessarily crushed by someone while the sticks or twigs are expressed as

244

a condition with no implication that this is the result of someone having broken them. Finally, the verb *gɾjak-~gɾek-* is not used as an independent verb in RAK-6 but *dahn-* 'press on, crush,' also used in free variation with *gɾaj-* 'grind.' In RAP-13, the verb *dahn-* 'press on, crush' is also used in free variation with the verb *gɾjak-~gɾek-* 'press on, pound, crush (cereals and condiments that have a chaff or outer dry layer).' It is likely that originally, the semantic valence of the transitive verbs *gɾaj-* and *gɾjak-~gɾek-* was different.

The transitive verb *ne-* 'sting' in the compound *ne-ŋaʔ* 'batasio fish,' which literally means 'the fish that stings,' entails that *ŋaʔ* 'fish' holds a syntactic role of A argument.

The two verb-noun compounds *sjak-ʥaʔ* 'tiger spirit of living people' and *tɕʰak-ʥaʔ* 'tiger spirit of deceased people' are less clear when it comes to understanding their formation.

The compound *sjak-ʥaʔ* 'tiger spirit of living people' is formed with the stative intransitive verb *sjak-* 'be alive.' This compound literally means 'the tiger that is alive' or 'living tiger' and refers to the spirit or soul of a person. But the meaning of the compound *sjak-ʥaʔ* really refers to 'the wandering tiger spirit of living people;' both spirits can leave the body of a person either while alive or after death. When alive, the person's spirit is expected to return to its host or the person would face death, whereas the spirit of the dead wanders looking for a living entity (human or animal) to possess. In the case of *sjak-ʥaʔ* 'tiger spirit of living people,' the morpheme *ʥaʔ* can be analyzed as the S argument of the verb *sjak-* 'be alive,' as the result of semantic analogy with the aliveness of the person to whom the spirit belongs.

In the case of the compound *tɕʰak-ʥaʔ* 'tiger spirit of deceased people,' *tɕʰak-* is found as a transitive verbal root *tɕʰak-* 'put sth above.' It could be that the verb *tɕʰak-* 'put sth above' metaphorically expresses that the spirit put itself above or upon a targeted living entity. Another possibility, which is rather unlikely because the vocabulary related to Chepang shamanism is native and the phonological forms of these two roots are not quite alike, is that the verbal form *tɕʰak-* was borrowed from the Nepali verb चढ्नु <caḍnu> 'mount, climb, ascend,' used in Nepali to talk about a spirit possessing a

245

shaman or person. In the compound tɕʰak-ʣaʔ 'tiger spirit of deceased people,' the nominal root ʣaʔ 'tiger, leopard' can be analyzed as an underlying A argument.

Table 84. Verb – Noun compounds

V-N	meaning	root 1	meaning	root 2	meaning
gɾaj-baŋ	'handmill stone'	gɾaj-	'grind'	baŋ	'stone'
gɾaj-sukmel (RAK-6)	'cardamom powder'	gɾaj-	'grind'	sukmel	'cardamom' (<N.)
gɾjak-~gɾek-siŋ	'sticks, twigs'	gɾjak-~gɾek-	'press on, crush'	siŋ	'tree, wood'
ɾok-saj	'chili'	ɾok-	'be spicy'	saj	'fruit, seed'
jʰak-lak	'bitter yam'	jʰak-	'be bitter'	lak	'yam (Dioscoreaceae)'
bɾes-ljam	'fork (path)'	bɾes-	'be forked'	ljam	'path'
jaɾ-njam	'sunset'	jaɾ-	'be yellow'	njam	'sun'
bɾok-sjaʔn	'caterpillar sp.'	bɾok-	'have white hair'	sjaʔn	'insect'
ahj-ŋat	'soft birdlime'	ahj-	'be soft'	ŋat	'birdlime'
tɕak-ŋat	'hard birdlime'	tɕak-	'be hard'	ŋat	'birdlime'
du-ŋaʔ	'Balitoridae fish'	du-	'be red'	ŋaʔ	'fish'
ne-ŋaʔ	'Batasio fish'	ne-	'sting'	ŋaʔ	'fish'
sjak-maj	'raw meat'	sjak-	'be alive'	maj	'meat'
sjak-ʣaʔ	'tiger spirit of the alive'	sjak-	'be alive'	ʣaʔ	'tiger, leopard'
tɕʰak-ʣaʔ	'tiger spirit of the dead'	tɕʰak-	'put sth above'	ʣaʔ	'tiger, leopard'

3.3.1.1.4.   Noun – Compound and Compound – Noun

Some compounds are formed with more than two elements, either a compound modified by a noun, or a compound modifying a noun. In this type of compound, the modifying or modified compound is formed by two nouns.

Some examples of noun-compound compounds are presented in Table 85 and compound-noun compounds in Table 86.

An example of a compound modified by a noun would be saʔ-liaŋ-saj 'strawberry,' where saʔ means 'earth, dirt, ground,' and liaŋ-saj 'Himalayan raspberry.' A compound modifying a noun is illustrated by a-jam-ahm 'Foxtail millet porridge,' where a-jam means 'Foxtail millet,' and ahm 'porridge.'

As mentioned above, some elements in this type of compounds are more difficult to analyze. The compound aj-laŋ-kliʔ 'first stool of newborn' for instance, has the element laŋ that independently means 'bread.' In this case, it is possible that laŋ recalls the white-yellowish color and/or shape of the stool.

246

Another example of a compound modifying a noun is *siŋ-rut-dʑaʔ-wa* 'cheer pheasant,' where *siŋ-rut-dʑaʔ*, which refers to the 'large Indian civet,' modifies *wa* 'bird.' The resemblance between the two species likely lies in their colors and triggered the formation of the compound that refers to the 'cheer pheasant.' The form *siŋ-rut* includes *siŋ* which means 'tree' and *rut* which means 'root.' The natural habitat of the large Indian civet (*siŋ-rut-dʑaʔ*) is the under the roots of trees.

Table 85. Noun – Compound compounds

N-comp.	meaning	root	meaning	compound	meaning
saʔ-ljaŋ-saj	'strawberry'	*saʔ*	'earth, dirt, ground'	*ljaŋ-saj*	'Himalayan raspberry'
aj-laŋ-kliʔ	'first stool of newborn'	*aj*	'mother-in-law; *first, before'	*laŋ-kliʔ*	'bread-secretion, shit'

Table 86. Compound – Noun compounds

comp.-N	meaning	compound	meaning	root	meaning
a-jam-ahm	'Foxtail millet porridge'	*a-jam*	'Foxtail millet'	*ahm*	'porridge'
siŋ-rut-dʑaʔ	'large Indian civet'	*siŋ-rut*	'tree-root'	*dʑaʔ*	'tiger'
siŋ-rut-dʑaʔ-wa	'cheer pheasant'	*siŋ-rut-dʑaʔ*	'large Indian civet'	*wa*	'bird'

3.3.1.1.5.   Synthetic Compound – Noun

Some compounds are also formed with a synthetic compound followed by a nominal semantic head. The synthetic compound consists of a deverbal noun that functions as the semantic head and its argument. The argument of the deverbal noun in the synthetic compound can be S or P. The synthetic compound attributively modifies the nominal semantic head of the compound.

 The few examples found in our data refer to species of birds, denoting their physical characteristics or behaviors. Some are presented in Table 87.

The two synthetic compounds that refer to species of flycatcher present synthetic compounds that consist of a stative intransitive verb and its S argument *meʔ* 'tail': *meʔ-*

mʌj-wa and *meʔ-jal-wa*, respectively literally meaning 'tail being small bird' and 'tail being long bird.'

The synthetic compounds can also consist of a transitive verb and its P argument: *ti-lʌʔn-wa* 'lapwing' literally means 'water lifting bird,' and *saʔ-ur-wa* 'kingfisher, bee-eater' literally means 'earth scratching bird.' Note that *saʔ-ur-wa* 'kingfisher, bee eater' is formed by calque with the Nepali compound माटो कोरे <māṭo kore> 'kingfisher,' literally 'earth scratching.'

In the case of *um~uhm-luk-wa* 'quail,' the form *luk* is not attested as an independent verbal root but can likely be reconstructed as such with the meaning of 'be round.' The form *luk* is also found in the compound *mik-luk-wa* 'owl' (RAK-6), which got clipped and became *luk-wa* 'owl' in RAP-13, which inspired the name of a local king as well, Lukwa Raja. In these compounds, *luk* denotes roundness, since owls have round eyes and quails have round eggs. Note that the morpheme *um* 'egg' developed an allomorph *uhm* in RAP-13 only in this compound, and not when *um* is independently used to mean 'egg.'

Caughley (2016) reports another form *luŋ* associated with roundness: *luŋ* 'ball-shaped object' and *mik-luŋ* 'eyeball.' While these forms are not recognized by the speakers I have been working with, it is possible that *luŋ* and *luk* are cognates. Both in Caughley (2016) and our data, the verb *luŋ-* is attested and means 'compliment, encourage, coax,' which could come from a metaphorical meaning of 'round off,' in the sense of 'smooth the rough edges.'

The two compounds *um~uhm-luk-wa* 'quail' and *mik-luk-wa* 'owl' (RAK-6) can likely be analyzed as formed with synthetic compounds that consist of a deverbal noun, which possibly etymologically reconstructs as **luk-* 'be round,' and its S arguments *um* 'egg' and *mik* 'eye,' and that modify the noun *wa* 'bird.' The literal meanings of *um~uhm-luk-wa* 'quail' and *mik-luk-wa* 'owl' (RAK-6) are respectively 'egg being round bird' and 'eye being round bird.'

Table 87. Synthetic Compound – Noun compounds

synth. comp.-N	meaning	synth. comp	meaning	root	meaning
meʔ-mʌj-wa	'flycatcher *sp.*'	*meʔ-mʌj-*	'tail-be small'	*wa*	'bird'
meʔ-jal-wa	'flycatcher *sp.*'	*meʔ-jal-*	'tail-be long'	*wa*	'bird'
ti-lʌʔn-wa	'lapwing'	*ti-lʌʔn-*	'water-lift'	*wa*	'bird'
saʔ-uɾ-wa	'kingfisher, bee-eater'	*saʔ-uɾ-*	'dirt-scratch'	*wa*	'bird'
um~uhm-luk-wa	'quail'	*um~uhm-luk-*	'egg-*be round'	*wa*	'bird'
mik-luk-wa	'owl' (RAK-6)	*mik-luk*	'eye-*be round'	*wa*	'bird'

3.3.1.1.6.   Lexicalized noun-phrases

Two types of compounds involve nominalizations, in which a modifying root is nominalized by the genitive morpheme =*ko* or nominalizer =*o*.

The first type is formed with a nominal root that functions as a modifier of a second root through the encliticization of the genitive morpheme =*ko*. The semantic relationship between the modified semantic head and the modifying first element can be seen as inalienable, such as for instance, the finger and the foot, as in *dom=ko bɾeh* 'finger of foot,' or the upper eyelid and the eye, as in *mik=ko kos* 'eyelid.'

In this type of compound, the second element, i.e., semantic head, can be found independently used as a verb. In this case, it is difficult to determine whether the semantic head is a nominalized form of the verb, or the verb originally comes from the noun. In both cases, they are the result of a zero-derivational process, i.e., derived with no presence of morphology.

The root *bɾeh* means 'finger' and *bɾeh-* is also a verbal root that means 'scatter, fork off, go in different directions.' It has two allomorphs: *bɾeh-* in RAP-13 *bɾes-* in RAK-6. The root *pak* means 'palm' and is also found as a verb *pak-* to mean 'suck up the juice and leave the rest (fruit).' Here again both meanings are linked since the process expressed by the verb involves the use of the palm of the hand. Such lexicalized noun-phrases are productive, as illustrated with the English borrowing *miniwatʌɾ* in the compound *miniwatʌɾ=ko ti* 'mineral water.' Some examples of lexicalized noun-phrases formed with the genitive morpheme =*ko* are illustrated in Table 88.

Table 88. Genitive =*ko* compounds

N=GEN N	meaning	root 1	meaning	root 2	meaning
dom=ko bɾeh	'finger of foot'	*dom=ko*	'foot.leg=GEN'	*bɾeh*	'finger'
kɾut=ko bɾeh	'finger of hand'	*kɾut=ko*	'hand=GEN'	*bɾeh*	'finger'
dom=ko pak	'sole'	*dom=ko*	'foot=GEN'	*pak*	'sole'
kɾut=ko pak	'palm of hand'	*kɾut=ko*	'hand=GEN'	*pak*	'palm'
kɾaʔ=ko dʰuʔŋ	'termite's egg'	*kɾaʔ=ko*	'termite=GEN'	*dʰuʔŋ*	'nest'
kɾaʔ=ko saʔ	'termite hill'	*kɾaʔ=ko*	'termite=GEN'	*saʔ*	'earth, dirt, ground'
mik=ko kos	'upper eyelid'	*mik=ko*	'eye=GEN'	*kos*	'eyelid'
ɾu=ko lʰjun~lʰjon	'snake's shed skin'	*ɾu=ko*	'snake=GEN'	*lʰjun~lʰjon*	'shed skin'
mʰeʔ=ko ljuŋ	'flame'	*mʰeʔ=ko*	'fire=GEN'	*ljuŋ-*	'burn (fire)'
meʔ=ko ɾʰus	'coccyx'	*meʔ=ko*	'tail=GEN'	*ɾʰus*	'bone'
ɾo=ko ɾoʔŋ	'pistil'	*ɾo=ko*	'flower=GEN'	*ɾoʔŋ*	'horn'
ŋaʔ=ko tɕʰap	'scale of fish'	*ŋaʔ=ko*	'fish=GEN'	*tɕʰap*	'scale'
ɾu=ko tɕʰap	'scale of snake'	*ɾu=ko*	'snake=GEN'	*tɕʰap*	'scale'
tuʔm=ko tɕoʔ	'bee larvae'	*tuʔm=ko*	'bee=GEN'	*tɕoʔ*	'child, son'
sja=ko tɕoʔ	'calf'	*sja=ko*	'cow=GEN'	*tɕoʔ*	'child, son'
miniwatʌɾ=ko ti	'mineral water'	*miniwatʌɾ=ko*	'mineral.water=GEN'	*ti*	'water'

In the second type of compound, the verbal root is nominalized by encliticization of the morpheme =*o*; the nominalization modifies the semantic head. The nominalized verb can be stative intransitive, like *bɾaw-* 'be tall' in the compound *bɾaw=o manta* 'adult.' When the verb is transitive, its argument, either A or P, can be expressed as well in the compound. In the compound *tuʔm sjas=o manta* 'beekeeper,' which literally means 'a person raising bees,' the noun *tuʔm* 'bee' is the P argument, while in the compound *no saʔj=o manta* 'deaf person,' *no* 'ear' is the A argument. None of these arguments are marked with case marking, which is further evidence for considering them as lexicalized noun-phrases.

In this type of compound, the semantic head can be a root referring to a hypernym, like the root *manta* 'person, people,' or else, *gʰaŋ* 'hole.' In both examples, the semantic head is the A argument, as in *git me-~ke=o manta* 'singer,' literally 'the person who sings songs,' or *tɕoʔ naʔ=o gʰaŋ* 'vagina,' literally 'the hole that gives birth to a child.' Negation can also be present in compounds formed with the nominalizer =*o*. Some examples of lexicalized noun-phrases formed with the nominalizer =*o* are illustrated in Table 89.

Table 89. Nominalizer =o compounds

V=NMZ N	meaning	V=NMZ	meaning	N root	meaning
bɾaw=o manta	'adult'	*bɾaw=o*	'be tall=NMZ'	*manta*	'person'

V=NMZ N	meaning	V=NMZ	meaning	N root	meaning
git me-~ke=o manta	'singer'	*git me~ke=o*	'song sing=NMZ'	*manta*	'person'
tuʔm sjas=o manta	'beekeeper'	*tuʔm sjas=o*	'bee raise=NMZ'	*manta*	'person'
tɕʰjus tɕʰjuʔ=o gʰaŋ	'urethra'	*tɕʰjus tɕʰjuʔ=o*	'urine urinate=NMZ'	*gʰaŋ*	'hole'
tɕoʔ naʔ=o gʰaŋ	'vagina'	*tɕoʔ naʔ=o*	'child give.birth=NMZ'	*gʰaŋ*	'hole'

N V=NEG=COP=NMZ N	meaning	V=NMZ	meaning	N root	meaning
no saʔj=ma=l=o manta	'deaf person'	*no saʔj=ma=l=o*	'ear hear=NEG=COP=NMZ'	*manta*	'person'
mik tɕjew=ma=l=o manta	'blind person'	*mik tɕjew=ma=l=o*	'eye see=NEG=COP=NMZ'	*manta*	'person'

Lexicalized noun-phrases that I consider compounds show morphosyntactic differences from non-lexicalized noun-phrases, although they still share similarities.

In the case of lexicalized compounds formed with the genitive marker =*ko*, the plural marker does not occur on the possessor but follows the possessed noun, as illustrated in (52) to (53). The possessive construction in (52), *lan=lʌm=ko tɕoʔ=lʌm* 'the children of the spirits,' contrasts with the compound *tuʔm=ko tɕoʔ=lʌm* 'bee larvae' in (53). The former refers to the children of particular (or definite) spirits and do not refer to a particular semantic category relative to a type of child, i.e., a spirit child. In fact, despite the existence of the compound *tuʔm=ko tɕoʔ=lʌm* 'bee larvae' that refers to a specific semantic category of larvae, if the presence of the plural morpheme =*lʌm* occurs on *tuʔm* 'bee,' as in (54), the possessive construction would be used to refer to the larvae of specific bees in a particular situation, i.e., contextually definite.

(52)   *o*   *Lanrang=sʌj*   *klaʔ=ti*   *waŋ=o*
        DIST   Lanrang=ABL   climb.down=SEQ1   come=NMZ:REL

        lan=lʌm=ko   *tɕoʔ=lʌm=paj,*
        spirit=PL=GEN   child= PL=GEN

        'Those spirits' children who came climbing down from Lanrang,'

(53) *tuʔm=ko*   *tɕoʔ=lʌm*
      bee=GEN   child=PL

      'bee larvae'

      CH_CTW_SPC_POL_E

(54) *tuʔm=lʌm=ko*   *tɕoʔ=lʌm*
      bee=PL=GEN   child=PL

      'the bees' larvae'

      CH_CTW_SPC_POL_E

The same is attested with compounds formed with a verb nominalized with the morpheme *=o*. The position of the plural marker *=lʌm* entails different constructions. Similarly to lexicalized noun-phrase compounds formed with the genitive morpheme *=ko*, the plural morpheme *=lʌm* is encliticized to the modified semantic head, as in (55) and (56), and not to the argument of the nominalized verb. A construction like (57) would correspond to a nominalization and not a compound, a construction that would not just mean 'singer' but rather 'the person who sang the songs' in reference to the particular songs sung by someone in an event. The same applies when the semantic head is also plural, as in (58).

(55) *tuʔm* *sjas=o*   *manta=lʌm*
      bee   raise=NMZ:REL   person=PL

      'beekeepers'

      CH_CTW_SPC_POL_E

(56) *git*   *me~ke=o*   *manta=lʌm*
      song   sing=NMZ:REL   person=PL

      'singers'

      CH_CTW_SPC_POL_E

(57)　　*git=lʌm*　　　*me~ke=o*　　　　*manta*
　　　　song=SML　　sing=NMZ:REL　　person

　　　　'the person who sung the songs'
　　　　CH_CTW_SPC_POL_E

(58)　　*git=lʌm*　　　*me~ke=o*　　　　*manta=lʌm*
　　　　song=SML　　sing=NMZ:REL　　person=PL

　　　　'the persons who sung the songs'
　　　　CH_CTW_SPC_POL_E

　　　　Lexicalized noun-phrase compounds formed with *=o* can occur without the
semantic head *manta* 'person, people,' only encliticized with the plural marker *=lʌm*, as
in (59). This construction is likely a calque from Nepali, which uses the same strategy, as
in (60). This construction conserves the same meaning than with the presence of the
semantic head. In this type of construction, the morpheme *=lʌm* can be reanalyzed as a
derivational morpheme, specifically an agentive nominalizer for plural animate referents.
Another example is given in (61).

(59)　　*tuʔm*　*sjas=o=lʌm*
　　　　bee　　raise=NMZ:REL=PL

　　　　'beekeepers'
　　　　CH_CTW_SPC_POL_E

(60)　　माउरी पाल्नेहरु
　　　　<*māuri*　　　　*pāl=ne=haru*>

　　　　bee　　　　　　raise=NMZ=PL

　　　　'beekeepers'
　　　　CH_CTW_SPC_POL_E

(61)　　*sjas=o=lʌm=i*　　　　　　*gʌm=na=n=i,*
　　　　raise=NMZ:REL=PL=ERG　　put=NPST=DIR/TR=PL

ŋi{=ko}   *dʑʌŋʌl=kʰa=paj*   *na=lʌ.*

1PL{=GEN}   jungle=LOC2=DIS   COP=NEG

'Those who cultivate (broom grass) keep it, but in our jungle, there is no.'

CH_MKW_RC_JMC_SIL_120119_Conversation

Semantically compounds formed through lexicalized noun-phrases derive a nominal that expresses a new referent or denotation. The noun *bɾeh* 'finger' refers to any type of finger, as in (62), while *kɾut=ko bɾeh* refers to a 'hand finger.'

(62)   *kʌn=nu*   *didi=ko*   *bɾeh*   *i=h*

       look=2PL.IMP.TR   elder.sister=GEN   finger   PROX=DEIC

'Look, elder sister's fingers, here.'

CH_CTW_JMC_PYK_101920_Cing_Lan

3.3.1.2.   Coordinative compounds

Coordinative compounds[40] denote the referents of both elements of the compounds. They do not express a modifying relationship. Some examples are presented in Table 90.

Table 90. Coordinative compounds

N-N	meaning
ba-ama	'parents; father and mother'
ahm-kjan	'meal; porridge and dish'
la-ge~la-gi	'fishing rod; thread/line and rod'
laʔ-luiʔ	'arrow and bow'

---

[40] Referred to as *dvaṅdva* in Sanskrit grammatical tradition.

**3.3.1.3.	Synthetic compounds**

Synthetic compounds present a deverbal noun as the semantic head and its argument. Synthetic compounds are rare in Chepang. Only two sets of synthetic compounds are attested: a set formed with the nominal root *mʰeʔ* as first element followed by different verbal roots (§ 3.3.1.3.1), and a set with different nominal or verbal roots as first element followed by a form *ɾaj* (§ 3.3.1.3.2), that I analyze as originally verbal and reconstruct as **ɾaj-* 'be perceived, felt.'

The nominalized verb or deverbal noun of synthetic compounds involve zero-derivation, i.e., they are bare verbal roots. The underlying argument of the deverbal noun can either hold the syntactic role of S or P, which is cross-linguistically often the case. Aikhenvald (2007: 32) notes that in synthetic compounds, underlying A arguments, or *transitive subjects*, "can hardly ever get compounded." She compares this restriction with that found with noun incorporation, i.e., a lexical derivational process where a noun combines with a verb to derive a complex verb form or verbal compound where the verb is the head (Sapir, Edward 1911; Mithun 1984). Cross-linguistically, in noun incorporation, P, S, and oblique arguments (location, instrument) are more often found incorporated to verbs (Mithun 1984; Keenan 1984).

**3.3.1.3.1.	*mʰeʔ* 'fire' compounds**

There is a set of five compounds formed with the nominal root *mʰeʔ* 'fire' as the first element. This type of compound is exocentric, that is, the semantic head is exterior to the compound and the meaning of the compound is not a subset of the meaning of either element of the compound. The meaning of the compound refers to the result of the process expressed by the deverbal noun. The semantic valence of these verbs varies, and the nominal root *mʰeʔ* 'fire' can hold the syntactic role of either S or P. They are presented in Table 91.

Table 91. *mʰeʔ* - Noun compounds

N-V	meaning	root 1	meaning	root 2	meaning
mʰeʔ-ku	'smoke (fire)'	*mʰeʔ*	'fire'	*ku-*	'smoke (fire)'
mʰeʔ-mʰut	'ash'	*mʰeʔ*	'fire'	*mʰut-*	'ignate, blow (fire)'
mʰeʔ-juɾ	'coals, embers'	*mʰeʔ*	'fire'	*juɾ-*	'squeeze a furuncle'
mʰeʔ-ɾiʔŋ	'soot'	*mʰeʔ*	'fire'	**ɾiŋ-*	'shake, agitate'
mʰeʔ-tʰap	'fireplace'	*mʰeʔ*	'fire'	*tʰap*	BHJ *cogn. tʰapu* 'ashes.'

Two of these compounds have verb roots that can occur independently, and whose meanings clearly relate to the meanings of the compounds: *mʰeʔ-ku* 'fire smoke' is formed with the intransitive verb *ku-* 'smoke (fire)' and *mʰeʔ-mʰut* 'ash' with the transitive verb *mʰut-* 'ignite, light up, blow (fire).' Both compounds can be analyzed as nominalizations expressing the result of the process denoted by the meaning of the verb. Literally, *mʰeʔ-ku* means 'fire smoking' and *mʰeʔ-mʰut* 'fire igniting.' The 'smoke' is the result of the process of *ku-* 'smoke (fire)' and the 'ashes' that of the process of *mʰut-* 'ignite, light up, blow (fire).'

The three other compounds are more difficult to analyze at first sight since their second element is either synchronically not used independently as a verb or its current meaning does not necessarily relate to the meaning of the compound. However, we may recover the original meaning and valence properties through internal reconstruction based on its comparison with other verbs. Indeed, Chepang has pairs, triplets, or groups of verbs, that show morphological traces of old derivational morphology, entailing that they synchronically express distinct semantic valency and argument structures (§ 5.2). As for the following description, note that final or pre-final sonorant glottal stop *-ʔ* is associated with P labile or anticausative constructions, while verbs with a final or pre-final sonorant glottal fricative *-h* is associated with intransitive middle causative constructions.

The second element *juɾ* in the compound *mʰeʔ-juɾ* is historically related to three verbs: *juɾ-* 'squeeze a furuncle or boil;' *juʔɾ-* 'extract juice (pressing with hand);' *juhɾ-* 'dive, sink into; be buried or shoved under (sand, ashes, hay).' These verbs have different

valency and thus entail different morphosyntactic properties at the clause level. The verb *juɾ-*, while synchronically specifically denoting the process of squeezing a furuncle, may likely have originally meant something like 'extract, express *x*.'

Indeed, the verb *juɾ-* 'squeeze a furuncle or boil' is transitive, as in (63). This verb is not P labile in that it does not occur in an anticausative construction with a P argument functioning as an S, as shown in (64) and (65).

(63) *o=i*   *lis*   *juɾ=ka=n.*
DIST=ERG   furuncle   squeeze=2/3.PST=DIR/TR

'S/he squeezed the furuncle.'
CH_MKW_PC_SIL_E

(64) **lis*   *juɾ=o*   *mu=na.*
furuncle   squeeze=PERF   COP=NPST

'The furuncle is squeezed.'
CH_MKW_PC_SIL_E

(65) **lis*   *juɾ=a.*
furuncle   squeeze=PST

'The furuncle was squeezed.'
CH_MKW_PC_SIL_E

The verb *juʔɾ-* 'extract juice (pressing with hand)' is P labile or anticausative. It can be used transitively with both A and P arguments, as in (66). But it can also occur intransitively in an anticausative construction with the presence of an S which holds the function of a P argument, as in (67), (68), and (69). The argument is then in an absolute form, and the verb presents intransitive morphology.

(66) *ŋa=i*   *kagʌti=ko*   *ɾos*   *juʔɾ=alaŋ.*
1SG=ERG   lemon=GEN   juice   extract=1.PST

'I extracted the lemon juice.'

(67)   *kagʌti=ko   ɾos   juʔɾ=a.*
       lemon=GEN   juice   extract=PST

'The lemon juice got extracted.'

(68)   *kagʌti=ko   ɾos   juʔɾ=o   mu=na.*
       lemon=GEN   juice   extract=PERF   COP=NPST

'The lemon juice is extracted.'

(69)   *kagʌti=ko   ɾos   pe=to   juʔɾ=na.*
       lemon=GEN   juice   be.nice=NMZ:ADV2   extract=NPST

'The lemon juice gets extracted / extracts nicely.'

Finally, the verb *juhɾ-* 'dive, sink into; be buried or shoved under (sand, ashes, hay)' is an intransitive middle causative verb. It can occur intransitively with an S argument whose participant performs the process expressed by the verb on itself. In that sense, this construction can be thought as middle since the Sp argument is referentially equal to both A and P, as in (70) and (71). Here, the referent of the S argument shoves themself under water or hay. In addition, the verb *juhɾ-* cannot have a P argument expressed as an S in absence of additional morphology if this latter is not also the initiator of the process, as in (72). The verb *juhɾ-* is thus not used in a transitive anticausative construction, which would then entail the presence of the causative marker *=tak* on the verb, as in (73).

(70)   *ŋa   ti=kʰa   juhɾ=alaŋ.*
       1SG   water=LOC2   dive=1.PST

'I dived into the water.'

(71) *ŋa   pʌral=kʰa   juhɾ=alaŋ.*
       1SG   hay=LOC2   dive=1.PST

‘I shoved myself into the hay.’

CH_MKW_PC_SIL_E

(72) **gwej mʰeʔ-juɾ=kʰa   juhɾ=o   mu=na.*
       taro   coal_ember=LOC2   shove=PERF   COP=NPST

‘The taro is shoved under the embers.’

CH_MKW_PC_SIL_E

(73) *gwej   mʰeʔ-juɾ=kʰa   juhɾ=tak=o   mu=na.*
       taro   coal_ember=LOC2   shove=CAUS=PERF   COP=NPST

‘The taro is shoved under the embers.’

lit. ‘The taro is caused to be shoved under the embers.’

CH_MKW_PC_SIL_E

This analysis shows that *juɾ-* is a transitive verb which likely originally meant ‘extract, express *x*’ although it has semantically been reduced to the context of extracting pus from a furuncle. Like *mʰeʔ-ku* ‘smoke (fire)’ and *mʰeʔ-mʰut* ‘ash,’ the synthetic compound *mʰeʔ-juɾ* was formed with the verb *juɾ-* in a nominalized antipassive form meaning ‘fire extracting, expressing’ to refer to the result of this process developing the meaning of ‘embers, coals.’

A little more challenging to understand is the compound *mʰeʔ-riʔŋ* ‘soot,’ since the form *riʔŋ* is not independently attested used as a verb, but as a noun meaning ‘black potter wasp.’ As seen, Chepang has verbs that present different valence properties associated with specific morphology. If a verb *riʔŋ-* was part of a triplet, it would be associated to the verb *rihŋ-* which means ‘tremble, shiver.’ For this triplet to be complete we would expect a verbal form *riŋ-*. But a verb *riŋ-* is not attested independently in Chepang. The only verbs attested with a form *riŋ* are *riŋ-kaj-* ‘search everywhere (small objects lifting things under which they may hide)’ as in (74), and *pitɕi-riŋ-* ‘have pins and needles,’ as in (75). If we compare the forms *riŋ* in the verb *riŋ-kaj-* ‘search everywhere’ and the verb *pitɕi-riŋ-* ‘have pins and needles,’ with the nominal root *riʔŋ* ‘black potter

wasp' and the verbal root *ɾihŋ-* 'tremble, shiver,' it is possible to understand the original meanings of the verbs **ɾiŋ-* and **ɾiʔŋ-* which again, are not used synchronically.

(74) *ŋa=ko*   *mobajl*   *ɾiŋ-kaj=ti*   *boŋ=ŋʌ=lʌ.*
      1SG=GEN   mobile   search=SEQ1   find=1=NEG

      'I searched my mobile all over and did not find it.'
      CH_CH_PPC_TAP_E

(75) *ŋa=kaj*   *dom*   *pitɕi-ɾiŋ=a=ta=ŋ.*
      1SG=DAT   leg   have.pins.needles=PST=INV=1

      'I have pins and needles in my leg.'
      lit. 'Leg pins and needles agitates me.' or 'I got agitated by leg pins and needles.'
      CH_MKW_PC_SIL_E

The verb *ɾiŋ-kaj-* means 'search everywhere (small objects lifting things under which they may hide).' The morpheme *-kaj* likely comes from *=kaj* (§ 5.7.8) which can derive an intransitive or transitive verb into a reciprocal, as in (76), (77) and (78). When this occurs, the verbs carry intransitive morphology, and the underlying arguments are left unmarked in the absolutive. It can also derive a transitive verb into a reflexive or middle, as one can see comparing (79) and (80). In this case, the underlying A argument is in the absolutive, the underlying P argument is marked as an oblique, as in (80), and the verb carries intransitive morphology, not reflexive, as shown with (80) and (81). Finally, the morpheme *=kaj* can also derive stative intransitive verbs into reflexive, as shown with (82) and (83).

(76) *ŋi*   *Chepang*   *bhasa*   *sjan=kaj=alaŋ=su.*
      1PL   Chepang   language   teach_learn=R/M=1.PST=1PL.EXCL

      'We taught each other the Chepang language.'
      CH_CH_PPC_TAP_E

(77)   *ʌni*   *kɾus=kaj=ti*   *noʔ=kaj=alaŋ=tɕʌ.*
        then   meet=R/M=SEQ1   speak=R/M=1.PST=1/3DU

    'And having met each other, we two talked to each other.'

    CH_MKW_BMB_BAN_090118_6_Love_and_marriage

(78)   *i-mi*   *tʰo=kaj=ka=i.*
        PROX-PL.H   beat=R/M=2/3.PST=PL

    'They beat each other.'

    CH_MKW_PC_SIL_E

(79)   *o=i*   *o*   *tɕoʔ=kaj*   *tʰo=ka=n.*
        DIST=ERG   DIST   child=DAT   beat=2/3.PST=DIR/TR

    'S/he beat that child.'

    CH_MKW_PC_SIL_E

(80)   *i*   *o*   *tɕoʔ=kusi*   *tʰo=kaj=a.*
        PROX   DIST   child=COM   beat=R/M=PST

    'S/he fought with that child.'

    CH_MKW_PC_SIL_E

(81)   *i*   *mjaŋ*   *kɾʌhn=ka=sʌ.*
        PROX   hair   comb=2/3.PST=REFL

    'S/he combed her/his hair.'

    CH_MKW_PC_SIL_E

(82)   *kim*   *lok=to*   *mu=na.*
        house   be.far=NMZ:ADV2   COP=NPST

    'The house is far.'

    CH_MKW_PC_SIL_E

(83) *ŋa kim=sʌj lok=kaj=alaŋ.*
  1SG house=ABL move.away=R/M=1.PST

'I moved away from home.'

CH_MKW_PC_SIL_E

The verb *riŋ-kaj-* 'search everywhere' likely literally means 'shake, agitate self; to bustle about' since the form *=kaj* has a detransitivizing function, and more specifically reflexive or reciprocal function. It is likely that the verb **riŋ-* was originally a transitive verb that meant 'shake, agitate *x*.' In fact, the form currently used to mean 'shake, agitate' is the verb *hʌlaw-* borrowed from Nepali. While the verb *riŋ-kaj-* 'search everywhere' was possibly formed with a transitive verb **riŋ-* 'shake, agitate *x*' suffixed with the reflexive or reciprocal morpheme *=kaj*, it is now completely morphologized since it takes transitive morphology, by contrast with other verbs derived with *=kaj*. This is illustrated in (84) to (87).

(84) *i=nis=tɕak tuʔm=kaj=ka=tɕʌ.*
  PROX=DU=CL2 kiss=R/M=2/3.PST=1/3DU

'These two kissed each other.'

CH_MKW_PC_SIL_E

(85) *tuʔm=kaj=dʑʌ !*
  kiss=R/M=2DU.IMP.INTR

'(You two) kiss each other!'

CH_MKW_PC_SIL_E

(86) *o=kusi tʰo=kaj=ʌ !*
  DIST=COM beat=R/M=2SG.IMP.INTR

'Fight with him/her.'

CH_MKW_PC_SIL_E

(87) *ɾiŋ-kaj=u !*
      search=2SG.IMP.TR

      'Search it everywhere!'

      CH_MKW_PC_SIL_E

In addition to the verb *ɾiŋ-kaj-* 'search everywhere,' the form *ɾiŋ* occurs in the verbal compound *pitɕi-ɾiŋ-* to mean 'have pins and needles (legs).' This verbal compound is formed with the verb *pitɕi-~pitɕej-* 'rub between fingers or two hand palms, or between a finger on the hand palm.' The form *pitɕi-~pitɕej-* is possibly an onomatopoeia denoting the process of rubbing. The verb *pitɕi-ɾiŋ-* 'have pins and needles (legs)' may have originally literally meant 'a rubbing (sensation) shakes, agitate *x*.' In fact, as illustrated above in (75), the P argument is marked with the dative and the verb takes inverse morphology which occurs when 1st person is the P argument. In this case, the literal translation of the above example (75) would be 'Leg pins and needles agitates me.' This analysis shows that we can likely reconstruct a transitive verb **ɾiŋ-* that originally meant 'shake, agitate *x*.'

The existence of the verb *ɾihŋ-* 'tremble, shiver' also support positing an earlier independent verb **ɾiŋ-* 'shake, agitate *x*.' While this latter must have been transitive and potentially A labile (antipassive), the verb *ɾihŋ-* 'tremble, shiver' is intransitive, as shown in (88) and (89). It is also used to describe the shaman in their trance. The form *ɾihŋ* is also a noun that means 'shaman's drum,' as in (90).

(88) *ɾihŋ=alaŋ.*
      shiver_tremble=1.PST

      'I shivered.'

      CH_MKW_PC_SIL_E

(89) *naŋ=i*   *ŋa=kaj*   *ɾihŋ=tak=te=a=ta=ŋ.*
      2SG=ERG   1SG=DAT   shiver_tremble=CAUS=2=PST=INV=1

      'You made me tremble.'

      CH_MKW_PC_SIL_E

263

(90) *ɾihŋ=ma*   *tajk=na=ŋ,*   *ama=ma*
      shaman.drum=ADD   beat=NPST=1   mother=ADD

      pande   *mʌ=na,*
      shaman   COP=NPST

      'I also beat the drum, my mother is also a shaman,'
      CH_MKW_PBPC_CHI_110219_5_Being_Shaman

It is possible that the verbs **ɾiŋ-* and *ɾihŋ-* formed a triplet with a former transitive verb **riʔŋ-* that was used as a P labile verb to describe P arguments occurring as S arguments. The verb **riʔŋ-* possibly meant something like 'get shaken, agitated' and the noun *riʔŋ* 'black potter wasp' likely literally meant 'the one that gets agitated,' referring to the behavior of wasps when they are disturbed. This analysis allows to better understand the formation of the compound *mʰeʔ-riʔŋ* 'soot' which literally means 'fire getting shacked, agitated,' the result of which being the presence of soot.

As for *mʰeʔ-tʰap* 'fireplace,' there is no other form *tʰap* attested in the lexicon. It is cognate with the Bhujel word *tʰapu* 'ashes,' Tibetan word *tʰap* 'fireplace,' and may simply have been borrowed from Tamang which uses the word *kutʰap* 'fireplace.'

3.3.1.3.2.   *-ɾaj* compounds

Another set of compounds is formed with *-ɾaj* as second element. While it is not used as a verb synchronically, I analyze *-ɾaj* as originally a verb that meant 'be perceived, felt.' Caughley (1982: 136–137) analyzes this form as part of a *nominal compound*, i.e., a compound verb in his terms, formed carrying the meaning of "the noise or sensation of an action." It can be interpreted as a deverbal noun forming nominalized compound similarly to the series of compounds formed with *mʰeʔ* (§ 3.3.1.3.1). All *-ɾaj* compounds attested in the studied varieties are presented in Table 92.

Table 92. *-ɾaj* compounds

N/V *-ɾaj*	meaning	root 1	meaning	root 2	meaning
wos-ɾaj	'axe'	*wos*	?	*ɾaj*	?
juin-ɾaj	'story, folktale'	*juin*	'bat'	*ɾaj*	?
gi-ɾaj	'broom grass'	*gi*	?	*ɾaj*	?
soʔ-ɾaj	'tiredness'	*soʔ*	'vein'	*ɾaj*	?
kliʔ-dʰam-ɾaj	'dung beetle'	*kliʔ-dʰam*	'secretion, shit-round'	*ɾaj*	?

Additional compounds formed with *-ɾaj* as second element were collected by Caughley (2016): *ŋiʔ-ray* 'laugh,' *noʔ-ray* 'talk, words, account,' *hah-ray* 'heat, hot season,' *ʔawh-ray* 'season of heat, sickness,' or else in the verb *chəh-ray-* 'be strong, vigorous.' Another set of compounds transcribed with a final glottal stop is also present in his dictionary: *yo-rayʔ* 'display, sight,' *hlak.rayʔ* 'account of events, information,' *chyanʔ-rayʔ* 'example, sign,' *dyumʔ-rayʔ* 'pastime, occupation to fill in time,' or else the verb *nə-rayʔ* 'defile, contaminate, weaken, cause sickness or wasting away (through contamination from dead body, by having dirty, unwashed skin),' *buh-rayʔ* 'heat and dryness (causing yellowing of leaves), season when this occurs.' None of these compounds is recognized by the Chepang speakers of the Lothar varieties.

The form *-ɾaj* in some of these compounds is transcribed with a glottal by Caughley (2016), as *rayʔ*. The compound *winʔ-rayʔ* 'bat' is transcribed with final glottals in Caughley (2016) while the form we collected is *juin-ɾaj* 'bat.' The form *wos* in the compound *wos-ɾaj* 'axe' is transcribed with a mid-central vowel [ə] in Caughley (2016), such as *wəs-ray* 'axe.' Another compound found in both our data and in Caughley (2016) is *soʔ-ɾaj*. Caughley (2016) translates it as 'joints (of legs).' When used in a sentence, *soʔ-ɾaj* occurs with the verb *dʑe-* 'eat' to mean being tired because of walking up a slope, as illustrated in (91). In this case, the compound *soʔ-ɾaj* functions as an A argument and *ŋa* 1SG as a P argument, marked with the dative morpheme *=kaj*. It is likely that *soʔ-ɾaj* means something like 'tiredness' since it metaphorically "consumes" the P argument marked with *=kaj* DAT in the context of a physical effort. The nominal root *soʔ* occurs independently with the meaning of 'vein.' The two forms *soʔ* could be homophonous but could as well have evolved from the same morpheme.

(91)   ŋa=kaj   soʔ-ɾaj   ʥe=ti,   wah=sa
       1SG=DAT   tiredness   eat=SEQ1   walk=NMZ1

kʰaj=ŋʌ=lʌ.

be.able=1=NEG

'Being tired, I can't walk.'

lit. '*x* having eaten me, I can't walk.'

CH_MKW_PC_SIL_E

Some of the elements present in the compounds that are not attested in our data are defined in Caughley's dictionary (2016). Although many are either not recognized by the Chepang speakers I have been working with, or show formal differences, they could inform the nature of -ɾaj compounds. It is indeed possible that all these compounds were formed with the same verb ɾaj- as second element, whose original meaning was 'be perceived, felt.' It was used to create deverbal compounds denoting the result of a process, a formation similar to that of the compounds formed with mʰeʔ 'fire' (§ 3.3.1.3.1).

For instance, the verb wəs- is defined (Caughley 2016) as 'splash (water, etc.) with instrument; fling or flick up (e.g. dirt with cow's horns)' and also as 'reverberate; sound well (singing);' the verb ɾayʔ- is defined as 'make noise (unintentionally, esp. with movement)'. In our data, the form wos- of the compound wos-ɾaj is not attested as a verb, nor is wʌs-. The verb ɾaj- is attested and its meaning is similar to that found in Caughley (2016), i.e. 'make noise (unintentionally, esp. with movement).' The compound wos-ɾaj 'axe' or wəs-ɾaj 'axe' in Caughley (2016), could literally mean 'the reverb making noise.' Given the compounds formed with -ɾaj and the current meaning of the verb ɾaj- 'make noise (unintentionally, esp. with movement),' it may originally have meant something like 'be perceived' or 'be felt.' While its exact semantic valence remains unclear, it may have been an intransitive verb. The compound wos-ɾaj 'axe' could have literally meant 'the reverb being perceived.' The compound juin-ɾaj 'story, folktale' may have meant something like 'bats being perceived,' since the Chepang traditionally practice storytelling by night when bats come out. If this analysis holds, the compound soʔ-ɾaj would have literally meant 'the veins being felt,' to denote physical tiredness and

266

pain.' Finally, the compound *kliʔ-dʰam-ɾaj* 'dung beetle,' could have literally meant 'dung ball being perceived,' since the dung beetle is visible from far away going around rolling dung balls. The form *dʰam* in this compound is not used independently. It is found in other compounds describing spherical or circular things, some of which are only attested in Caughley's dictionary (2016), such as: *krama-dhama* 'bulging (bag),' *bəw-dham siŋʔ* 'a medium-size tree with circular leaves,' *layʔ-dham-loʔ* 'circular leaf of the *layʔ-dham* tree.'

The form *dʰam* is attested in a compound found as well in RAK-6 and RAP-13 that describes a thorny tree: *ma-dʰam-siŋ* 'madham tree,' *ma-dʰam-saj* 'madham seed/fruit,' *ma-dʰam-tɕuʔ* 'madham thorn.' The compound *ma-dʰam-saj* can be interpreted literally as 'big round fruit.' The skin and flesh of the fruit is traditionally pounded to make soap to wash the body and clothes. The thorny tree was likely given the name of *madʰam* in reference to its big round fruits of the size of a guava. Another compound, *ma-dʰam-bɾeh* is used in all studied varieties to mean 'thumb,' literally, 'big round finger.'

The reconstruction of possible literal original meanings for *-ɾaj* compounds are presented in Table 93, including the compounds collected by Caughley (2016). In some cases, the meaning of the first element of this type of compounds remains unknown.

Table 93. Reconstruction of *-ɾaj* compounds

N/V *-ɾaj*	meaning	root 1	meaning
wos-ɾaj	'axe' < 'reverb being perceived'	*wəs-~wos-*	'reverberate'
juin-ɾaj	'story, folktale' < 'bats being perceived'	*juin*	'bat'
gi-ɾaj	'broom grass' < '? being perceived'	*gi*	?
soʔ-ɾaj	'tiredness' < 'vein being felt'	*soʔ*	'vein'
kliʔ-dʰam-ɾaj	'dung beetle' < 'dung ball being perceived'	*kliʔ-dʰam*	'shit ball'

other compounds from Caughley (2016)[41]			
ŋiʔ-raj	'joke' < 'laugh being perceived'	*ŋiʔ-*	'laugh'
noʔ-raj	'speech' < 'speaking being perceived'	*noʔ-*	'speak'

[41] Although they may be different, Caughley's (2016) transcriptions are preserved.

hah-~awʔ-raj	'hot season' < 'heat being perceived'	*hah-~awʔ-*	'heat' ?
chəh-raj-	'be strong, vigorous' < 'strength being felt'	*chəh-*	'strength' ?
jo-rajʔ	'display, sight' < 'sight being perceived'	*jo-*	'see'
hlak-rajʔ	'account of events' < 'telling being perceived'	*hlak-*	'relate, tell'
chyanʔ-rayʔ	'example, sign' < 'explanation being perceived'	*chyanʔ-*	'explain'
dyumʔ-rayʔ	'pastime' < 'the end being perceived'	*dyumʔ-*	'be finished'
nə-rayʔ	'contaminate through body' < ? being perceived'	*nə-*	?
buh-rayʔ	'heat and dry season' < ? being perceived'	*buh-*	'heat' ?

3.3.1.4.   Complex and onomatopoeic compounds

There are also nominal compounds that are morphologically more complex, presenting an element of more than one syllable preceding the semantic head. These compounds often refer to animal species, in particular birds, insects, bats, or fish. These need to be analyzed investigating the zoological knowledge and folk taxonomy speakers have of these species. I present some examples of these in Table 94.

Table 94. Complex compounds

compound	meaning	compound	meaning
intɛʌ-ŋaʔ	'shrimp'	*tanapɾuk-sjaʔn*	'caterpillar *sp.*'
njalja-bop	'slug'	*dewkʌl-wa*	'Himalayan bulbul'
paltot-wa	'parakeet *sp.*'	*ɾapaŋ-juin*	'bat *sp.*'
daŋsaɾ-wa	'heron, hoopoe'	*ɾiki-njaʔm*	'grasshopper *sp.*'

Some nominal compounds can be formed through onomatopoeia. This is often found in names of birds, where the syllables that form the first element of the compound mimic the birds' chirp. For instance, the bird referred to as *ʥʰiʥʰi-wa* 'citrine wagtail' sings [ʥʰi.ʥʰi], or the one called *tuŋkuɾ-wa* 'bulbul *sp.*' sings [tuŋ.kuɾ].

Some compounds show only one onomatopoeic element. In this case, the meaning of the other element is often difficult to understand. The bird called *paŋkʌt-wa* 'owl *sp.*' sings [paŋ.paŋ] while the element *kʌt* is unknown.

In the compound *njam-tɕi-wa* 'Nepal wren babbler,' the onomatopoeic element *tɕi* combines with *njam* 'sun.' The onomatopoeic element *tɕi* recalls the Nepal wren babbler's chirp [tɕi.tɕi] one can hear when the sun rises. These onomatopoeic compounds are reported in Table 95.

Table 95. Onomatopoeic compounds

compound	bird's chirp	meaning
dʑʰidʑʰi-wa	[dʑʰi.dʑʰi]	'citrine wagtail'
tuŋkuɾ-wa	[tuŋ.kuɾ]	'bulbul *sp.*'
hudʑuɾuŋ-wa	[huː.dʑu.ɾuŋ]	'Indian eagle owl'
paŋkʌt-wa	[paŋ.paŋ]	'owl *sp.*'
njamtɕi-wa	[tɕi.tɕi]	'Nepal wren babbler'

3.3.2. Echo words

Echo words (Abbi 1994) or modified reduplication (Diffloth 1976) are found in IA languages like Hindi and Nepali. Echo words formation is productive and consists of replacing the initial consonant of the noun or adding a consonant if the noun starts with a vowel. The new denotation refers to the referent of the noun and things associated with it. This construction was borrowed into Chepang from Nepali. In both Chepang and Nepali, the alveolar fricative /s/ is mostly frequently chosen as the replacing consonant. Some echo words have been borrowed wholesale from Nepali (or English via Nepali) and others can be productively created using Chepang words.

Echo words are used as a humorous cover for talking about the referent in a derogatory manner, i.e., presenting the referent as meaningless or not important. The use of echo words often conveys the speaker's feeling of anger, or irritation. Some echo words are presented in Table 96, and examples are given in (92) to (97).

When nominal morphology is present, it can occur twice after the noun and after the echo form, as in (96), or once after the echo form, as in (97). Syntactically, a noun and an echo form can either constitute a single noun-phrase or two noun-phrases.

Table 96. Echo words

echo words	meaning
tɕija-sija	'tea and stuff' (<N.)
ahm-sahm	'cooked rice and stuff'
maj-saj	'meat and stuff'
dal-sal	'lentil and stuff' (<N.)
kjan-sjan	'vegetable/dish and stuff'
kam-sam	'work and stuff' (<N.)
manta-santa	'people and stuff'
kwi-swi	'dogs and stuff'
metɕʰja-setɕʰja	'goats and stuff'
pjak-sjak	'hogs and stuff'
kukʌr-sukʌr	'pressure cooker and stuff' (<N. <E.)
tʰal-sal	'plates and stuff' (<N.)

(92)   *e,*   *ŋa=i*   *dʑe=ŋʌ=lʌ,*   *ahm-sahm !*
        EXPR   1SG=ERG   eat=1=NEG   cooked.rice-ECHO

        'Argh, I won't eat rice and stuff!'

        CH_CTW_SPC_POL_E

(93)   *e,*   *ŋa=i*   *waʔn=ŋʌ=lʌ,*   *budi-sudi !*
        EXPR   1SG=ERG   eat=1=NEG   wife-ECHO

        'Argh, I won't bring back any kind of wife at all!'

        CH_CTW_SPC_POL_E

(94)   *e,*   *ŋa*   *pahj=ŋʌ=lʌ,*   *bʰaw-kʌ-saw-kʌ !*
        EXPR   1SG   leave=1=NEG   husband-NMZ:LOC-ECHO-NMZ:LOC

        'Argh, I won't go to just any husband!'

        CH_CTW_SPC_POL_E

(95) *e,*   *ŋa=i*   *tɕoʔ-soʔ*   *tɕjew=ŋʌ=lʌ !*
      EXPR   1SG=ERG   child-ECHO   see_find=1=NEG

'Argh, I won't have any kind of children whatsoever!'

CH_CTW_SPC_POL_E

(96) *i*   *kwi=kaj-swi=kaj*   *ahm*   *bʌj=sa*   *pʌɾ=ʌ=lʌ.*
      PROX   dog=DAT-ECHO=DAT   rice   give=NMZ1   have.to_fall=LN=NEG

'No need to give rice to this dog or whatever it is!'

CH_CTW_SPC_POL_E

(97) *i*   *kwi-swi=kaj*   *ahm*   *bʌj=sa*   *pʌɾ=ʌ=lʌ.*
      PROX   dog-ECHO=DAT   rice   give=NMZ1   have.to_fall=LN=NEG

'No need to give rice to this dog or whatever it is!'

CH_CTW_SPC_POL_E

3.3.3. Denonymic suffix

The derivational denonymic suffix *-mʌj* is a productive morpheme used to create nouns that refer to the inhabitants of a locality. The nouns to which the morpheme *-mʌj* attaches mainly denote locations expressed in particular by toponyms, i.e., place names, as illustrated in (98) and (99). It is also found attached to the relator nouns *tjaw-* and *kaʔm-* which express higher and lower locations, respectively, resulting in the forms *tjaw-mʌj* 'people of upwards' and *kaʔm-mʌj* 'people of downwards.'

Note that this type of derived denonym always carries plural semantics. It likely originally developed from a noun meaning 'people' in a noun-noun type determinative or descriptive compound where it functioned as a semantic head, before being reanalyzed as a derivational denonymic suffix.

(98) *Dhirang=ko*   *bʰasa=paj,*   *kʰwej ?!*
      Dhirang=GEN   language=DIS   PART

      Dambarang=ko=l=o   *le*   *dʌhja*
      Dambarang=GEN=COP=NMZ:REL   COP   maybe

      Dhirang-mʌj=ko   *bʰasa=ma,*   *ma=ba ?*
      Dhirang-DEN=GEN   language=ADD   PART=PART

      'The language of Dhirang, what do I know? It is probably like that of Dambarang,
      the language of the people of Dhirang too, right?'
      CH_MKW_CMC_BC_SIL_120619_2_Conversation_Dhirang

(99) *ŋi=ko*   *Polkim-mʌj=i*   *kʌdʌw*   *suk=ka=n=i.*
      1PL=GEN   Polkim-DEN=ERG   millet   plant=2/3.PST=DIR/TR=PL

      'The people of our Polkim planted millet.'
      CH_CTW_SP_POL_051021_E

Denonyms derived with *-mʌj* are found used as clan names, as shown in (100)
with Yomtimai (< *Yomti-mʌj*) and Sarlingmai (< *Sarling-mʌj*), recalling the location from
where the people likely originate.

(100) *ʌ,*   *Yomtimai,*   *Sarlingmai,*   *Daruwai,*   *Nakdaruwai,*
       uh   Yomtimai   Sarlingmai   Daruwai   Nakdaruwai

       i-mʌj=ko   *tes-ɾaŋ*   *beglʌj*   *beglʌj*   *sjaw=na.*
       PROX-PL.H=GEN   grave   different   different   become=NPST

       'The Yomtimai, the Sarlingmai, the Durawai, the Nakdaruwai, their graves are
       different for each of them.'
       CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

The morpheme *-mʌj* is found in RAP-13 following the locational nominalizer *=kʌ*
which derives the head of a noun-phrase that expresses the usual living location of a
referent (§ 3.4.2.2, § 3.4.4.2). The resulting construction denotes the people who live at
the place of the referent, i.e., usually the referent's family members. This is illustrated in

(101) and (102). The fact that the morpheme *-mʌj* occurs in this position shows further evidence for its nominal origin as a free morpheme meaning 'people.' Indeed, its position in this construction is shared with genitively modified heads of noun-phrases, relator nouns and postpositions (§ 3.4.2.3).

(101) *ŋa=kʌ-mʌj=i*   *ahm*   * dʑe=ka=n=i*   *ja ?*
       1SG=NMZ:LOC-DEN=ERG   porridge   eat=2/3.PST=DIR/TR=PL   or

       'Did my people eat?' //in the case of a food distribution in a village//
       CH_CTW_SP_POL_051021_E

(102) *ŋa=kʌ-mʌj=kaj*   *to=nu*   *da !*
       1SG=NMZ:LOC-DEN=DAT   tell_say=2PL.IMP.TR   PART

       'Tell it to my family, hey!'
       CH_CTW_SP_POL_051021_E

The morpheme *-mʌj* can also occur with nouns that denote a category or kind of people holding a higher-ranked social status or position. This construction is mainly found in the speech of shamans, shamanic chants and narratives.

For instance, *-mʌj* is found on the noun *pande* 'shaman,' such as *pande-mʌj* 'shamans, shaman people,' as in (103), and with the compound *bʌn-sʌkti* 'jungle power,' such as *bʌn-sʌkti-mʌj* 'those having the power of the jungle,' which may refer to higher-ranked shamans or gurus who appear to the shaman in dreams when passing on to them their knowledge, as in (104). In a shamanic chant, it is found with the nouns *ɾani* 'queen' and *ɾadʑa* 'king,' illustrated in (105). This construction or "socionymic" derivational function possibly developed through metonymic analogy between people coming from a location in space and a location in society.

The morpheme *-mʌj* is cognate with the Bhujel morpheme *-may* [mʌj] presented as a honorific plural marker attached to nouns (Regmi 2007: 150–151). These nouns denote any type of people, unlike the Chepang denonymic morpheme *-mʌj*. This function is likely a Bhujel innovation motivated by the derivational property of *-mʌj* to denote people's higher social status or position.

(103) ʌ=ŋ=sʌj   o   pande-mʌj   ʌrko   pande
      DIST=LOC1=ABL   DIST   shaman-DEN   other shaman

sat=na=n=i,   ŋa=i=tɕʌj,   ŋa=i=tɕʌj   sat=ŋʌ=lʌ.
kill=NPST=DIR/TR=PL 1SG=ERG=DIS   1SG=ERG=DIS kill=1=NEG

'From there, (it happens that) those shamans kill another shaman, as for me, I don't kill any.'

CH_CTW_SMC_JIM_101920_1_Being_Shaman

(104) bʌn-sʌkti-mʌj   Satchyurya-mʌj   Churya=haŋ=le
      jungle-power-DEN   Satchyurya-DEN   Chyurya=LOC1=DIS

aʔl=o   ŋa=kaj,   kja !
take.away=PERF   1SG=DAT   PART

'The people of the power of the jungle, the people of Satchyurya, have taken me to Chyurya, imagine!'

CH_CTW_BBC_PID_011520_1_Being_Shaman

(105) rani-mʌj   niʔ=ma=tʌ   dʌj,
      queen-DEN   laugh=NEG=NMZ:ADV1   PART

rad͡za-mʌj   niʔ=ma=tʌ   dʌj !
king-DEN   laugh=NEG=NMZ:ADV1   PART

'Queens, without laughing hey, Kings, without laughing hey!'

CH_MKW_KMC_SK_082918_5_Shaman_Song

The morpheme =mʌj also functions as a plural nominalizer directly encliticizing to verbal roots, as in (106). Such a nominalized form is used as an honorific imperative, as in (107). This type of nominalized construction is also found in the speech of shamans as well as in the Lothar varieties of RAP-11 and RAP-13.

Like when -mʌj attaches to nouns, this construction may as well originate from a determinative or descriptive verb-noun compound where -mʌj functioned as modified semantic head with the meaning of 'people.'

Bhujel (Regmi 2012: 194) features a cognate nominalized construction formed with -*maɣ* [mʌj].

(106) *tɕʰak-ʣaʔ=paj*   *si=mʌj=ko*   *hʌ,*
      dead.tiger.spirit=DIS   die=NMZ:H=GEN   COP

      sjak=mʌj=ko=tɕʌj   *sjak-ʣaʔ.*
      be.alive=NMZ:H=GEN=DIS   alive.tiger.spirit

      'The Chak-jaˈ spirit is that of the dead, and that of the alive is Syak-jaˈ.'
      CH_CTW_SMC_JIM_101920_2_Syak_Ja'

(107) *baba=kʰe=tʌ*   *"pʌjsa*   *lʰok=ti*   *bʌj=mʌj,"*
      father=DIS=DIS   money   send=SEQ1 give=IMP.HON

      to=o   *kʰe=to,*
      tell_say=PERF   COP=REM.PST

      '"Send and give money," was said (I said) to (their) father,'
      CH_MKW_SC_SIL_010220_1_Life

The morpheme **mʌj* 'people' can possibly be reconstructed back to PCB. It has cognates in Bhujel (Regmi 2012: 194) and other TH languages where it occurs as a nominal root or as suffix, such as: *³mi* 'people, person' in Tamang (Mazaudon 1978; 2003: 308), *mih* 'person' in Gurung (Glover 1972), *mi* 'person' in Newah (Genetti 2007), *mi* 'person' in Bunan (Widmer 2014), *mju* 'person' in Apatani (Abraham 1985), *miːn* 'person' in Dumi (Driem 1993), *maʔi* 'person' in Belhare (Bickel 2003: 548). It is reconstructed **r-mi(ɣ)* (Chou 1972; Benedict 1972), **r-mi(ɣ)-n* (Matisoff 2003) with the meaning of 'person, people.'

Finally, although they could be historically related, it is worth noticing that the denonymic derivational morpheme -*mʌj* is different from the morpheme -*mi*, which features an allomorph -*mʌj*, used in the formation of 3rd person plural pronouns referring to humans (§ 4.1). It attaches to the 3rd person or demonstrative pronouns *i* PROX, *o* DIST and *u* REM. While the two allomorphs are attested in free variation, -*mi* (236 tokens) is more frequent in our corpus than -*mʌj* (32 tokens).

Instances of *-mi* and *-mʌj* are illustrated in (108) to (114). In (114), both forms occur in a single sentence. The sound change from /i/ to /ʌj/ is also attested in the same environment with other roots,[42] conditioned by the presence of a bilabial nasal consonant in initial position and the absence of final consonant. It is therefore likely that the original form of the 3ʳᵈ person plural was *-mi*, since this sound change is not attested with the derivational denonymic suffix *-mʌj* and the nominal bound root *mʌj-* 'banana.'

(108)   *nʌ*   *i-mi=kaj*   *dʑʌŋʌl=haŋ*   *mʌ=sa*   *sjaw=lʌ,*
        neither   PROX-PL.H=DAT   jungle=LOC1   COP=NMZ1   become=NEG

        dukʰʌ   *tɕjew=ka=n=i*
        sorrow   see_find=2/3.PST=DIR/TR=PL

        'Neither they are allowed to live in the jungle, they found sorrow.'
        CH_CTW_BBC_PID_011520_8_Chepang_Raute

(109)   *ŋaʔ*   *sat=ti*   *tɕʌn*   *boŋ=ti*   *o-mi*
        fish   kill=SEQ1   crab   look.for=SEQ1   DIST-PL.H

        dʑe=na=n=i,   *kʰe=lʌ ?*
        eat=NPST=DIR/TR=PL   COP=NEG

        'They feed on fishing, looking for crabs, isn't it?'
        CH_CTW_BBC_PID_011520_9_Chepang_Raute

(110)   *lʌj=ko*   *dʰʌrmʌ*   *lʌj=ko*   *sʌskar*
        SLF.INTS=GEN   religion   SLF.INTS=GEN   ritual

        lʌj=ko   *sʌskriti=kaj*   *o-mʌj*   *mʰeʔ=ka=n=i.*
        SLF.INTS=GEN   culture=DAT   DIST-PL.H   forget=2/3.PST=DIR/TR=PL

        'They forgot their own religion, their own rituals, their own culture.'
        CH_CTW_ABC_JIM_101920_1_Language and Culture

---

[42] Such a sound change is attested with the roots *mi-* ~ *mʌj-* 'be small' and *mri-* ~ *mrʌj-* 'mix millet and corn flour to make porridge.'

(111) ʌni   o-mi=kaj=ma   pɾʌstʌ   tʰaha   na=lʌ,
      then   DIST-PL.H=DAT=ADD   abundant   knowledge   COP=NEG

      dʌjti   bʰʌne   lipi   na=lʌ.
      why   say.NMZ   writing   COP=NEG

'And they also don't know (the history of the Chepangs) well because there is no writing system.'

CH_MKW_SCBKC_SIL_081918_2_Chepang_king

(112) u-mi=ko   bʰasa=le   beklʌj   sjaw=na.
      REM-PL.H=GEN   language=DIS   different   become=NPST

'Their language is different.'

CH_CTW_BBC_PID_011520_7_Chepang_Raute

(113) u-mʌj   tɕiʔ=ma=l=o   kuɾa,
      REM-PL.H   know=NEG=COP=NMZ:REL   language

      ŋi=ko   kuɾa   sek=ti,
      1PL=GEN   language   pick.up=SEQ1

      u-mʌj=ma   tɕiʔ=pa=n=i   bʌɾu !
      REM-PL.H=ADD   know=OPT=DIR/TR=PL   instead

'They, rather than a language that one doesn't know, having learned our language, may they too know it instead!'

CH_CTW_JBC_BHR_111720_3_Agreement

(114) o-mi=ko,   o-mʌj=ko,   kʰe=lʌ,   o   bela=kʰa
      DIST-PL.H=GEN   DIST-PL.H=GEN   COP=NEG   DIST   time=LOC2

      jam   bʰʌkʰʌɾ   suk=ti   mu=o   kʰe=to.
      rice   just   plant=SEQ1   COP=PERF   COP=REM.PST

'Their, their, no, at that time they were just planting (their) rice.'

CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

3.3.4. Other nominal formations (non-productive)

Non-productive suffixes are also found to form nominal stems. A suffixing morpheme -*si* attaches to bound roots that denote species of trees (§ 3.3.4.1). Another set consists of two suffixing morphemes -*tɕa* and -*tɕoŋ* attached to a restricted number of nouns to mean a 'pair of' and a 'group of,' respectively (§ 3.3.4.2). These morphemes are not productive and attach to a closed set of lexical bound roots.

3.3.4.1.　　Suffix -*si*

The suffix -*si* only occurs with native bound roots denoting species of trees. There are a few examples where the suffix -*si* denotes species that grow as a vine, or a shrub.

Some illustrative examples are given in Table 97.

Table 97. Suffix -*si*

N-**si**	**meaning**
bʰaɾ-si	'Indian chesnut tree (male)
bun-si	'Ougeinia oojeinense tree'
daʔɾ-si	'Indian laurel tree'
de-si	'chayote squash vine'
haɾ-si	'Nepalese alder tree'
jʌ~jo-si	'Indian butter tree'
kok-si	'drooping fig tree'
lek-si	'flame tree'
mʌj-si	'banana stem/tree'
ɾak-si	'pine tree'
tak-si	'mango tree'
tiŋ-si	'cashew tree'
tup-si	'bastard myrobalan tree'
tʌw-si	'Indian gooseberry shrub'

As seen in § 3.3.1.1.1, a bound root suffixed with *-si* is often followed by the free root *siŋ* 'tree, wood' forming a compound. This is illustrated in (115). There are very few examples of a bound root suffixed with *-si* occurring in absence of *siŋ*, as in (116).

(115) *gʌ=tɕuk=ko*   *naŋ=i*   *mʌhj=te=na=u*
      INT=QTY=GEN   2SG=ERG   cite_invoke=2=NPST=3O/DIR

      u=tɕuk=ko   *lek-si-siŋ*   *braw=na.*
      ANA=QTY=GEN   flame-tree-tree   be.big=NPST

      'The more you invoke its name, the more the flame tree grows big.'
      CH_CTW_JMC_PYK_101920_Cing_Lan

(116) *sat=jakbʌti,*   *jat=dʑjo*   *tak-si=taŋ*
      kill=SEQ2   one=CL1   mango-tree=ATT

      suk=ka=n=i.
      plant=2/3.PST=DIR/TR=PL

      'After having killed (the king), it is said that they planted a mango tree.'
      CH_CTW_SBC_BBC_GUN_012120_Chepang Kings

The morpheme *-si* comes from the root *siŋ* 'tree, wood,' and its combination with bound roots specifying the species probably reflect a compound from an earlier stage, later renewed with the combination of the root *siŋ*, i.e., to form a new compound, such as *rak-si-siŋ* 'pine tree.'

Besides combining with *siŋ* 'tree,' a bound root suffixed with *-si* can also combine with other free roots. The root *raŋ* 'field' modified by a bound root suffixed with *-si* forms compounds often used as toponyms. If the name of a place refers to a type of tree, it is likely that this tree was growing in the area, as for the name of the village of Maisirang (RAK-8) formed with *mʌj-si* 'banana stem/tree' modifying *raŋ* 'field' to mean 'field of banana trees,' although the most common pronunciation of the name of this village is [maj.si.ɾãŋ] and not [mʌj.si.ɾãŋ]. The free roots *loʔ* 'leaf' and *ro* 'flower' can also be modified by a bound root suffixed by *-si* to form compounds, as in *jʌ-~jo-si-loʔ* 'Indian butter tree leaf' or *jʌ-~jo-si-ɾo* 'Indian butter tree flower.'

As seen in § 3.3.1.1.1, the bound root denoting the species can also combine directly with *saj* 'fruit, seed' without the presence of *-si* to form a compound like with *jʌ-~jo-saj* 'Indian butter fruit/seed.' While the absence of *-si* is rare in combination with other roots, it is also attested with *ro* 'flower,' or *sati* 'oil, butter,' as illustrated in (117) and (118), respectively.

(117) *djah   jʌ-ro   oɾ=na,*
        now   Indian.butter-flower   bloom=NPST,

        'Now, the Indian butter tree flowers bloom,'
        CH_MKW_RC_JMC_SIL_120119_Conversation

(118) *ʌ,   jʌ-sati=kʰa   kʰaŋ=ti,*
        uh,   Indian.butter-oil=LOC2   cook=SEQ1

        'Uh, having cooked (the bats) in Indian butter tree oil,'
        CH_MKW_PNC_SIL_081818_2_Bat_hunting

The absence of *-si* in combination with other roots than *saj* 'fruit, seed,' has so far only been attested in RAK-6, and specifically with the bound root *jʌ-~jo-* 'Indian butter tree,' an area where the Indian butter tree has traditionally been extremely precious, since it has been a nutritional source as well as a source of income: its fruits are eaten; its fruits' pits are pressed to make oil to cook food or apply as ointment; what is left from the pits in this process is used as a poison to catch fish; and its flowers feed the bats that are hunted and the bees that produce honey. The absence of *-si* may be triggered by frequency, or by analogy through contact with Nepali, since borrowed species of flowers for instance are directly followed by *ro* 'flower,' as in *tɕʌmeli ro* 'jasmine flower.'

Some examples of species of trees expressed without *-si* may be found in shamanic chants. Beside the fact that these forms may be older since they are found in shamanic chants, and thus reflect a possible earlier stage of compound where *siŋ* has not yet become *-si*, the absence of *-si* in this case could also be explained by the fact that the trees the shaman refers to here are deities. This is illustrated in (119).

(119) *haha hoho siŋ=taŋ,   siŋ,   wiʔ   waŋ=o   siŋ,*
       haha hoho tree=ATT   tree   blood come=NMZ:REL   tree

       bʰam bun   ja
       white Ougeinia.oojeinense_tree   or

       jaɾ   bun   siŋ ?
       yellow   ougeinia.oojeinense_tree   tree

'The tree, the tree, the tree that bled, (is it) the white Ougeinia oojeinense or

the yellow Ougeinia oojeinense tree?'

CH_MKW_BLC_SIL_113019_1_Shaman_chant

3.3.4.2.   Suffixes *-tɕa* and *-tɕoŋ*

The morphemes *-tɕa* and *-tɕoŋ* are not productive but were used to derive nouns to refer to a pair and a group. All attested examples are reported in Table 98.

Most of the nouns derived with *-tɕa* represent a pair of two referents. These nominals are built on a nominal root denoting a single referent that semantically acts like the representative of the pair including a "younger" or "smaller" version of them. For instance, *ma* 'mother' with *-tɕa* will mean a 'mother and her child (daughter or son),' *pa-* 'father' with *-tɕa* a 'father and his son,' or *naʔ* and *-tɕa* an 'elder sister and her younger sister,' or else *aj* 'mother-in-law' and *-tɕa* a 'mother-in-law and her daughter-in-law.'

In some Chepang varieties, the forms *ama-tɕa* and *baba-tɕa*, using the Nepali borrowings of *ama* 'mother' and *baba* 'father,' replace respectively *ma-tɕa* and *pa-tɕa*. In a variety of MAN-4 (han), the form *ma-tɕa* may also convey the meaning of 'wife.'

The nouns suffixed with *-tɕa* are always introduced with the numeral *nis* 'two' (§ 3.4.3.1). This does not mean that the pair is doubled, but simply determined as being two referents, as shown in (120).

(120) - *gʌhʌt,*   *niŋ=dʑi=ko*   *pewa*   *ja ?*
       vetch   2=DU=GEN   dowry   or

       - *kʰe=lʌ,*   *nis*   *puʔ-tɕa=ko.*
       COP=NEG   two   elder.brother-pair=GEN

       '- Is the vetch the property of the two of you?

       - No, it's that of my elder and younger brother.'

       CH_MKW_RC_JMC_SIL_120119_Conversation

A derived noun with -*tɕa* can also be used to address the two referents, similarly to when a 2nd person pronoun is present in an imperative construction to soften a command, making it a little less directive. This is illustrated in (121).

(121) *nis*   *aj-tɕa=i=ma*   *kʰaŋ=dʑu !*
       two   mother-in-law-pair=ERG=ADD   cook=2DU.IMP.TR

       'You two mother-in-law and daughter-in-law cook as well!'

       CH_MKW_DBC_MAI_2_020320_Newa Dung

Only two examples are found with the suffix -*tɕoŋ.* This form is used to refer to a group. Like -*tɕa*, it denotes a representative referent and its "younger versions," such as *puʔ-tɕoŋ* 'elder and younger brothers.'

An additional stem formed with *pʌjk* and -*tɕoŋ* is present in Caughley's dictionary (2016), and defined as 'husband and wives' in the case of a polygamous marriage, but this form is not recognized by the speakers I have been working with. It is also possible that it was borrowed from Nepali, i.e., पोइ <poi> 'husband.'

Table 98. Suffixes *-tɕa* and *-tɕoŋ*

N-*tɕa*	aj-tɕa	aj 'mother-in-law'	'mother-in-law and daughter-in-law'
	kwa-tɕa	(<N.) *khwāmit* 'bound friend' (arch.)	'two bound friends'
	kʌma-tɕa	kʌma 'daughter-in-law'	'two daughter-in-laws'
	ma-tɕa	ma 'mother'	'mother and child'
	naʔ-tɕa	naʔ 'elder sister' (arch.)	'elder and younger sisters'
	pa-tɕa	pa 'father'	'father and son'
	puʔ-tɕa	puʔ 'elder brother' (arch.)	'elder and younger brothers'
	pʌjk-tɕa	(<N.) *poi* 'husband' ?	'husband and wife'
N-*tɕoŋ*	naʔ-tɕoŋ	naʔ 'elder sister' (arch.)	'sisters; elder and younger sisters'
	puʔ-tɕoŋ	puʔ 'elder brother' (arch.)	'brothers; elder and younger brothers'

The suffix *-tɕa* could be a reflex of the PTH form **tsa~*za* reconstructed for 'child' (Benedict 1972). A diminutive morpheme *-ca* attaches to nominal roots that refer to animals in Newah, deriving nouns that express animal's offspring (Genetti 2007: 93). A collective morpheme *-tso* is attested in Hayu with kinship terms (Michailovsky 1988: 205). Another possible cognate is found in Khamci (Pons, data), with the form *pa-tɕa* used to refer the paternal lineage.

3.3.5. Traditional first or given names

The Chepangs traditionally gave their children first or given names that can be considered affectionate nicknames, although they may not be flattering. These names are often native to Chepang or of Chepang creation, when derived from or based on a borrowing from Nepali. Chepang first names or nicknames are generally given during childhood, and sometimes later in life. This is not a custom specific to the Chepang people but present all over the Himalayan region, in TH and IA languages.

First names based on Nepali have also been given to Chepang children for a long time, even before the existence of identity cards, which goes back to 1978 (BS 2035). The Chepangs follow the Hindu custom of naming their children either on their 5th, 7th, or 9th day. It is difficult to grasp when the Chepangs started to give their children first names

borrowed from Nepali. Indeed, little is known about Chepang traditional spiritual beliefs (Varya 1971; Gurung 1987; Riboli 2000), and the influence that Hinduism had on the Chepangs or other indigenous communities of Nepal (Hitan Magar 1992; Lecomte-Tilouine 2010), is likely tied to the presence of Brahmin, Chhetri and Kami communities in Chepang villages. Some first names have also been borrowed from Tamang, since there is a long history of contact between the two communities (Riboli 2000).

People may have both a first name (or nickname) in Chepang and a first name in Nepali, or only have one or the other. In cases of having both types of first name, the Chepang one is usually socially used over the Nepali one. In some cases, the Nepali first name may not even be recognized by the person's peers or people living in the same village, being merely used for administrative purposes. Most of the first names in Chepang that were collected were of elders or people who died. New generations are not given such names anymore.

These traditional Chepang first names or nicknames are formed through a single root, or compounds, or else verbs that are nominalized.

They can either be gender neutral, or have morphemes dedicated to marking gender (feminine, masculine). The morphemes attested to mark gender are all of IA origin (Masica 1991: 218–222), i.e., Nepali: *-i, -ni* for feminine, and *-a, -ja*, for masculine.

There is a small set of words showing gender nominal stem alternation through non-concatenative morphology, a process that is not attested elsewhere. By non-concatenative, I mean that the stem morphonology is not expressed through a segmentable string of morphemes. In this case, different vowel qualities are dedicated to marking gender: the low central vowel /a/ marks masculine gender, as in *tɕapana* (♂) 'born with a flat and big nose' and *tɕapaɾ* (♂) 'born with a flat and small nose,' while the close-mid front vowel /e/ is used for their feminine gender counterparts, i.e., *tɕepene* (♀) and *tɕepeɾ* (♀). It is possible that these forms were built on the root *tɕep-* चेप् <cep> of the adjective चेप्टो <cepṭo> 'flat' borrowed from Nepali. Another example of this kind is a name given to a person who has a twin of a different gender: *tapala* (♂) and *tepele* (♀).

Semantically, Chepang first names or nicknames convey physical or behavioral traits associated with the person, such as: protruding navel, flat nose, curly hair, thin body, having a small head, being irritating, uncontrollable, and so on. Relating to a person's appearance and behavior translates as well through using traditional names that include animal and bird names.

First names can also be chosen according to the day the child is born. It is likely that these are relatively old, since the names of the days were borrowed from Nepali and adapted to Chepang phonology when necessary. For instance, the two-syllable root सोम्बार् <sombār> 'Monday,' which became *sub-i~ʌ* (suffixed with gender marking morphology), was adapted to fit a CV.CV structure closer to Chepang native syllable structure.

The word *wa* 'bird' is also found combined with the name of the days, such as *ajtʌ-wa* 'Sunday-bird.'

Some polysyllabic first names of possible compound origin, may be morphologically opaque. This is the case with *kɾiduŋ* 'thick and curly (of hair),' which is given to children with thick and curly hair, and which functions as an adjective to qualify hair. Its two syllables do not correspond to any recognizable morpheme, suggesting a possible formation based on an ideophone. Another example is *kjaŋkul*, which is given to children born with long arms and legs. This term is not found anywhere else, and none of its syllables corresponds to any known morpheme.

Finally, first names or nicknames can also be chosen for their esthetic or poetic aspect, with or without conveying any specific meaning, such as *jʌ-saj* 'Indian Butter tree fruit,' or else *takula*, *takuli*, and *tantula*, which do not express any physical of behavioral trait.

In Table 99, I present examples of Chepang first names traditionally given to children.

Table 99. Chepang traditional first names

first name	gender	formation	meaning/origin	physical/behavioral traits
juʔ	F/M	N	'mouse'	born with fronting lips
tɕjusu	F/M	N	'shrew'	born with fronting lips
ʌnu	M	N	(<N.) *anu* 'plough handle'	uncontrollable
heɾ-dʑaʔ	F/M	V.INTR-N COMP.	'small tiger *sp.*'	born with rolls of fat on belly
sak~sag-gu	F/M	N-N COMP.	'be irritating, pungent'-'yam (*Dioscoreaceae*)'	annoying, irritating
kjaŋkul-i	F	?-FEM	'?'	born with long arms and legs
kjaŋkul-a	M	?-MASC	'?'	born with long arms and legs
bu-tɕa	M	N-MASC	'boarlet/piglet m.'	born crying like a boarlet/piglet
bʰu-ni	F	N-FEM	'boarlet/piglet f.'	born crying like a boarlet/piglet
pit-ja	M	N-MASC	*pitjak* 'tongs'	born very thin, no strong muscle
pit-i	F	N-FEM	*pitjak* 'tongs'	born very thin, no strong muscle
kot-o	F/M	V.INTR-NMZ	'protrude (of navel)'	born with protruding navel
goʔŋ-o	F/M	V.INTR-NMZ	'be twisted'	born with twisted mouth
dʑʰjaʔn-a	M	V.INTR.-MASC	'be loose (of skin of arm and testicles)'	born with loose testicles' skin
kɾiduŋ	F/M	ADJ	'thick and curly (of hair)'	born with thick curly hair
tɕepene	F	ADJ.FEM	(<N.) <cepṭo> 'flat'	born with a flat and big nose
tɕapana	M	ADJ.MASC	(<N.) <cepṭo> 'flat'	born with a flat and big nose
tepele	F	?	'?'	female born with a male twin
tapala	M	?	'?'	male born with a female twin
tɕepeɾ	F	ADJ.FEM	(<N.) <cepṭo> 'flat'	born with a flat and small nose
tɕapaɾ	M	ADJ.MASC	(<N.) <cepṭo> 'flat'	born with a flat and small nose
tomplik-wa	F/M	ONO-N COMP.	'common tailorbird'	born with a small body
momɾiŋ-wa	F/M	?-N COMP.	'common tailorbird'	born with a small body
tɕene-wa	F/M	?-N COMP.	'?' (<N.) <lāca>	born crying like this bird
kutuɾuk-wa	F/M	ONO-N	'?'	born with a big head

first name	gender	formation	meaning/origin	week days-based
ajtʌ lal	M	N-PR. N	(<N.) <āitabār> 'Sunday' (<N.) <lāl> 'Lāl'	born on Sunday
ajtʌ-wa	M	N-N	(<N.) <āitabār> 'Sunday' wa 'bird'	born on Sunday
sub-i maja	F	N-FEM N	(<N.) <sombār> 'Monday' (<N.) <māyā> 'Māyā'	born on Monday
sub-ʌ lal	M	N-MASC N	(<N.) <sombār> 'Monday' (<N.) <lāl> 'Lāl'	born on Monday
sub-i wa	F	N-FEM N	(<N.) <sombār> 'Monday' wa 'bird'	born on Monday
moŋʌl-i	F	N-FEM	(<N.) <maṅgalbār> 'Tuesday'	born on Tuesday
moŋʌl-ja	M	N-MASC	(<N.) <maṅgalbār> 'Tuesday'	born on Tuesday
biʃu	M	N	(<N.) bihibār 'Thursday'	born on Thursday

bihi-ni	F	N-FEM	(<N.) bihibār 'Thursday'		born on Thursday

first name	gender	formation	origin	source language
pʌbitrʌ	F	PR. N	borrowed	Nepali
santoʃ	M	PR. N	borrowed	Nepali
toksoɾ	M	PR. N	borrowed	Tamang
kʌrnja	M	PR. N	borrowed	Tamang

3.3.6. Kinship terms

In this section, I describe the morphological formation of Chepang kinship terms (§ 3.3.6.1). I list and discuss all of them, grouping them by generation, with special attention given to the way people use these terms (§ 3.3.6.2). I briefly present traditional and modern marriage practices, in addition to the favored matrilateral cross-cousin marriage still practiced by the Chepang communities (§ 3.3.6.3). Finally, I report cognate kinship terminology attested in other TH languages (§ 3.3.6.4).

3.3.6.1. Kinship terms morphology

In this section, I describe the morphological formation of Chepang kinship terms. I give some observation on gender distinction (§ 3.3.6.1.1), show that compounding remains one of the most attested formation process for kinship terms (§ 3.3.6.1.2), and briefly discuss borrowings from Nepali with regard to adding new terms, replacing native terminology, or else undergoing semantic shift (§ 3.3.6.1.3).

3.3.6.1.1. Gender distinction

While a fair amount of Chepang native kinship terms have been preserved, a lot have been borrowed from Nepali, and with them, gender distinction, using forms of IA origin.

In borrowed kinship terminology, the vowel endings /i/ and /a/ often characterize feminine and masculine gender, respectively, as with: *sali* (<N.) 'wife's younger sister,'

and *sala* (<N.) 'wife's younger brother.' The Nepali morpheme /ni/ (which can be considered a productive morpheme in other domains of the lexicon) is also associated with feminine gender, as with: *bʰʌdʌjni* (<N.) 'brother's daughter,' and *bʰʌdʌj* (<N.) 'bother's son.'

Chepang does not natively encode gender differences through overt morphology. With certain kinship terms however, initial bilabial nasals and initial bilabial stops are found to be associated to feminine and masculine gender, respectively, such as: *mʌh* 'mother's elder sister; father's elder brother's wife' and *bʌh* 'father/husband's elder brother, father/mother's elder sister's husband,' or else *moʔm* 'granddaughter' and *ploʔm* 'grandson.'

3.3.6.1.2. Compounding

Determinative or descriptive compounds (§ 3.3.1.1) are the main type of formation process attested with native kinship terms. As in other lexical domains, they can either express an attributive or possessive relationship. In determinative or descriptive compound, the first root modifies the second root, i.e., semantic head of the compound.

In attributive determinative compounds, the characteristics associated to the modifying root are applied to the semantic head. For instance, *bʌh-puʔ* 'husband's elder sister's husband (arch.)' is formed with *puʔ* 'elder brother (arch.),' modified by *bʌh* 'father/husband's elder brother; father/mother's elder sister's husband,' which translates as an elder brother of the kind of an uncle or brother-in-law.

Possessive determinative compounds express a possessor as modifier of the semantic head. The compound *ʨoʔ-bʰaw* 'son-in-law' expresses a possessive relationship, referring to the child's husband with the parents as ego. In the compounds *hʌw-bʰaw* 'younger sister's husband' and *hʌw-kʌma* 'younger brother's wife,' the morpheme *hʌw* that relates to the sibling, functions as possessor of the spouse.

3.3.6.1.3.   Borrowing

Chepang has borrowed a fair amount of kinship terms from Nepali, such as single nominal roots or compounds, either directly borrowed, or combining both a native term and a Nepali loan.

In Figure 71, I present the percentage of kinship terms that can be characterized as native terms, native terms considered archaic and being replaced by Nepali loans, compounds formed through the combination of a native term and Nepali loan, Nepali loans bearing identical semantics, and Nepali loans having undergone a semantic change.

The percentage values of the kinship terms' origins are based on 141 types of ego's kin relationships (§ 3.3.6.2).

While Chepang uses native terms to describe 40% of its kin relationships, it features 36% of Nepali loans (including both those whose meaning remained and those whose meaning changed, and those whose native archaic form still exists but is being replaced in most varieties by Nepali loans), and 24% of compounds that are formed mixing both Chepang native terms and Nepali loans.

Figure 71. Origin of Chepang kinship terms
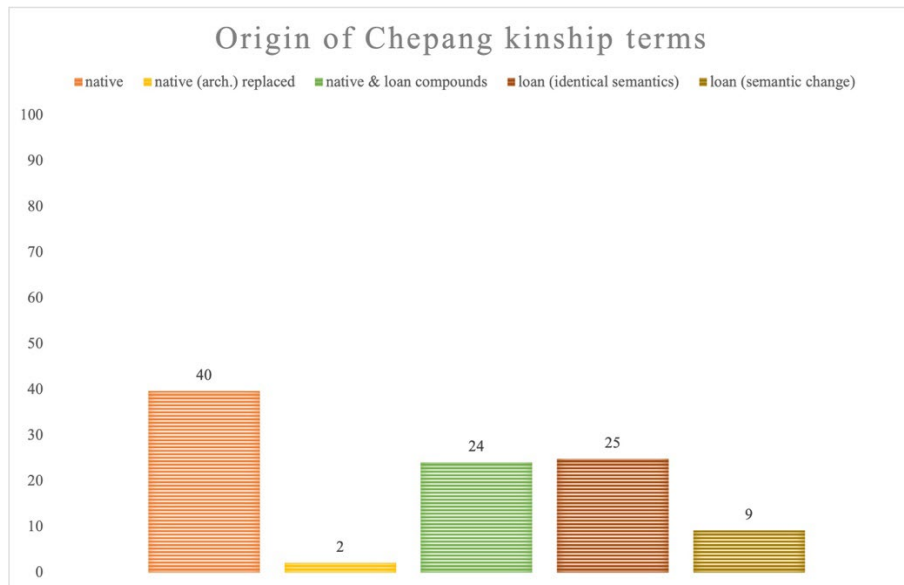
Since it concerns kinship terms, although this cannot be completely certain, one can suppose that some borrowings have replaced native words, rather than having been a source for a non-existing term. For example, the kinship terms *sali* (<N.) 'wife's younger sister' and *sala* (<N) 'wife's younger brother' were borrowed, while native terms referring to the husband's younger siblings exist, such as: *mom* 'husband's younger sister,' and *paŋ* 'husband's younger brother.'

There are kinship terms that were borrowed from Nepali and that underwent a semantic shift in Chepang. For example, the term *kaka* (<N.) means 'father's younger brother' in both Chepang and Nepali, but this term is also used to refer to the mother's sister's husband, which in Nepali is सानो-बा <sāno-bā>. This semantic change likely arose from clipping the original Chepang compound *kaka-ba*, used interchangeably with *kaka*.

Some native kinship terms have become archaic while still in use amongst elders or found in traditional stories. These are replaced by borrowed terms from Nepali, for example the native terms *naʔ* 'elder sister (arch.)' and *puʔ* 'elder brother (arch.)' are being replaced by *didi* (<N.) 'elder sister' and *dadzu* (<N.) 'elder brother.'

Some compounds have been borrowed as is from Nepali, such as *dzetʰi-sasu* (<N.) 'wife's elder brother's wife,' while others have combined a native and a borrowed root, like *mama-to* 'brother's wife's father' for the Nepali term मामा-ससुरा <māmā-sasurā> 'brother's wife's father.' The series of kinship terms that refers to the children through their position as siblings has been borrowed, combining the native term for 'son' and 'daughter' with their position in Nepali, such as: *kantɕʰa-tɕoʔ* 'sixth son' and *kantɕʰi-tɕoʔ-djaŋ* 'sixth daughter,' where *kantɕʰa* and *kantɕʰi* were borrowed from Nepali, meaning the youngest or the sixth child (since there is a maximum of six terms categorizing children's positions).

3.3.6.2.   Kinship terms by generation

This section presents Chepang kinship terms that relate to each generation. I specify whether the terminology was borrowed, and whether a kinship term is used as a term of address or a term of reference, or both.

The standard abbreviations traditionally used to describe kinship terms are reported in Table 100. For example, the abbreviation MeB would refer to the mother's elder brother.

In addition, the abbreviation ^N on a kinship term stands for 'Nepali' to signify that the term is of Nepali origin. Some kinship terms borrowed from Nepali underwent semantic shifts.

Most kinship terms are employed as both a term of address and a term of reference. These are noted ^{TAR}, i.e., 'term of address and reference.' When it is not the case, I use ^{TA} and ^{TR} to mean respectively 'term of address,' and 'term of reference.'

When a kinship term can occur as a term of address, the vocative morpheme =jʌ can be used as well, conveying a more respectful or soften form of address. Only kinship terms allow the presence of the vocative, not proper nouns.

Also, note that when it is socially acceptable for the speaker to address their interlocutor by their first name, the kinship term often follows the first name, as with: *Santosh bʰandza* 'Santosh HZS.' Only ego's own children, own younger siblings, and wife are usually directly addressed by their first name with no mention of the kin relationship.

Kinship terms related to ego's parents' generation are presented in Table 101. They include ego's parents, and ego's parents' siblings and spouses.

Kinship terms related to ego's generation are presented in Table 102. They include ego's siblings, egos' siblings' spouses, egos' spouses, egos' spouses' siblings, ego's parents' siblings' children, and ego's spouses' siblings.

Kinship terms related to ego's parents' parents' generation and beyond are presented in Table 103. They include ego's parents' parents, and ego's parents' siblings' spouses' parents.

Kinship terms related to ego's children and grandchildren's generations are presented in Table 104. They include ego's children and grandchildren, ego's siblings' children and grandchildren, ego's great-grandchildren and great-great-grandchildren, ego's spouses' siblings' children, and ego's parents' siblings' grandchildren.

Table 100. Standard abbreviations used for kinship terms

F	father	M	mother
B	brother	Z	sister
S	son	D	daughter
H	husband	W	wife
e	elder	y	younger
♂	male	♀	female

3.3.6.2.1. Ego's parents' generation

The kinship terms related to ego's parents' generation are presented in Table 101.

The forms *apa* 'father' and *ama* 'mother,' may not have the same origin. Both are found in other TH languages. The term *ama* 'mother' however, is also used in Nepali and can likely be considered of IA origin. The term *apa* 'father' has cognates in different TH languages of different clades, such as *pa* in Limbu (Eastern-Kiranti (Michailovsky 1994)) (Driem 1987), *papa* in Bahing (Western-Kiranti (Michailovsky 1994)) (Michailovsky 1989a), *apa* in Thangmi (Newaric (Turin 2004a)) (Turin 2004b), *a-pa* in Puroik (Kho-Bwa (Bodt & Lieberherr 2015; Lieberherr 2017; Post 2020)) (Sun & et al. 1991; Lieberherr 2017), and Damu (Tani (Sun 1994; Driem 2014b; Post 2020)) (Sun & et al. 1991). The initial vowel /a/ is analyzed as a possessive marker in some languages, such as Puroik and Damu, and is likely a reflex of the PTH nominal 3rd person possessive prefix **a-* (Benedict 1972: 121–123). If this prefix is cognate with the Chepang initial vowel /a/ of *apa* 'father,' it would be the only noun that would have retained it, morphologizing with its inherited PTH host root **pa* 'father' (Benedict 1972).

Note that *apa* is considered archaic in all studied varieties. It is replaced by *buwa*, *baba*, or *ba*, borrowed from Nepali.

The kinship terms of *tɕʰjama* (<N.) is dedicated to referring to the mother's younger sister and borrowed from Nepali, where it is used as "an address by a woman to a somewhat older woman" (Turner 1931: 204), while *mom* is also used to refer to ego's spouse's younger daughters (§ 3.3.6.2.2).

Finally, ego's parents' elder siblings and their spouses are referred to as *mʌh* (♀) and *bʌh* (♂). Only the father's elder sister is referred through the compound *mʌh-pʰupu*, combining the Chepang form *mʌh* and the Nepali form *pʰupu*, which by itself carries the same semantics as *mʌh*. Both *mʌh* and *bʌh* are also used to refer to ego's spouse's elder siblings (§ 3.3.6.2.2). Very few potential cognates are found in other TH languages for the terms *mʌh* (♀) and *bʌh* (♂) (§ 3.3.6.4).

Table 101. Ego's parents' generation

EGO'S				
parents				
F	*apa* (arch.)[TAR], *baba*[N|TAR], *ba*[N|TAR], *buwa*[N|TAR]	M	*ama*[N|TAR]	

parents' siblings		**parents' sibling's spouses**	
MB	*mama*[N|TAR], first name[TA]	MBW	*majdzu*[N|TAR], first name[TA]
MeZ	*mʌh*[N|TAR], first name[TA]	MeZH	*bʌh*[TAR], first name[TA]
MyZ	*tɕʰjama*[N|TAR], first name[TA]	MyZH	*kaka*[N]-*ba*[N|TAR], *kaka*[N|TAR], first name[TA]
FeB	*bʌh*[TAR], first name[TA]	FeBW	*mʌh*[TAR], first name[TA]
FyB	*kaka*[N|TAR], first name[TA]	FyBW	*mom*[TAR], first name[TA]
FeZ	*mʌh-pʰupu*[N|TAR], first name[TA]	FeZH	*bʌh*[TAR], first name[TA]
FyZ	*pʰupu*[N|TAR], first name[TA]	FyZH	*pusʌj*[N|TAR], first name[TA]

3.3.6.2.2. Ego's generation

The kinship terms related to ego's generation are presented in Table 102.

Ego's siblings are referred to with native terms, two of which are considered archaic in Lothar, Rapti, and Handikhola varieties, being replaced by Nepali borrowings: *naʔ* 'elder sister' and *puʔ* 'elder brother.'

The form *puʔ* 'elder brother' has cognates in a lot of TH languages from different clades, while fewer cognates are found for *naʔ* 'elder sister,' like *a-na* 'elder sister' in Yakkha (Eastern Kiranti) (Schackow 2015b).

The form *hʌw* 'younger brother' has cognates in Thangmi (Newaric) (Turin 2004b) and Meche (a.k.a. Boro, Bodo in India) (Bodo-Garo) (Weidert 1987). It is likely

that *hʌw-djaŋ* 'younger sister' be an innovation, a compound whose second element carries a female gender meaning also found in *tɕoʔ-djaŋ* 'daughter.'

The elder brother's wife is referred to as *kʌma*. When the ego is the parents, another term applies, *kʰon*. This latter is cognate with *me-kʰon* 'daughter-in-law' used in Magar (Centra-Himalayan) (Pons data).

Ego's parents' siblings' children (or ego's cousins) are referred to as siblings, i.e., as *hʌw* 'younger brother,' *dadʑu* 'elder brother,' *hʌw-djaŋ* 'younger sister,' and *didi* 'elder sister.'

Wives are referred to as *meɾu* 'wife,' a native term considered archaic in Lothar, Rapti, and Handikhola varieties, where it is replaced with Nepali borrowings. It is preserved in Manahari. The term *bʰaw* 'husband' is also native. It is used widely in addition to Nepali borrowings.

While a husband can call his wife by her name, a wife will completely avoid calling her husband by his name. If she needs to call him, she will go look for him or send someone to look for him.

Reference to spouses' elder siblings is analogous to that of parents' elder siblings: *bʌh* (♂) and *mʌh* (♀) for ego's husband's elder siblings, and *mʌh* (♀) for ego's wife's elder sister. The term *bʌh* (♂) is also used to refer to ego's father's elder sister's husband (§ 3.3.6.2.1).

Ego's husband's younger siblings are referred to as *paŋ* (♂) and *mom* (♀). The term *mom* (♀) also applies to ego's wife's younger sister and ego's father's younger brother's wife (§ 3.3.6.2.1).

Table 102. Ego's generation

EGO'S

siblings		sibling's spouses	
eB	*puʔ* (arch.)^TAR, *dadʑu*^N|TAR, first name^TA	eBW	*kʌma*^TAR, *bʰawdʑu*^N|TAR
yB	*hʌw*^TAR, first name^TA	yBW	*hʌw-kʌma*^TR, *kʌma*^TA, *buhaɾi*^N|TAR, first name^TA
eZ	*naʔ* (arch.)^TAR, *didi*^N|TAR, first name^TA	eZH	*bʰena*^N|TAR
yZ	*hʌw-djaŋ*^TAR, first name^TA	yZH	*hʌw-bʰaw*^TR, *hʌw-djaŋ-bʰaw*^TR, *hʌw-dʑw̃ʌj*^N|TR, first name^TA

		EGO (=PARENTS)	
		SW	*kʰon*^TAR, *dulʌj*^N|TA
		DH	*bʰandʑa*^N|TAR, *dʑw̃ʌj*^N|TAR, *babu*^N|TAR

parents' siblings' children			
MBD	*hʌw-djaŋ*^TAR, *didi*^N|TAR	MBS	*hʌw*^TAR, *dadʑu*^N|TAR
MZD	*hʌw-djaŋ*^TAR, *didi*^N|TAR	MZS	*hʌw*^TAR, *dadʑu*^N|TAR
FBD	*hʌw-djaŋ*^TAR, *didi*^N|TAR	FBS	*hʌw*^TAR, *dadʑu*^N|TAR
FZD	*hʌw-djaŋ*^TAR, *didi*^N|TAR	FZS	*hʌw*^TAR, *dadʑu*^N|TAR

spouses			
H	*bʰaw*^TR, *buda*^N|TR, *budo*^N|TR, *dula*^N|TR, not directly addressed by ego	W	*meɾu* (arch.)^TR, *budi*^N|TR, *dulʌj*^N|TR, first name^TA

spouses' siblings	
HeB	*bʌh*^TR, *dadʑu*^N|TA
HyB	*paŋ*^TR, *babu*^N|TA, first name^TA
HeZ	*mʌh*^TR, *mʌh-didi*^TR, *didi*^N|TA
HyZ	*mom*^TAR, *nani*^N|TAR, first name^TA
WeB	*dʑetʰan*^N|TAR, *dʑetʰan-puʔ* (arch.)^TR, *dadʑu*^N|TA, *daj*^N|TA
WyB	*sala*^N|TAR, *bʰaj*^N|TA, first name^TA
WeZ	*mʌh*^TR, *mʌh-didi*^TR, *didi*^N|TA, *dʑetʰi*^N-*sasu*^N|TAR
WyZ	*mom*^TAR, *sali*^N|TR, *bʌhjni*^N|TA (RAP-13), first name^TA

spouses' siblings' spouses	
HeBW	*dʑetʰajni*^N|TR, *didi*^N|TA
HyBW	*dewɾajni*^N|TR, *bʌhjni*^N|TA, first name^TA

HeZH   *bʌh* (arch.)[TR], *bʌh-puʔ* (arch.)[TR], *dadʑu*[N|TAR]

HyZH   *mom-bʰaw*[TR], *mom-dʑw̃ʌj*[N|TR], *babu*[N|TA], *bʰaj*[N|TA], first name[TA]

WeBW   *dʑetʰan-didi*[N|TR], *didi*[N|TA]

WyBW   *sala-ko budi*[N|TR], *hʌw-djaŋ*[TA], *bʌhjni*[N|TA], first name[TA]

WeZH   *mʌh-dadʑu*[N|TR], *dadʑu*[N|TA]

WyZH   *hʌw*[TA], *bʰaj*[N|TA], *mom-bʰaw*[TR], *mom-dʑw̃ʌj*[N|TR], *nʌnde*[N]-*bʰaj*[N|TR], first name[TA]

3.3.6.2.3.   Ego's parents' parents' generation and beyond

Kinship terms that correspond to ego's parents' parent's generation and beyond are presented in Table 103.

Ego's parents' parents, and ego's parents' parents' parents are differentiated by the compound elements *nunu~dudu* and *dʑu~dʑi* to refer to the mother's ancestors, while the element *had(ʌ)* is used to refer to the father's ancestors.

The maternal grandmother is referred to as *nunu~dudu-bʌdʑe* and the maternal grandfather as *nunu~dudu-badʑja.* The allomorph elements *nunu~dudu* carry the meaning of 'breast, milk' associated with the figure of the mother. These elements of the compound are of Nepali origin, although these specific compounds are not present in Nepali.

The paternal grandmother and grandfather are referred to as *had(ʌ)-bʌdʑe* and *had(ʌ)-bʌdʑja*, respectively. The term *had(ʌ)* is also borrowed from Nepali while not being used in such compounds in Nepali. It reflects the meaning of 'bone' associated to the paternal ancestors' lineage. This symbolic association is present in other TH (in Tibetan for example, the term རུས <rus> 'bone' also means 'lineage') and Nepali speaking communities, and likely of IA origin (Lecomte-Tilouine 1993).

Ego's spouses' parents are referred to as *aj* 'mother-in-law' and *to* 'father-in-law.' These terms are also used to refer to ego's siblings' spouses' parents.

Ego's in-laws' parents (or spouses' grandparents) are referred to through the use of the compounds *aj-bʌdʑe* 'spouse's grandmother' and *to-badʑja* 'spouse's grandfather,' literally 'mother-in-law type of grandmother' and 'father-in-law type of grandfather,'

respectively. These compounds merely reflect in-laws' gender but no difference in lineage.

Ego's siblings' spouses' parents are referred to as *bʌdʑe* (♀) and *badʑja* (♂).

Finally, the terms *sʌmdi* 'ego's children's father-in-law,' and *sʌmdini* 'ego's children's mother-in-law,' are borrowed from Nepali.

Table 103. Ego's parents' parents' generation and beyond

	EGO'S		
	parents' parents		
MM	*bʌdʑe*$^{N|TAR}$, *nunu~dudu-bʌdʑe*$^{N|TR}$	MF	*badʑja*$^{N|TAR}$, *nunu~dudu-badʑja*$^{N|TR}$
FM	*bʌdʑe*$^{N|TAR}$, *had(ʌ)*^{N}-*bʌdʑe*$^{N|TR}$	FF	*badʑja*$^{N|TAR}$, *had(ʌ)*^{N}-*badʑja*$^{N|TR}$
	spouse's parents		
HM	*aj*^{TR}, *ama*$^{N|TA}$, *pʰupu*$^{N|TA}$, *bʌdʑe*$^{N|TA}$	HF	*to*^{TR}, *buwa*$^{N|TA}$, *pusʌj*$^{N|TA}$, *badʑja*$^{N|TA}$
WM	*aj*^{TR}, *ama*$^{N|TA}$, *majdʑu*$^{N|TA}$, *bʌdʑe*$^{N|TA}$	WF	*to*^{TR}, *buwa*$^{N|TA}$, *mama*$^{N|TA}$, *badʑja*$^{N|TA}$
	siblings' spouses' parents		
ZHM	*aj*^{TR}, *ama*$^{N|TA}$, *pʰupu*^{TA}, *bʌdʑe*$^{N|TA}$	ZHF	*to*^{TR}, *ama*$^{N|TA}$, *pʰupu*$^{N|TA}$, *badʑja*$^{N|TA}$
BWM	*aj*^{TR}, *ama*$^{N|TA}$, *majdʑu*$^{N|TA}$, *bʌdʑe*$^{N|TA}$	BWF	*to*^{TR}, *buwa*$^{N|TA}$, *mama*$^{N|TA}$, *badʑja*$^{N|TA}$
	spouse's parents' parents		
HMM	*aj-bʌdʑe*$^{N|TR}$, *bʌdʑe*$^{N|TA}$	HMF	*to-badʑja*$^{N|TR}$, *badʑja*$^{N|TA}$
HFM	*aj-bʌdʑe*$^{N|TR}$, *bʌdʑe*$^{N|TA}$	HFF	*to-badʑja*$^{N|TR}$, *badʑja*$^{N|TA}$
WMM	*aj-bʌdʑe*$^{N|TR}$, *bʌdʑe*$^{N|TA}$	WMF	*to-badʑja*$^{N|TR}$, *badʑja*$^{N|TA}$
WFM	*aj-bʌdʑe*$^{N|TR}$, *bʌdʑe*$^{N|TA}$	WFF	*to-badʑja*$^{N|TR}$, *badʑja*$^{N|TA}$
	parents' parents' parents		
MMM	*dʑu-bʌdʑe*$^{N|TR}$	MMF	*dʑu-badʑja*$^{N|TR}$
MFM	*dʑu-bʌdʑe*$^{N|TR}$	MFF	*dʑu-badʑja*$^{N|TR}$
FMM	*had(ʌ)*^{N}-*bʌdʑe*$^{N|TR}$	FMF	*had(ʌ)-badʑja*$^{N|TR}$
FFM	*had(ʌ)*^{N}-*bʌdʑe*$^{N|TR}$	FFF	*had(ʌ)-badʑja*$^{N|TR}$

parents' sibling's spouses' parents			
MBWM	*bʌdʑe*[N|TAR]	MBWF	*badʑja*[N|TAR]
MZHM	*bʌdʑe*[N|TAR]	MZHF	*badʑja*[N|TAR]
FBWM	*bʌdʑe*[N|TAR]	FBWF	*badʑja*[N|TAR]
FZHM	*bʌdʑe*[N|TAR]	FZHF	*badʑja*[N|TAR]

3.3.6.2.4. Ego's children, grandchildren, and great-grandchildren

The kinship terms that relate to the generations of ego's children, grandchildren and great-grandchildren are presented in Table 104.

The formation of kinship terms related to ego's children is calqued from Nepali: children are referred to according to the position they hold amongst their siblings of the same gender, from the eldest, i.e., *dʑetʰa-tɕoʔ* (♂) / *dʑetʰi-tɕoʔ-djaŋ* (♀), to the youngest, i.e., *kantɕʰa-tɕoʔ* (♂) / *kantɕʰi-tɕoʔ-djaŋ* (♀). Similarly in Nepali, the eldest son will be referred to as जेठा छोरा <jeṭhā chorā> 'the first, eldest son.' The first element that specifically expresses the position of the sibling is borrowed from Nepali.

Three additional positions can be added in case there are more than seven same gender siblings, such as *ãtʌrja-tɕoʔ* 'the eighth son,' *dʑãtʌrja-tɕoʔ* 'the ninth son,' *kʰantʌrja-tɕoʔ* 'the tenth son.' The first elements *ãtʌrja*, *dʑãtʌrja*, and *kʰantʌrja* are borrowed from Nepali.

 Regardless of the number of children, the youngest can be referred to as *kantɕʰa-tɕoʔ* (♂) or *kantɕʰi-tɕoʔ-djaŋ* (♀) 'the last, the youngest,' or according to their position. That is, if there are four female siblings, the youngest can be referred to as *kantɕʰi-tɕoʔ-djaŋ* (♀) or *kajli-tɕoʔ-djaŋ* (♀).

Parents mostly call their children using their first name but can also use their rank order position name, such as: *dʑetʰi* (♀) 'the first, the eldest,' *majli* (♀) 'the second,' *sajli* (♀) 'the third,' *kajli* (♀) 'the fourth,' *tʰajli* (♀) 'the fifth,' *ɾajli* (♀), 'the sixth,' and *kantɕʰi* (♀) 'the seventh, the last, the youngest.'

When a sibling goes to live with their spouse's family, they take the position that their spouse holds within their family. While it is more common that a daughter goes live with her husband's family, it also happens that a husband goes live with his wife's

family, becoming *kim-bʰaw*, literally 'house-husband.' For instance, in both cases, they can be referred to by the family members of their spouse as *kantɕʰi* (♀) or *kantɕʰa* (♂) if their spouse is the youngest.

When it comes to refer to ego's relatives' offspring, such as ego's siblings' children, ego's spouse's siblings' children, or ego's parents' siblings' grandchildren, the kinship terms' organization also recalls that attested in Nepali, combining Nepali and Chepang terms.

For ego's siblings' children, ego's gender is considered. This is the only configuration where this applies, apart from reference to ego's spouse and ego's spouse's relatives. As in Nepali, when ego and their sibling share the same gender, the sibling's children are referred to similarly as ego's own children, by the native terms *tɕoʔ* (♂) 'son' and *tɕoʔ-djaŋ* (♀) 'daughter.' Semantically, it is similar to Nepali, which uses the terms छोरा <chorā> and छोरि <chori> to express the same kin relationships, respectively. Conversely, when ego and their sibling do not share the same gender, the sibling's children will be referred to using kinship terms borrowed from Nepali. For instance, a male ego will refer to his sister's children as *bʰandʑa* (♂) 'nephew' and *bʰandʑi* (♀) 'niece.'

While the organization of these kinship terms is also present in Nepali and was likely borrowed into Chepang, it reflects a significance in Chepang that does not apply per se to Nepali since Nepali speakers of IA ancestry do not practice cross-cousin marriage. The fact that in Chepang, parallel nephews and nieces (same gender sibling's children) are referred to as daughters and sons underlines the fact that a marriage between them and ego's children is not socially acceptable. Cross-cousin marriage has traditionally been the prevalent practice (§ 3.3.6.3).

Ego's wife's sister's children, and ego's husband's brothers' children are also referred to as *tɕoʔ* (♂) 'son' and *tɕoʔ-djaŋ* (♀) 'daughter,' whereas ego's wife's brother's children are referred to as *bʰʌdʌj* (♂) 'nephew' and *bʰʌdʌjni* (♀) 'niece,' and ego's husband's sister's children are referred to as *bʰandʑa* (♂) 'nephew' and *bʰandʑi* (♀) 'niece.'

This also reflects the fact that ego's son can be given away to his maternal uncle's daughter, i.e., ego's wife's brother's daughter, and ego's daughter to her paternal aunt's son, i.e., ego's husband's sister's son.

Ego's parents' sibling's grandchildren, or ego's cousins' children are referred to as *tɕoʔ* (♂) 'son' and *tɕoʔ-djaŋ* (♀) 'daughter.' This underlines as well that ego's children and ego's cousin's children cannot be married. Recalls that ego's cousins, referred to as 'sister(s)' and 'brother(s),' constitute potential marital alliances, i.e., specifically cross-cousins (§ 3.3.6.3). If it's the case, this union entails offspring that cannot marry each other, since it would constitute an incestuous union between a brother and a sister.

Finally, the formation of the kinship terms used for great-grandchildren and great-great-grandchildren illustrates an innovation from a borrowing from Nepali. The term पनाति <panāti> is used in Nepali to mean 'great-grandson,' while खनाति <khanāti> is used to mean 'great-great-grandson.' While Chepang borrowed the form पनाति <panāti> to mean 'great-great-grandson,' i.e., *pʌnati*, its form got reduced through the deletion of the initial consonant /p/, i.e., *ʌnati*, to mean 'great-grandson.'

Table 104. Ego's children's, grandchildren's, and great-grandchildren's generations

	EGO'S		
	children		
S	*tɕoʔ*[TR], *goj-tɕoʔ-tɕoʔ*[TR], first name[TA]	D	*tɕoʔ-djaŋ*[TR], first name[TA]
eS1	*dʑeʈʰaᴺ-tɕoʔ*[TR], first name[TA]	eD1	*dʑeʈʰiᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
eS2	*majlaᴺ-tɕoʔ*[TR], first name[TA]	eD2	*majliᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
eS3	*sajlaᴺ-tɕoʔ*[TR], first name[TA]	eD3	*sajliᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
eS4	*kajlaᴺ-tɕoʔ*[TR], first name[TA]	eD4	*kajliᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
eS5	*tʰajlaᴺ-tɕoʔ*[TR], first name[TA]	eD5	*tʰajliᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
eS6	*rajlaᴺ-tɕoʔ*[TR], first name[TA]	eD6	*rajliᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
[yS7]	*ãtʌrjaᴺ-tɕoʔ*[TR], first name[TA]	[yD7]	*ãtʌriᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
[yS8]	*dʑãtʌrjaᴺ-tɕoʔ*[TR], first name[TA]	[yD8]	*dʑãtʌriᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
[yS9]	*kʰʌntʌrjaᴺ-tɕoʔ*[TR], first name[TA]	[yD9]	*kʰʌntʌriᴺ-tɕoʔ-djaŋ*[TR], first name[TA]
yS7[10]	*kantɕʰaᴺ-tɕoʔ*[TR], first name[TA]	yS7[10]	*kantɕʰiᴺ-tɕoʔ-djaŋ*[TR], first name[TA]

siblings' children			
(EGO♀) BD	*bʰʌdʌjni*[N|TAR], first name[TA]	(EGO♀) ZD	*tɕoʔ-djaŋ*[TR], first name[TA]
(EGO♀) BS	*bʰʌdʌj*[N|TAR], first name[TA]	(EGO♀) ZS	*tɕoʔ*[TR], *bʰanis*[TR], first name[TA]
(ego♂) BD	*tɕoʔ-djaŋ*[TR], first name[TA]	(EGO♂) ZD	*bʰandʑi*[N|TAR], first name[TA]
(EGO♂) BS	*tɕoʔ*[TR], first name[TA]	(EGO♂) ZS	*bʰanis*[TR], *bʰandʑa*[N|TAR], first name[TA]

spouses' siblings' children			
HZD	*bʰandʑi*[N|TAR], first name[TA]	WZD	*tɕoʔ-djaŋ*[TR], first name[TA]
HZS	*bʰandʑa*[N|TAR], first name[TA]	WZS	*tɕoʔ*[TR], first name[TA]
HBD	*tɕoʔ-djaŋ*[TR], first name[TA]	WBD	*bʰʌdʌjni*[N|TAR], first name[TA]
HBS	*tɕoʔ*[TR], first name[TA]	WBS	*bʰʌdʌj*[N|TAR], first name[TA]

grandchildren			
♂	*ploʔm*[TAR], first name[TA]	♀	*moʔm*[TAR], first name[TA]

great-grandchildren			
♂	*ʌnati*[TAR], first name[TA]	♀	*ʌnatini*[TAR], first name[TA]

great-great-grandchildren			
♂	*pʌnati*[TAR], first name[TA]	♀	*pʌnatini*[TAR], first name[TA]

siblings' grandchildren			
♂	*ploʔm*[TAR], first name[TA]	♀	*moʔm*[TAR], first name[TA]

parents' siblings' grandchildren			
♂	*tɕoʔ*[TR], first name[TA]	♀	*tɕoʔ-djaŋ*[TR], first name[TA]

3.3.6.3.   Traditional matrilateral cross-cousin marriage and beyond

For the past two generations, marriage practices have drastically changed. In the past, which here roughly corresponds to the generation of those who are now in their mid-forties and beyond, the most widespread practices were capture, arranged marriage, and elopement. At the time, marriage would often occur at a young age, between 9 and

14 years old. It is reported to be between 12 and 20 years old for boys and between 10 and 12 years old for girls by von Nebesky-Wojkowitz (1959), and between 12 and 25 years for both girls and boys by Bista (1967). Beyond 14 years old, teenagers were considered already too old to get married. The girl would usually be younger than the boy.

In the case of a marriage by capture, the boy (along with his male friends) would get a girl either by convincing her to go meet her relatives at her paternal uncle's house along with them (in a cross-cousin marriage configuration in particular) while not letting her come back home, or by coming to her home or wherever she would then be to take her away by force: either during the day (in absence of the girl's parents) or by night (while sleeping for instance in a cowshed when away from home grazing the cattle for several days). In this type of marriage, no discussion between the parents of the boy and that of the girl occurs prior to the capture. Accounts of this practice are countless and condemned by today's generations.

Arranged marriages are still practiced, but at a later age (although child marriage still occurs in some places). Arranged marriages are decided by the children's parents, while privileging the traditional marital alliances between cross-cousins. Once their choice is made, they inform the children who are pushed to respect their decision.

Marriage by elopement is also still practiced by younger generations. When a boy and a girl fall in love, they flee their respective families for some time (a few weeks to some months if it is disapproved by their family). They go on a trip to a different place than their respective villages. By the sole act of eloping, the couple shows to their families that their decision to get married is taken, and even that their marriage "has already taken place." When the couple comes back, they visit their parents to overtly inform them of their desire to get married. If there is no opposition, they organize the marriage festivities. If there is an opposition to the union, the couple may be separated by the parents.

Today's young generations find their partner through common manners: they meet someone, talk and message each-other through their phone for a while, and decide or not to pursue the relationship. A widespread reported practice in the Himalayan region is to use the technique of the *ɾoŋ-nʌmbʌɾ* 'wrong number,' which consists in calling

random phone numbers until they turn out to be a good match. The person who starts the call excuses themselves for having dialed a wrong number, while inquiring where the person lives or is located, and other details. Today'soung generations get in contact through Facebook, Messenger or other online chat and call platforms, offering endless possibilities of meeting someone.

Matrilateral cross-cousin marriage is still practiced amongst the Chepangs (although more present amongst people living in the hills). Matrilateral cross-cousin marriage consists in giving away a son to his maternal uncle's daughter (mother's brother's daughter). The term *matrilateral* is used through a male perspective. Indeed, matrilateral cross-cousin marriage equivales to giving away a daughter to her father's sister's son. In fact, this is what better corresponds to the reality, since women leave their home to go live at their husband's.

Matrilateral cross-cousin marriage is practiced in these basic terms (1 in Figure 72) amongst other TH language speaking communities, like the Magars, Khams, Gurungs, and Tamangs (Hitchcock 1966; Pignède 1966; Doherty 1974; Oppitz 1982; Oppitz 1988; Lecomte-Tilouine 1993; Grunow-Hårsta 2008; Baral Magar 2012; Yonjan Tamang 2012). Some however do not practice at all cross-cousin marriage, such as the Thangmi (Shneiderman & Turin 2006; Regmi 2017).

The clan lineage is patrilinear, that is, children take on the clan's name of their father, referred to as गोत्र <gotra> (and sometimes rather mistakenly as थर् <thar>) in Nepali. Clans are exogamous, which means people marry outside their own clan.

In addition to marrying her father's sister's son, a daughter can also be given away to her father's sister's husband's son (2 in Figure 72), or to her father's father's sister's son's son (that is, her grandfather's sister's son's son, or paternal great-aunt's son's son) (3 in Figure 72). The latter two marital alliances have never been mentioned in the literature but are as much favored as matrilateral cross-cousin.

These three possibilities of marital alliances are illustrated in Figure 72, based on three clans. Note that today's matrilateral cross-cousin marriage does not restrain daughters from being given and received in specific directions as described for instance by Oppitz (1982). That is, a clan does not specifically always give away daughters to the same clan while always receiving daughters from another (or different) clan. They can be

given to any clan, for as long as this latter differs from theirs. There is no clan that is considered taboo amongst the Chepangs, by contrast with the Tamangs for instance, who disallow marital alliances between certain clans.

In Chepang, the kinship terms employed to refer to the kin relation in all three favored marital alliances, by the female and male egos, are *hʌw* 'younger brother' and *hʌw-djaŋ* 'younger sister,' respectively.

Finally, when a woman loses her husband (death, separation), if she does not want to stay alone, she has the possibility to marry her former husband's younger brother (ego's *paŋ*) (only if single). Another possibility for her would be to go live with her elder sister and her husband.

Figure 72. Matrilateral cross-cousin marriage amongst the Chepangs

MATRILATERAL CROSS-COUSIN MARRIAGE*
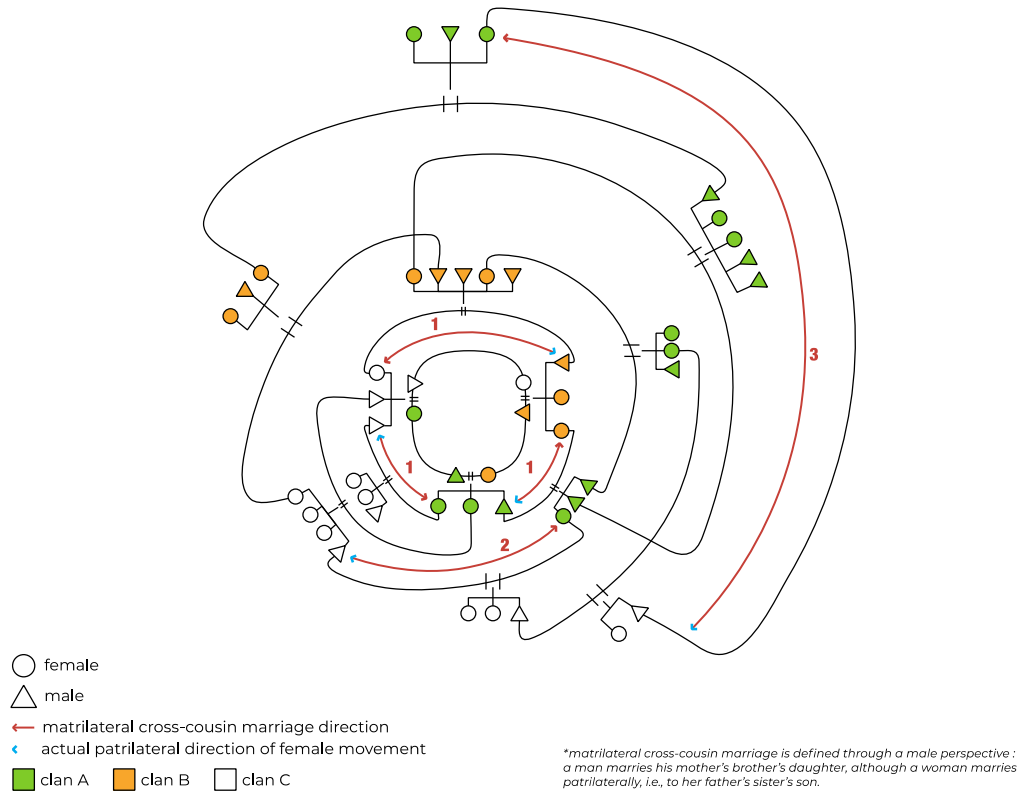
AMONGST THE CHEPANGS

female

male

← matrilateral cross-cousin marriage direction

◄ actual patrilateral direction of female movement

clan A clan B clan C

*matrilateral cross-cousin marriage is defined through a male perspective :
a man marries his mother's brother's daughter, although a woman marries
patrilaterally, i.e., to her father's sister's son.

3.3.6.4.   Chepang kinship terms and TH cognates

In this section, I present some Chepang cognate kinship terms found in other TH languages. I do not intend to analyze these cognates and propose a reconstruction, but simply recognize and acknowledge the possible cognacy of these forms for further discussion. Note that part of the forms presented as cognates with Chepang kinship terms may already have been recognized and acknowledged in previous literature.

Chepang kinship terms and TH cognates are presented in Table 105. In this table, I use the following abbreviations to refer to the sources for the reported forms: A75 for Allen (1975); B20 for Breugel (2020); D09 for Doornenbal (2009); DD85 for Davids and Driem (1985); E97 for Ebert (1997); G90 for Genetti (1990); gg77 for Glover and Gurung (1977); HD92 for Huang Bufan and Dai Qingxia (1992); K03 for Krishan (2003); K10 for Kansakar (2010); L98 for Lahaussois (1998); L17 for Lieberherr (2017); M89a for Michailovsky (1989a); M89b for Michailovsky (1989b); O82 for Oppitz (1982); Pdata for Pons (data); P09 for Post; R85 for Rai (1985); RHH75 for Rai et al (1975); S15 for Schackow (2015b); S91 for Sun (1991); S93 for Sun (1993); T99 for Tolsma (1999); T04 for Turin (2004b); V79 for Vinding (1979); VD87 for van Driem (1987); W87 for Weidert (1987); W02 for Watters (2002); W04 for Watters (2004).

The forms presented in (122) are reconstructed back to PTH by Benedict (1941; 1972).

(122)   *pa   'father'
        *tsa~*za   'child'

        *puw   'elder brother'

Table 105. Chepang kinship terms and TH cognates

Chepang	meaning	cognate
apa	'father'	Limbu *pa* 'father' VD87
		Thangmi *apa* 'father' T04
		Bahing *papa* 'father' M89a
		Thulung *pap* 'father' A75, L98
		Atong *(a-)wa* 'father' B20
		Puroik (Sulung, Bulu) *a-pa* 'father' S91, L17
		Damu *a-pa* 'father' S93
tɕoʔ	'child, son'	Magar *dʑʌ-dʑa*; *-dʑa* 'child' PDATA
		Kham *za* W02
		Limbu *saʔ* 'child, offspring' VD87
		Thulung *tsɯ* 'child' L98
		Yakkha *a-cya* 'child' S15
		Thangmi *ca* 'son' T04
		Newah *mu-ca* 'child' G90
		Baram *u-ca* 'child' K10
		Tamang *ca* 'son' V79
		Gurung *cxa* 'son' GG77
hʌw	'younger brother'	Thangmi *hu* 'younger brother' T04
		Meche *hɤu-aʔ* 'younger brother' W87
		Atong *haw* 'mother's younger brother' B20
puʔ	'elder brother' (arch.)	Limbu *phu* 'elder brother' DD85, VD87
		Limbu *-ŋ-phuq* 'elder brother' M89b
		Limbu *phuʔ* 'elder brother' W87
		Bahing *ho-po* 'elder brother' M89a
		Athpare *bhu* 'elder brother' E97
		Yakkha *a-phu* 'elder brother' S15
		Kulung *bu* 'elder brother' T99
		Kulung *bua* 'elder brother' RHH75
		Bantawa *bua* 'elder brother' R85

Bantawa *buwa* 'elder brother' D09

Thangmi *bubu* 'elder brother' T04

Atong *phaw-jong~phawong* 'elder brother'
B20

Tibetan *phu* 'elder brother'

Jingpho *phu, hpu* 'elder brother' HD92

Darma *pu* 'elder brother' K03

Chaudangsi *po-hya* 'elder brother'

Byangsi *po-hya* 'elder brother'

naʔ	'elder sister' (arch.)	Yakkha *a-na* 'elder sister' S15

Limbu *nɛʔnɛʔ* 'elder sister' DD85

Kham *nana* 'elder sister' W02

Magar (Northern) *nana* O82

Kham (Taka) *nana* 'elder sister' W04

Atong *janaw* 'elder sister' B20

paŋ	'husband's younger brother'	Limbu *paŋmi* 'daughter's husband; siblings' daughter's husband'

paŋmik 'younger sister's husband' VD87

pʰoŋaʔ 'mother's younger sister's husband'
DD85

Khamci *pamma~pomba* 'father's younger
brother; mother's sister's husband' PDATA

Yakkha *a-pʰaŋ* 'mother's younger sister's
husband; father's younger brother' S15

Atong *(a-)waŋ* 'father's younger brother' B20

Bokar *a paŋ* 'father's brother' HD92

bʌh	'father's elder brother; father's elder sister's husband; mother's elder sister's husband; husband's elder brother'	Limbu *tumba* 'father's elder brother; mother's elder sister's husband' VD87, DD85 *-ndumba* 'my father's elder brother; my mother's elder sister's husband' VD87

		Magar (Northern) *pago* 'father's elder brother; mother's elder sister's husband' O82
		Kham (Taka) *pahgo* 'father's elder brother; mother's elder sister's husband' W04
mʌh	'mother's elder sister; father's elder brother's wife; spouse's elder sister'	Limbu *tumma* 'mother's elder brother' DD85
		Magar (Northern) *mago* 'mother's elder sister; father's elder brother's wife' O82
		Kham (Taka) *mahgo* 'mother's elder sister; father's elder brother's wife' W04
meɾu	'wife' (arch.)	Limbu *meʔl* 'wife' DD85
to	'father-in-law'	Bokar *a to* 'father-in-law' HD92
		Apatani *a-to* 'father-in-law' S93
		Galo *a-to* 'father-in-law' P09
kʰon	'daughter-in-law' (ego=parents)	Magar (Western) *me-kʰon* 'daughter-in-law' PDATA

3.3.6.5.   Addressing genealogically unrelated people

The terms used to address people with whom one is not genealogically related, whether they are friends or strangers, are chosen according to the age of the individual and the generation they belong to.

If both individuals belong to the same generation, they call each other 'sister' or 'brother' with the terms of address presented in Table 106. If strangers, both individuals will estimate each other's age and choose the term of address they think is appropriate. If they are or are to become friends, they will investigate each other's exact date of birth and will modify the way they address each other if necessary. When people are born on the same day, they will further consider their birth hour.

Table 106. Same generation terms of address

hʌw-djaŋ	'younger sister'
bʌhjni	'younger sister' (<N.)
hʌw	'younger brother'
bʰaj	'younger brother' (<N.)
daj	'elder brother' (<N.)
dadʑu	'elder brother' (<N.)

If the addressee belongs to the generations above the speaker, they are addressed with the terms of address presented in Table 107.

Table 107. Higher generations terms of address

+1	*ãti*	'aunt' (<N. <E.)
	ãkʌl	'uncle' (<N. <E.)
+2	*ama*	'mother' (<N.)
	buwa, baba	'father' (<N.)
+3	*bʌdʑe*	'grandmother' (<N.)
	badʑja	'grandfather' (<N.)
	hʌdʑuɾ	
any	*sahuni* (♀), *sahudʑi* (♂)	
	sʌrkaɾ~sʌskaɾ	
	sahep	

Finally, when addressing people of a higher position in a professional setting (work, school, etc.), the terms presented in Table 108 are used.

Table 108. Higher in a professional hierarchy terms of address

mjam	'madam' (<N. <E.)

sʌɾ	'sir' (<N. <E.)

3.4. Nominal morphology

This section is dedicated to the description of nominal morphology attested with nouns (noun stems or nominals) forming the head of a noun-phrase.

Nominal morphology can be functionally described as inflectional and derivational (nominal-based nominalization), in addition to adverbial and adpositional since it includes relator nouns or postpositions.

Inflectional morphology comprises number (§ 3.4.3) and case marking (§ 3.4.5), which expresses the syntactic function of an argument in a clause.

Derivational morphology (§ 3.4.4) includes the genitive morpheme *=ko* (§ 3.4.4.1), and the locational nominalizer *=kʌ* (§ 3.4.4.2). These morphemes do not share the same paradigmatic and syntagmatic distributions with inflectional markers and do not have the same syntactic properties. They follow number markers and can be followed by case markers. I follow Shibatani (2019) in analyzing the genitive as a nominal-based nominalization device. This split from the traditional view, analyzing the genitive as a derivational nominalizer rather than as an inflectional device, is also found in Zhu (1982) who describes the Mandarin Chinese genitive morpheme *de* and verb-based nominalizer *de* as the same morpheme, i.e., a noun- and verb-based nominalizer.

All nominal inflectional and derivational devices are enclitics attaching to the noun stem head of a noun-phrase. Relator nouns and postpositions (§ 3.4.6) occupy the syntactic position of the head of a modified noun-phrase. They can be expressed in the presence or absence of the genitive morpheme *=ko* and be followed by locational case markers.

I start by specifying and describing the morphological type of markers used to express inflectional and derivational morphology, i.e., clitics (§ 3.4.1). I then present the paradigmatic and syntagmatic distribution of nominal morphology in relation to a minimal noun-phrase (not including determiners), over four constituent positions (SLOT 1 to 4) following a nominal stem (§ 3.4.2). Finally, I describe number markers (§ 3.4.3),

derivational genitive and locational morphemes (§ 3.4.4), case markers (§ 3.4.5), and relator nouns and postpositions (§ 3.4.6).

3.4.1. Nominal enclitics

Nominal derivational and inflectional morphology is expressed using enclitics, i.e., clitics attaching to the right edge of their host. Like affixes, clitics are bound morphemes, i.e., they need a host to attach to.

Aikhenvald (2003) and Haspelmath & Sims (2010) provide several typological criteria to distinguish affixes from clitics in a scalar approach since there is no such thing as a clear boundary between affixes, clitics and independent or free morphemes. Using these criteria, I show that all nominal inflectional and derivational morphology is expressed using enclitics. However, it is not possible to establish a set of criteria that would define all of them homogenously since they present differences which may be explained as results of their historical development on a continuum between their morphological origin to their clitic status. Haspelmath & Sims (2010: 202) recall that clitics do not form a typological "uniform group" and that they may represent "intermediate stages" from a diachronic perspective.

In Table 109, I summarize a selection of typological characteristics of affixes and clitics according to Aikhenvald (2003: 43–57) and Haspelmath & Sims (2010: 196–102) that are relevant to Chepang. I use the notations A03 and HS10 to respectively highlight the differences between the authors' criteria, if any. I use the + and - symbols to show that the mentioned criterion respectively applies positively or negatively to affixes or clitics, and the +/- symbol when the mentioned criterion applies both positively and negatively, i.e., thus making no distinction between affixes and clitics. The symbol / indicates that the corresponding author does not report such criterion.

Table 109. Selected typological criteria for affixes and clitics

	affixes		clitics	
	HS10	A03	HS10	A03
1 bound forms – need a host	+	+		
2 stressed	-	-/+	-/+	
3 prosodic integration with host	+	-/+	-/+	
4 freedom of host selection	-	-/+	+	-/+
5 morphophonological changes at host boundary	+	-/+	-/+	
6 trigger host morphological suppletive alternation	+	/	-	/
7 possible idiosyncratic meaning with host	+	/	-	/
8 expected combination with host fails to exist	+	/	-	/
9 possible pause at host boundary	/	-	/	+
10 orthographic variation (separated from host or not)	/	-	/	+

The morphosyntactic behavior of Chepang nominal inflectional and derivational enclitics correspond to what is typologically observed. In Figure 73, I describe the particularities of Chepang nominal enclitics with regard to the selected typological criteria described in Aikhenvald (2003) and Haspelmath & Sims (2010) presented in Table 109.

Figure 73. Morphosyntactic behavior of Chepang nominal enclitics

(1) Like affixes, Chepang nominal enclitics are bound forms. They need a host to attach to, i.e., they are not used independently as free morphemes.

(2/3) In terms of prosodic integration, stressed clitics are typologically attested. In Chepang, inflectional number markers and derivational genitive and locational nominalizer morphemes are stressed, while case markers are unstressed. The fact that some enclitics bear stress while others do not may find an explanation in diachrony.

(4) Chepang nominal enclitics' selectivity regarding their host depends on their syntagmatic distribution and what one means by host. If the host is characterized in terms of part of speech, then all nominal enclitics are selective, since they may attach to a host that is a noun or noun stem, head of a noun-phrase. If the host is characterized in terms of morpheme, then nominal enclitics can be considered less selective, since they can attach

313

to a noun stem, to the nominalizer *=o* which forms verb-based nominalizations, to number markers, or else to the derivational genitive and locational morphemes *=ko* and *=kʌ*.

The attested combinations between nominal enclitics and their host are reported in Table 110.

Table 110. Selectivity of nominal enclitics regarding their host

nominal enclitics	noun / noun stem / np head	number	deriv.
number	✓		
derivational genitive	✓	✓	
derivational locational	✓	✓	
case markers	✓	✓	✓

(5) Some Chepang nominal enclitics may entail morphophonological changes at morpheme boundary. For instance, the locational morpheme *=haŋ* in combination with the 3rd person pronouns *i* PROX and *o* DIST can reduce to *=ŋ*, such as *o=ŋ* 'there.' This change is observed in free variation, as a result of speech rate and frequency. Another morphophonological change that occur at morpheme boundary, in this case conditioned, is observed with the ergative case marker *=i* and instrumental case marker *=i*. These morphemes are deleted when attaching to a homorganic vowel, i.e., to a noun stem ending in an open syllable whose rhyme is a vowel /i/.

(6) A clitic is not likely to entail the presence of a suppletive form of the host by contrast with an affix. A suppletive form is understood as morphologized. The clitic *=haŋ* may however trigger its host morphophonological change when attaching to the 3rd person pronouns *i* PROX and *o* DIST. The vowel quality of the 3rd person pronouns can change to [e] for the proximate pronoun, and to [ʌ] for the distal pronoun. This change is however not morphologized but attested in free variation, as a result of speech rate and frequency.

(7) As posited for clitics, there is no idiosyncratic meaning attested when Chepang nominal enclitics attach to a particular host.

(8) Expected combinations between a host and an enclitic do not fail to exist. This is also the case for Chepang nominal enclitics.

(9) In Chepang, pauses may be observed between stressed enclitics (number, genitive and locational derivational morphemes) and their host, but not in the case of unstressed enclitics (case markers).

(10) Intuitively, Chepang speakers may transcribe stressed enclitics (number, genitive and locational derivational morphemes) as separate from their host. This has also been observed with unstressed clitics (case markers), but not as much.

Finally, another remark on Chepang enclitics is to be noted. Stressed enclitics (number, genitive and locational derivational morphemes) do not show any morphophonological change within the root, while unstressed clitics (case markers) may be morphologically reduced. These changes often occur as a result of speech rate, and are mainly observed with the dative case marker =*kaj* (which may be pronounced [ka] or [kʌ]), the ablative case marker =*sʌj* (which may be pronounced [sʌ]), and the locational case marker =*haŋ* (which may be pronounced [hã], [ãŋ], or else [hẽŋ], [ẽŋ], [hõŋ], [õŋ], when attaching respectively to the 3rd person pronouns *i* PROX and *o* DIST, harmonizing with the vowel quality of the pronouns).

To sum up, Chepang derivational and inflectional morphemes attested with nouns can be analyzed as enclitics rather than suffixes. There is no set of criteria that can uniformly apply to all of them, as seen for example through the fact that number markers and derivational morphemes are stressed while case markers are unstressed. As mentioned in typological scalar descriptions of clitics (Aikhenvald (2003) and Haspelmath & Sims (2010)), their non-uniformity may be explained through their diachronic developments. For instance, the number markers =*nis* DU and =*lʌm* PL likely originate from a coordinative compound construction, which would explain the fact that they are stressed.

3.4.2. Paradigmatic and syntagmatic distribution of nominal morphology

A nominal stem (derived for number or not) forms the head of a minimal noun-phrase that can be further encliticized with inflectional morphemes (number and case marking) or be followed by relator nouns and postpositions to function as a core argument or as an adverbial or oblique argument of a verb in a clause. A nominal stem (not inflected for case) can also be derived with the genitive morpheme =*ko* and locational nominalizer =*kʌ* to express a new denotation: the possessed entity of the referent and the living location of the referent, respectively. Relator nouns and postpositions occupy the same position of a modified noun-phrase, showing evidence that relator nouns and postpositions originally come from nouns.

In Table 111, I present a schematized template that shows the paradigmatic and syntagmatic distribution of nominal morphology; I aim at representing its distribution following a noun stem which can either or not function as a modifying or derived noun stem, in addition to the position of relator nouns and postpositions which occupy the slot of a modified noun-phrase.

Table 111. Paradigmatic and syntagmatic distribution of nominal morphology

minimal NP1 HEAD		NP1 head modifier	modified NP2 head	argument role
noun stem	inflectional	derivational	NP2 HEAD	inflectional
noun stem	+1 | number	+2 | nominalization	+3 | relator noun | postposition	+4 | case
	=*nis* DU	=*ko* GEN	<u>RELATOR NOUN</u>	=*i* ERG
	=*lʌm* PL/SML/ASS	=*kʌ* NMZ:LOC	*kaʔm* 'downwards'	=*i* INST
			tjaw 'upwards'	=*kaj* DAT
			mʌlgʌ~eleŋ 'under, below'	=*kusi* COM
			tajli 'above'	=*haŋ* LOC1
			madʑʰʌ 'middle'	=*kʰa* LOC2
			tuʔŋ 'foot of'	=*sʌj* ABL
			tɕjo 'top of'	=*haŋ*=*sʌj* ABL
			ljam 'from'	=*kʰa*=*sʌj* ABL
				=*taŋ* ALL
			<u>RELATOR NOUN (< N.)</u>	

			lagi 'for'
			nimti 'for'
			biʨe~biʨeʨeʌ 'middle'
			POSTPOSITION (< N.)
			ʌgaɖi 'front of, before'
			ʌgʰi 'before'
			pʌʨʰaɖi 'behind, back of, after'
			pʌʨʰi 'after, later'
			waɾi 'deictic side'
			paɾi 'deictic opposite side'
			waɾi-paɾi~paɾi-waɾi 'around'
			dekʰi(n) 'from, since'
			sʌmmʌ(n) 'up to, until'
			bʰitɾʌ 'inside (of)'
			bahjɾʌ 'outside (of)'
			matʰi 'upw., above'
			muni 'downw., under, below'
			mʌddʰe 'amongst'
			ʌnusaɾ 'according to'
			nimti 'for the sake of'
			bahek 'except'
			bina 'without'

The above Table 111 shows that four types of constituents can follow a noun stem: number (SLOT 1), derivation (SLOT 2), relator noun and postposition (SLOT 3), and case marking (SLOT 4).

These four types of constituents do not occur at the same syntactic level. SLOT 1 marks a noun stem for number (§ 3.4.2.1, § 3.4.3). SLOT 2 derives a new referent through the use of the genitive morpheme *=ko*, which can modify a noun-phrase (expressed or not) or allow the presence of relator nouns and postpositions, or through the use of the locational nominalizer *=kʌ* to express the usual living place of a referent (§ 3.4.2.2, § 3.4.4.1). SLOT 3 hosts relator nouns and pospositions or a possible modified noun-phrase (§ 3.4.2.3, § 3.4.6). SLOT 4 is the position occupied by case markers which encliticize to the last element of a noun-phrase, which can either be SLOT 1, 2 or 3,

depending on the semantic type of referent expressed and their syntactic role in the clause (§ 3.4.2.4, § 3.4.5).

What Table 111 does not show are the possible syntagmatic co-occurrences of the different paradigmatic elements over all four types of constituents. For instance, while the relator noun *tjaw* 'upwards' requires the presence of the genitive morpheme *=ko* and allows the optional presence of locational case markers, the Nepali borrowed postposition *mʌddʰe* 'amongst' by contrast, does not occur with any of these elements.

In the remainder of this section, I describe each of these four types of constituents. I give more details about the enclitic or independent status of the nominal morphological forms, about the type of possible syntagmatic co-occurrences that exist over their elements, and between them and the different semantic referents that can be expressed by the noun stem.

3.4.2.1.   Slot 1: Number

The first slot after the noun stem is occupied by number which may be dual or plural (§ 3.4.3). Dual and plural number markers are enclitics and bear stress. Syntactically, number attaches to the noun stem, head of a noun-phrase.

The plural morpheme *=lʌm* is illustrated in (123) and the dual morpheme *=nis* is illustrated in (124).

(123) *Chepang=lʌm,*   *badʑa=ko*   *sʌkti*   *dʑaʔ=haŋ*
       Chepang=PL   elder.M=GEN   power   tiger=LOC1

       lʌt=ti   *wah=sa*   *kʰaj=o.*
       climb=SEQ1   walk=NMZ1 be.able=PERF

       'The Chepangs, with the power of the ancestors, they are able to go around climbing on / ride tigers.'
       CH_MKW_BBC_SIL_032920_2_My grandfather

(124) *ʌni*   *naŋ=ko*   *hʌw=nis=ma*
       then   2SG=GEN   younger.brother=DU=ADD

lʌj=ko	*bʰasa=le*	*noʔ=na=tɕʌ ?*
SLF.INTS=GEN	language=DIS	speak=NPST=1/3DU

'And your two little brothers too, they also speak their own language?'

CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

Number markers do not encliticize to noun-phrases. While the plural enclitic *=lʌm* in (125) and (126) seems to attach to a complex noun-phrase formed with the coordinative morpheme *rʌ* 'and.' This type of construction is borrowed from Nepali and rather rare in our corpus. In (125), the morpheme *=lʌm* does not have scope over the entire complex noun-phrase result of a coordination of two noun-phrases but applies to the noun-phrase *rʌjti* 'subject,' since *radza* 'king' is singular and does have more than one subject. In other words, the morpheme *=lʌm* does not stand for the plurality of the king in addition to his subjects. Example (126) is prototypical of a borrowed construction from Nepali where the plural has scope over the complex noun-phrase result of coordination. Indeed, the speaker had several daughters and only one son who were studying in school at the time of speech.

In the native construction, *=lʌm* attaches to each noun stem head of its own noun-phrase, as illustrated in (127) and (128). The same applies when the plural marker *=lʌm* functions as a similative marker, as in (129). Finally, number can also attach to nominalized clauses, as in (130).

Note that independent pronouns, which also constitute a noun stem, do not additionally get inflected by dual and plural markers, since they are already formed and lexicalized with their own dual and plural markers if any (§ 4.1). The same applies for noun stems formed with place names attached with the denonymic derivational marker -*mʌj* which intrinsically denotes a plural referent (§ 3.3.3).

(125)	*radza*	*rʌ*	*rʌjti=lʌm*	*dzʌmma*	*hum=ti,*
	king	and	subject=PL	ALL	gather=SEQ1

o   ɾadʑa=kaj   gʰeɾ=ʌ=ka=n=i.

DIST   king=DAT   surround=LN=2/3PST=DIR/TR=PL

'They gathered together the king and the subjects and surrounded that king.'

CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

(126)   ŋa=ko   tɕoʔ-djaŋ   ɾʌ   tɕoʔ=lʌm=ma
        1SG=GEN   daughter   and   son=PL=ADD

        pʌɾ-ʌ=ti=le   mu=na.
        study-LN=SEQ1=DIS   COP=NPST

        'My daughters and son are studying too.'
        CH_MKW_BMB_BAN_090118_8_Marriage

(127)   ʌ,   ploʔm=lʌm   moʔm=lʌm   noʔ=waj=lʌ.
        uh   grandson=PL   granddaughter=PL   speak=CERT=NEG

        'Uh, my grandsons and granddaughters don't speak at all.'
        CH_MKW_MRC_LPK_080918_3_Chepang_Language

(128)   ʌni   tɕoʔ-deaŋ=lʌm=kaj   tɕoʔ=lʌm=kaj,   abʌ   doh   bʌj=sa ?
        then   daughter=PL=DAT   son=PL=DAT   now   what   give=NMZ1

        'Then what to give to my daughters and sons?'
        CH_MKW_KMC_SK_082918_3_The_Bear

(129)   sʌhʌjog   bʌj=na=u   hʌj,   bʌlʌ bʌlʌ   gʌhʌt,
        help   give=NPST=3O/DIR   PART   little little   vetch

        mas=lʌm   aʔl=ti   bʌj=na=u   ma !
        black.lentil=SML   bring=SEQ1   give=NPST=3O/DIR   PART

        'Help them hey, please bring and give them some vetch, black lentils and stuff.'
        CH_MKW_SC_SIL_120619_2_E_2

(130)   o   pʰek=paj   o   sjas=o
        DIST   broom=DIS   DIST   raise_breed=NMZ:REL

320

sjas=o=lʌm=ko   *mʌtrʌj.*

raise_breed=NMZ:REL=PL=GEN   only

'But these brooms are only of the kind that comes from those who cultivate them.'

CH_MKW_RC_JMC_SIL_120119_Conversation

3.4.2.2.   Slot 2: Derivation

The second slot, which follows the noun-phrase formed by a noun stem (followed or not by SLOT 1) is the position of the genitive morpheme =*ko* (§ 3.4.4.1). The genitive case is different from other case markers, since it does not share the same syntactic distribution and functionally derives new denotations through nominal-based nominalizations (§ 3.4.4.1), as in (131) with the noun *kam* 'work.'

(131) *niŋ=i,*   *i*   *pande=lʌm=ko*   *kam*   *dʑahŋ=ljam !*
        2PL=ERG   PROX   shaman=PL=GEN   work   do_make=IMP.NEG

'You shall not do the work of these shamans!'
CH_MKW_PSC_MAI_012620_2_Becoming_Christian

In addition, the genitive marker =*ko* allows the presence of relator nouns and certain postpositions as following constituents, as in (132) with the native relator noun *tajli* 'top of,' and in (133) with the postposition *bʰitrʌ* 'inside,' borrowed from Nepali. The presence of the morpheme =*ko* is required for the seven native relator nouns, i.e., *kaʔm* 'downwards,' *tjaw* 'upwards,' *mʌlgʌ~eleŋ* 'under, below,' *tajli* 'above,' *madʑʰʌ* 'middle,' *tuʔŋ* 'foot of' and *tɕjo* 'top of,' and three of the Nepali borrowed postpositions, i.e, *matʰi* 'upwards, above,' *muni* 'downwards, under, below,' *lagi* 'for' and *nimti* 'for.'

(132) *o*   *mʌkʌj*   *biruwa=ko*   *loʔ=ko*   *tajli=sʌj, (...)*
        DIST   corn   shoot=GEN   leaf=GEN   top=ABL

'At the top of that leaf of corn shoot, (…)'
CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(133)   *ow=ko*   *bʰitɾʌ*   *ow*   *kɾaʔ=ko*   *dʰuʔŋ*
         DIST=GEN   inside   DIST   termite=GEN   egg.bag

    mu=na   *dʌj !*
         COP=NPST   PART

    'Inside that, there is that termites' egg bag, hey!'
    CH_MKW_PMRC_LAM_081618_3_Kalitar

The second slot also hosts the derivational locational nominalizer morpheme *=kʌ*. When the referent denotes a human or an animal, or any living entity, the locational nominalizer *=kʌ* is required for that argument to function as a locational adverbial or oblique argument and be encliticized with locational case markers. Such adverbial or oblique argument expresses the referent's usual living place, as in (134), or by metaphor the idea conveyed by English prepositions such as 'amongst' or 'at,' as in (135).

    Like the genitive morpheme *=ko*, the locational nominalizer is an enclitic that has scope over the entire noun-phrase, as in (136) with a complex noun-phrase formed with two noun stems attached with the plural morpheme *=lʌm* here functioning as a similative.

(134)   *pande=kʌ=haŋ=tɕʌhe*   *ʌ,*   *lajt*   *na=jak=lʌ,*
         shaman=NMZ:LOC=LOC1=DIS   uh   light   COP=REM.PST=NEG

    o   *bela=haŋ,*
         DIST   moment=LOC1

    'At the shaman's, there was no light before, at that time,'
    CH_CTW_SRP_GUN_102620_5_Animism_to_Christianity

(135)   *biɕwaɕ*   *Chepang=lʌm=kʌ=haŋ*   *mu=na.*
         Christianity   Chepang=PL=NMZ:LOC=LOC1   COP=NPST

    'There is Christianity amongst the Chepangs.'
    CH_MKW_BBC_SIL_032920_1_Chepang people

(136)   *ŋa=paj*   *ama=lʌm*   *ba=lʌm=kʌ=taŋ*   *pahj=na=ŋ, (...)*
         1SG=DIS   mother=ASS   father=ASS=NMZ:LOC   leave=NPST=1

'I'm gonna leave to my mother and father's familly, (…)'
CH_CTW_BBC_POL_111720_1_Cing_Lan

Finally, this second slot also hosts the comparative morpheme =bʰʌnda borrowed from Nepali. This morpheme is optionally used in combination with nine of the postpositions borrowed from Nepali: matʰi 'upwards, above,' muni 'downwards, under, below,' ʌgaɖi 'front of, before,' pʌtɕʰaɖi 'behind, after,' bʰitrʌ 'inside,' bahjrʌ 'outside,' wari 'deictic side,' pari 'deictic opposite side,' and wari-pari~pari-wari 'around.' This is illustrated in (137).

(137) *wa*   *kim=bʰʌnda*   *ʌgaɖi*   *mʌ=na.*
       bird_hen   house=CMPR   front   COP=NPST

'The hen is in front of the house.'
CH_MKW_1_21-22_CPR_BAN_100817_1_E

3.4.2.3.   Slot 3: Relator Nouns and Postpositions

The third slot is the position dedicated to a possible additional noun-phrase when modified by the genitive morpheme =ko, as in (138), or to the presence of relator nouns and postpositions (§ 3.4.6), as in (139) and (140), respectively. The morphosyntactic differences between relator nouns and postpositions is discussed in § 3.4.6, but by contrast with number and case markers, they do not attach to a constituent; they are independent since they may as well be used as adverbs without a host, as in (141).

Indeed, the genitive morpheme =ko is required in presence of native relator nouns, while it may be optional or not required in presence of the postpositions borrowed from Nepali, as in (142). Note that Chepang also borrowed the construction illustrated in (143), that features the Nepali comparative morpheme =bʰʌnda which may precede postpositions borrowed from Nepali.

The native relator nouns and eight of the postpositions borrowed from Nepali, i.e., ʌgaɖi 'front of, before,' pʌtɕʰaɖi 'behind, after,' bʰitrʌ 'inside,' bahjrʌ 'outside (of),' wari 'deictic side,' pari 'deictic opposite side,' wari-pari~pari-wari 'around' and bitɕ~bitɕtɕʌ

'middle (of)' can further be followed by SLOT 4, that is, be encliticized with locational case markers, as in (144).

(138)   *gaw=ko*   *tol=ko*   *nam,*   *Polkim.*
        village=GEN   area=GEN   name   Polkim

        'The name of the village area is Polkim.'

        lit. 'The village's area's name is Polkim.'

        CH_CTW_BBC_POL_102420_1_Tu'm

(139)   *mʌndir=ko*   *tjaw=kʰe,*   *sjaʔm=lʌm*   *mu=na=i,*
        temple=GEN   upwards=DIS   Tamang=PL   COP=NPST=PL

        Tamang=lʌm   *mu=na=i.*

        Tamang=PL   COP=NPST=PL

        'Upwards the temple, live the Tamangs, live the Tamangs.'

        CH_MKW_MRNDC_SIL_081818_2_Chepang_Language_Culture

(140)   *tɕoʔ-deaŋ,*   *i=h,*   *ŋa=ko*   *pʌtɕʰadi*   *mu=na.*
        daughter   PROX=DEIC   1SG=GEN   behind_after   COP=NPST

        'My daughter, this one here, she is behind me.'

        CH_MKW_SC_SIL_122619_1_Nakko_Co'

(141)   *Syamrang,*   *bʌlʌ*   *tjaw=kʰa*   *pʌr=a.*
        Syamrang   a.little   upwards=LOC2   have.to_fall=PST

        'Syamrang is located a little upwards.'

        CH_CTW_BBC_POL_102520_1_Polkim

(142)   *o*   *pʌtɕʰadi*   *ramro rekʌt*   *gʌm=u.*
        DIST   behind_after   nice   recording   put_keep=2SG.IMP.TR

        'After that, keep well the recordings.'

        CH_MKW_BRC_CYO_120119_Yukdhung

(143) ŋa=ko   kim=bʰʌnda   pʌtɕʰaɾi   jat=ʥjaŋ   ɾaŋ   mu=na.
       1SG=GEN   house=CMPR   behind_after   one=CL1   field   COP=NPST

'There is a field behind my house.'

CH_MKW_1_73-79_CPR_BAN_102417_2_Verb

(144) ten=ko   lagi,   ŋa=ko   pɾʌbʰu=ko   ʌgaɖi=haŋ,
       today=GEN   for   1SG=GEN   lord=GEN   front=LOC1

       kam,   i   kuɾa   noʔ=alaŋ,
       work   PROX   thing   speak=1.PST

'For today, in front of my lord, the work (for Jesus), I spoke about it.'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian

3.4.2.4.   Slot 4: Case markers

The fourth slot consists of case markers (§ 3.4.5), either grammatical or adverbial/oblique. They are all enclitics. They can directly attach to a noun-phrase, as in (145), a =kʌ NMZ:LOC derived noun-phrase, as in (146), a =ko GEN derived noun-phrase in presence of a modified noun-phrase, as in (147), or a relator noun or postposition, as in (148). Core argument case markers and the adverbial comitative case marker can also directly encliticize to a =ko GEN derived noun-phrase without the presence of a head noun if its referent is contextually recoverable, as in (149).

(145) lan=lʌm=i   Silinge=sʌj   klaʔ=o
       spirit=PL=ERG   Silinge=ABL   fall_go.down=NMZ:REL

       manta=lʌm=kaj   hot=ka=n=i : (...)
       person=PL=DAT   ask=2/3.PST=DIR/TR=PL

'The spirits asked to the people who went down from Silinge: (…)'

CH_MKW_BBC_SIL_032820_1_Lanrang

(146) o   pʰeɾi   ama=kʌ=haŋ=ma   dah=na.
       DIST   again   mother=NMZ:LOC=LOC1=ADD   reach=NPST

'He again arrives at his mum too.'

(147) *mil=ʌ=ti,*   *ŋa=ko*   *ama=lʌm=kusi*   *naŋ*   *mu=ljam.*
agree=LN=SEQ1   1SG=GEN   mother=ASS=COM   2SG   COP=IMP.NEG

'Manage (the situation) and don't stay with my mother's family.'

CH_MKW_SC_SIL_122619_1_Nakko_Co'

(148) *ɾu*   *siŋ=ko*   *tuʔŋ=haŋ*   *mu=na.*
snake   tree=GEN   foot=LOC1   COP=NPST

'The snake is at the foot of the tree.'

CH_MKW_BBC_SIL_032920_Topological_Questionnaire_55

(149) *- su=ko*   *tɕoʔ=kaj*   *bʌj=sa ?*
who=GEN   child=DAT   give=NMZ1

- ŋa=ko=kaj !

1SG=GEN=DAT

'- Whose child to give it?

- To mine!'

CH_MKW_PC_SIL_E

3.4.3. Number

Two morphemes encoding number are attested encliticizing to noun stems: a dual morpheme *=nis* (§ 3.4.3.1) and a plural morpheme *=lʌm* (§ 3.4.3.2). Both forms attach to nouns denoting animate referents.

The morphemes *=nis* and *=lʌm* have additional functions: *=nis* is used to coordinate two nouns whose referents hold a kin relationship; *=lʌm* functions as a similative plural with nouns denoting inanimate referents and as an associative, denoting the family members of the referent when attached to a kinship term or a proper noun.

When describing the dual and plural markers, I also discuss questions of definiteness of the referent.

3.4.3.1.   Dual =*nis*

The morpheme *nis* functions as a free morpheme which means 'two.' However, *nis* 'two' always occurs either in constructions where it determines a noun referring to a year or quantity (150), or followed by the classifier =*dʑo~dʑjaŋ*, as in (151) , or in -*tɕa* constructions expressing a pair, as in (152) (see § 3.3.4.2). Its use as a free morpheme merely occurs when the speaker expresses the idea of "some" or a vague number of things in combination with *sum* 'three' as in (153). Even when a speaker counts the numbers, the classifier =*dʑo~dʑjaŋ* is often present.

(150)   *ŋa=i   i   tuʔm   gʌm=o,   nis   bʌrsʌ   sjaw=a.*
        1SG=ERG   PROX   bee   put=PERF   two   year   become=PST

        'It's been two years that I have set up these bees.'
        CH_MKW_MRNDC_SIL_081818_1_Bee_keeping

(151)   *teh=ma   nis=dʑjaŋ   sat=aka=n.*
        last.year=ADD two=CL1   kill=2/3.PST=DIR/TR

        'Last year too, (a tiger) killed two (goats).'
        CH_MKW_PMRC_LAM_081618_3_Kalitar

(152)   *nis   ama-tɕa=taŋ   mu=o,   kim=haŋ.*
        two   mother-pair=ATT   COP=PERF   house=LOC1

        'Mother and son were at home.'
        CH_MKW_DBC_MAI_2_020320_Newa_Dung

(153)   *nis   sum   bʌtɕʌn=kʰe   ŋa=kaj   bʌj=sa*
        two   three   promises=DIS 1SG=DAT   give=NMZ1

        pʌr=na   da !
        have.to_fall=NPST   PART

'One has to give me some promises, hey!'

CH_MKW_ARRKC_CHI_102919_1_Life

With the encliticization of the dual morpheme *=nis* to a noun, the referent denoted by the noun is always animate, as illustrated in (154). The referent can be definite, as in (154) and (155), or indefinite, as in (156). If the noun is inanimate, another construction occurs, i.e., with the determiner *nis=ʥjaŋ~ʥjo*, as in (157).

(154) *ʌni*   *ŋa=ko*   *o*   *bʌʨɕa=nis=i*   *ʌ,*   *kʰel=ʌ=na=ʨʌ.*
        then   1SG=GEN   DIST   child=DU=ERG   uh   play=LN=NPST=1/3DU

'So, those two children of mine uh, they two play.'

CH_MKW_BBC_SIL_042420_6_My two daughters

(155) *i*   *manta=nis*   *waŋ=ka=ʨʌ.*
        PROX   person=DU   come=2/3.PST=1/3DU

'These two persons came.'

CH_CTW_SPC_POL_E

(156) *ʌhʌ,*   *kim=kʰa*   *biɾalo=nis*   *mu=na=ʨʌ !*
        oh,   house=LOC2   cat=DU   come=NPST=1/3DU

'Oh! There are two cats in the house!'

CH_CTW_SPC_POL_E

(157) *(...)*   *nis=ʥjaŋ*   *tak-saj=taŋ*   *ɾʌh=a*   *kja !*
        two=CL1   mango.fruit=ATT   fall=PST   PART

'(…) two mangos fell, imagine!'

CH_MKW_KBKC_TNG_102617_2_Two_Mangos'_pits

The dual morpheme *=nis* occurs as well attached to nouns followed by another noun forming a coordinate construction. This is only found with two nouns that form a pair of referents holding a kin relationship, specifically, forming a couple. This is

illustrated in (158) and (159). In this construction, the order of the nouns can vary, as one can see comparing (159) and (160).

(158) *tɕoʔ=nis*   *kʰon=i*   *kʌdʌw*
       son=DU   daughter-in-law=ERG   millet

       tat=ka=tɕ=u.
       cut=2/3.PST=1/3DU=3O/DIR

       'My son and daughter-in-law cut the millet.'
       CH_MKW_NC_DAM_112819_1_Conversation with Bipana

(159) *baba=nis*   *ama :*   "*e,*   *ma ?*   *lʌw,*   *lʌt=i,*
       father=DU   mother   oh   yes   well   climb=1PL.IMP.INTR

       sjahŋ,   *leʔ=laŋ !"*   *dahj=ti=taŋ*   *bat*   *sjaw=a.*
       tomorrow   take_buy=PUR   say=SEQ1=ATT   matter become=PST

       'The father and mother saying "Oh, really? Okay, let's climb (up there) tomorrow to take (it)!" and the thing was set.'
       CH_MKW_SC_SIL_122619_1_Nakko Co'

(160) *lʌw,*   *ama=nis*   *baba,*   *ŋa=paj*   *pahj=na=ŋ.*
       well   mother=DU   father   1SG=DIS   leave_go.home=NPST=1

       'Well, mother and father, I will go back home.'
       CH_CTW_BBC_GUN_102620_1_Cing_Lan

3.4.3.2.   Plural =*lʌm*

Like the dual number morpheme =*nis*, the plural morpheme =*lʌm* attaches to nouns denoting animate referents, as in (161).

(161) *ni=ko*   *gãw=haŋ*   *na=to*   *manta=lʌm*   *mu=na=i.*
       1PL=GEN   village=LOC1 be.a.lot=NMZ:ADV2 person=PL   COP=NPST=PL

'A lot of people live in our village.'

CH_MKW_1_40-47_CPR_BAN_101217_1_E

The morpheme =lʌm functions as a plural marker on nouns denoting animate referents, and in particular human referents. It is optional on nouns denoting non-human animate referents, such as animals and insects, as illustrated in (162) to (166). In absence of =lʌm, the verb does not inflect for plural, as in (164).

(162)   tʰoɾ   rʰam=laŋ   gljuhŋ=alaŋ,   kaʔmʌ.
        cow   graze=PUR   go.out=1.PST   downwards

'I went out to graze the cows, downwards.'

CH_MKW_TRC_DAM_112819_Conversation_with_Bipana

(163)   ow=i   pʌhila   wa   ʥe=na=u,
        DIST=ERG   first   bird_hen   eat=NPST=3O/DIR

        ʌni   lonte, kwi   ʥe=na=u,
        then   later   dog   eat=NPST=3O/DIR

'It first eats the hens, and later eats the dogs.'

CH_MKW_BAN_081618_The_Leopard_Spirit

(164)   ŋa=ko   metɕʰja   mu=na.
        1SG=POSS   goat   COP=NPST

'I have a goat/goats.'

CH_MKW_1_40-47_CPR_BAN_101217_1_E

(165)   ŋa=kusi   metɕʰja=lʌm mu=na=i.
        1SG=COM   goat=PL   COP=NPST=PL

'I have goats.'

CH_MKW_1_40-47_CPR_BAN_101217_1_E

(166)   i=tʌ   haj=ti   mu=na=i,   kwi=lʌm.
        PROX=NMZ:ADV1   do=SEQ1   COP=NPST=PL   dog=PL

'They sit doing like this, the dogs.'
CH_MKW_SC_SIL_120619_4_E_2

With inanimate referents, plural is left unmarked, as shown in (167) and (168). Like with non-human animate referents, in absence of plural marker *=lʌm* on the noun, the verb is not indexed for plural, as in (168).

(167) *Chepang=lʌm=ko   kim,   baŋ   rʌ   saʔ=i   bʌn=ʌ=o.*
      Chepang=PL=GEN   house   stone   and   earth=INST   make=LN=NMZ:REL

      'Chepangs' houses are made with stones and earth.'
      CH_MKW_BBC_SIL_032920_1_Chepang people

(168) *tʌrʌ   ŋa=i   tɕiʔ=to   sʌmʌn tin=ta   kim*
      but   1SG=ERG   know=NMZ:ADV2   until three=CL2   house

      kʰe=to.
      COP=REM.PST

      'But as far as I know, there were three houses.'
      CH_MKW_LC_SIL_113019_1_Cave

When the morpheme *=lʌm* occurs on a noun denoting an inanimate referent, it functions as a similative plural. It conveys the meaning of "this referent and associated things," or "this referent and its varieties" (Moravcsik 2004; Daniel & Moravcsik 2013; Mauri & Sansò 2018; 2021). This is illustrated in (169) to (172).

For instance, in (169), *kjan=lʌm* dish=SML can refer to vegetables or meat, and other things needed to make a dish, such as spices, condiments, or garlic. In (170), *kitab=lʌm* book=SML refers to not only books but also textbooks, pens, or other things associated with the idea of studying. In (171), *ro=lʌm* flower=SML can denote flowers of different varieties, but also flowers and weeds.

In presence of the similative morpheme *=lʌm* on nouns denoting inanimate referents, the verb does not inflect for plural, as shown in (170) and (171).

However, in some cases, as in (172), where the trees are anthropomorphologized, the presence of *=lʌm* on nouns combines with the presence of the plural marker on the verb. In this case the function of *=lʌm* is rather plural than similative. In fact, in the Chepang variety of RAP-13 (SYAM and POL in particular), people also use the plural morpheme *=lʌm*, along with the plural marker *=i* on the verb, with inanimate nouns denoting referents that are considered alive in nature, like flowers and trees, as in (173). This can relate to the Chepangs' traditional animist beliefs.

(169) *ŋa*   *kjan=lʌm*   *leʔ=ŋʌ=lʌ*   *ten.*
      1SG   dish=SML   take_buy=1=NEG   today

      'I won't buy vegetables and such today.'
      CH_MKW_SC_SIL_120619_4_E_1

(170) *ŋa=kusi*   *kitab=lʌm*   *mu=na.*
      1SG=COM   book=SML   COP=NPST

      'The books and such are with me.'
      CH_CTW_SPC_POL_E

(171) *ɾaŋ=haŋ*   *ɾo=lʌm*   *oɾ=na.*
      field=LOC1   flower=SML   bloom=NPST

      'In the field, flowers and such / all sorts of flowers are blooming.'
      CH_MKW_1_73-79_CPR_BAN_102417_2_Verb_Archive

(172) *o*   *siŋ=lʌm=i=ma*   *maja* *dʑahŋ=o*   *dʑati,* *Chepang.*
      DIST   tree=PL=ERG=ADD   love   do_make=NMZ:REL   group   Chepang

      'The Chepangs are a group which is loved by those trees too.'
      CH_MKW_GBC_CYO_120119_Conversation

(173) *bʌn=kʰa,*   *na=to*   *siŋ=lʌm*   *mu=na=i.*
      jungle=LOC2   be.a.lot=NMZ:ADV2   tree=PL   COP=NPST=PL

      'In the jungle, there are a lot of trees.'
      CH_CTW_SPC_POL_E

In absence of a possessive or demonstrative determiner, nouns marked with the plural morpheme =lʌm can either be interpreted as indefinite, as in (174) and (175), generic, as in (176), or definite, as in (177).

(174)   *Simi*   *Bhumi*   *dekʰi,*   *siŋ=lʌm=kaj*   *sʌmmʌ,*
       Simi   Bhumi   since   tree=SML=DAT   until

       pudʑa   *dʑahŋ=sa,*   *pʌtɕʰi*   *pʌr=i=li.*
       worship   do_make=NMZ1   after   have.to_fall=PL=NEG

       'Worshiping from Simi and Bhumi to trees and such, wasn't required later.'
       CH_MKW_BBC_SIL_032920_1_Chepang people

(175)   *manta=lʌm*   *mu=na=i*   *dʌgi=ma !*
       person=PL   COP=NPST=PL   nowadays=ADD

       'There are people nowadays too!'
       CH_MKW_SC_SIL_120619_4_E_2

(176)   *ʌni*   *gʌjda=lʌm=ma*   *kwejkwej*   *bela*   *ʌ,*
       then   rhinoceros=PL=ADD   some   moment   uh

       o   *Khaptyauni*   *to=o*   *tʰãw,*   *Lothar*
       DIST   Khaptyauni   tell_say=NMZ:REL   place   Lothar

       bʌdʑar=bʰʌnda   *tjaw=taŋ,*   *ʌ,*   *waŋ=i=to.*
       market=CMPR   up=ALL   uh   come=PL=REM.PST

       'So, rhinoceroses as well sometimes, uh, they used to come to that place called Khaptyauni, above the Lothar bazar.'
       CH_MKW_SCBKC_SIL_081918_5_Hunting

(177)   *ŋa=kaj*   *bʰandʑa-tɕoʔ=lʌm*   *sʌpej*   *krus=laŋ*   *waŋ=ka=i.*
       1SG=DAT   nephew=PL   all   meet=pur   come=2/3.PST=PL

'The nephews, all came to meet me.'
CH_MKW_PMRC_SIL_081818_1_Life

The morpheme =*lʌm* can be encliticized to proper nouns or kinship terms to refer to the extended family members or relatives of the referent, as in (178). This function of =*lʌm*, which recalls that of similative plural with inanimate referents, is called "associative plural" (Corbett & Mithun 1996; Moravcsik 2004; Daniel & Moravcsik 2005; 2013; Mauri & Sansò 2019; 2021).

(178) *Kamala=lʌm   waŋ=ti   mu=na=i ?*
        Kamala=ASS   come=SEQ1   COP=NPST=PL

'Are Kamala and her family coming?'
CH_CTW_SPC_SIL_E

It is possible that Chepang has developed the functions of similative and associative plural under the influence of Nepali which uses the plural form हरु <haru> for the same functions, as shown in (179) with the similative plural and in (180) with the associative plural.

Nepali

(179) तिम्रो   किताब्हरु   कहाँ ?
        timro   *kitāb=haru   kahã̃ ?*
        2SG.POSS.NHON   book=SML   where

'Where are your books and stuff?'

Nepali

(180) कमलाहरु   आयो।

334

Kamalā=haru āyo.

Kamala=ASS   arrive.PST.3SG

'Kamala and her family arrived.'

Associative and similative plurals are nevertheless attested in other TH languages, such as Magar (Grunow-Hårsta 2008: 105–106), Limbu (Driem 1987: 30), Yakkha (Schackow 2015b: 124–125) and Japhug (Jacques 2021: 332) for the associative plural, and Karbi (Konnerth 2014: 322, 575–578) and Chintang (Paudyal 2015: 35–36) for both the associative and similative plural.

The associative and similative function of *=lʌm* can be considered derivational since it creates a form that expresses a different type of referent, a new denotation.

In fact, with nouns denoting abstract concepts, the morpheme *=lʌm* functions as an agentive plural nominalizer that express a type of people, such as soldiers, policemen, or linguists, as illustrated in (181) and (182).

(181) *o=haŋ,*   *o*   *Ratan=ko*   *talaŋ=paj*   *djahn=ti*
      DIST=LOC1   DIST   Ratan=GEN   head=DIS   return=SEQ1

      o   *sena=lʌm=kaj,*   *pulis=lʌm=kaj*   *djahn=ti*
      DIST   army=PL=DAT   police=PL=DAT   return=SEQ1

      ŋiʔ=ti   *ŋiʔ=ti*   *ton=ti=taŋ*   *dzal=a.*
      laugh=SEQ1   laugh=SEQ1   fall=SEQ1=ATT   go_leave=PST

      'There, the head of that king Ratan was falling (down the ravine) looking back at
      the soldiers and at the policemen while laughing.'
      CH_MKW_PSC_MAI_012720_Local_History_1

(182) *bʰasa*   *bigjan=lʌm=i*   *ʌdʰjan*   *dzahŋ=dʌj*
      language   science=PL=ERG   research   do_make=PROG

mʌ=na=i.

COP=NPST=PL

'The linguists are doing research.'

CH_MKW_BBC_SIL_032820_3_Lipi

3.4.4. Derivational genitive and locational nominalizers

In this section, I describe the derivational genitive morpheme *=ko* and locational morpheme *=kʌ*. They derive new denotations and are analyzed as nominal-based nominalizations (Zhu 1982; Shibatani 2019).

Unlike other case markers, the genitive *=ko* does not encode a category that expresses grammatical information relating to the argument structure of a clause. The genitive serves the possessive modification of the head of a noun-phrase which can either be overtly expressed or absent if its referent is understood contextually. The genitive construction does not denote the referent of the noun it attaches to, but a new referent, that of the modified noun.

The locational nominalizer *=kʌ* similarly modifies the head of a noun-phrase, which is not overtly expressed when denoting the usual living location of the referent or the referent itself as a location. The function of the head of the modifying noun-phrase is that of the possessor of the denoted location, and it is likely that the morpheme *=kʌ* historically developed from the genitive morpheme *=ko*. The nominal bound denonymic suffix *-mʌj* (§ 3.3.3), which likely come from a noun meaning 'people,' can also attach to a *=kʌ* construction to denote the family members of the referent expressed by the modifying noun-phrase, i.e., people sharing the same living place. In such construction, the form *-mʌj* holds an underlying syntactic position of head of a modified head the same way the semantic head of a noun-noun type of determinative or descriptive compound is modified by a nominal.

In the process of modifying or restricting the denotation of the head of a noun-phrase, the genitive morpheme *=ko* (§ 3.4.4.1) functions as a derivational morpheme that creates a new denotation. The same applies for the derivational locational morpheme *=kʌ* (§ 3.4.4.2), which derives a new denotation, i.e., the usual living place of the referent.

3.4.4.1.　　　Genitive =*ko*

　　　The genitive morpheme =*ko* is analyzed as a noun-based nominalizer, deriving new denotations. It marks the possessive relationship between a possessor noun-phrase and the head noun of a larger embedding noun-phrase. This is illustrated in (183) with a single =*ko* modifying construction, and with three =*ko* modifying constructions embedded in another one, as in (184).

(183)　*sahu=ko*　　　　　　*kʰet=haŋ*　　　*jam*　*suk=ŋʌ=to.*
　　　employer=GEN　　　　field=LOC1　　rice　plant=1=REM.PST

　　　'I planted rice in the field of the employer.'
　　　CH_MKW_BMB_BAN_103119_4_Meeting with Sahu

(184)　*ʌ,*　　*dʑʌstʌj,*　　　*Chitwan=ko*　　　　*Chepang=lʌm,*
　　　uh　　that.is　　　　Chitwan=GEN　　　　Chepang=PL

　　　Makawanpur=ko　　　*Chepang=lʌm,*
　　　Makawanpur=GEN　　Chepang=PL

　　　Dhading=ko　　　　*Chepang=lʌm =ko*
　　　Dhading=GEN　　　　Chepang=PL=GEN

　　　jat=dʑo　　　*ɾadʑa=taŋ*　　　　　*kʰe=to.*
　　　one=CL1　　　king=ATT　　　　　　COP=REM.PST

　　　'Uh, that is, it is said that the Chepangs of Chitwan, the Chepangs of Makawanpur, the Chepangs of Dhading had one king.'
　　　lit. 'Uh, that is, it is that said that there was one king of the Chepangs of Chitwan, the Chepangs of Makawanpur, the Chepangs of Dhading.'
　　　CH_MKW_SCBKC_SIL_081918_2_Chepang_king

　　　The possessor noun-phrase often syntactically precedes the modified head noun, but it can as well follow it, as in (185).

(185) ʌni   pʰeɾi,   tɕoʔ   si=a,   at   bʌɾsʌ=ko.
       then   again   son   die=PST   eight   year=GEN

‘And again, my son died, of eight years old.’

CH_MKW_BMB_BAN_090118_8_Marriage

The semantic types of relationships marked by *=ko* are diverse. It can relate for instance to a kin relation, as in (186) and (187), to body parts, as in (188) and (189), to ownership of concrete objects and abstract concepts, as in (190) and (191), respectively, or to locations, as in (192), or people, as in (193).

(186) ŋa=ko   badʑja=ko   kim   kʰe=to.
       1SG=GEN   elder.M=GEN house COP=REM.PST

‘There was the house of my grandfather.’

CH_MKW_LC_SIL_113019_1_Cave

(187) soɾa   bʌɾsʌ=haŋ,   ŋa=ko   tɕoʔ-djaŋ   naʔ=a.
       sixteen   year=LOC1   1SG=GEN   daughter   be.born=PST

‘At the age of sixteen, my daughter was born.’

CH_MKW_BMB_BAN_090118_8_Marriage

(188) ŋaʔ=ko   tuk=kʰa=taŋ   pok=ti   mu=o (...)
       fish=GEN   stomach=LOC2=ATT enter=SEQ1   COP=PERF

‘(The frog) entered in the stomach of the fish and was sitting there, (…)’

CH_CTW_SP_POL_111420_Ream_Tokrak

(189) juʔ=ko   sʌjk=l=o   tahŋ=to
       mouse=GEN   tooth=COP=NMZ:REL be.huge_be.like=NMZ:ADV2

sjaw=na,   kja !

become=NPST PART

'It becomes like the teeth of a mouse, imagine!'

CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

(190)   ʌ   *Golebang*   *to=o*   *tʰãw=haŋ=ma*
        uh   Golebang   tell_say=NMZ:REL   place=LOC1=ADD

        tɕiŋ-lan=ko   *kim*   *le,*
        Cing-spirit=GEN   house   COP

        'Uh, there is a house of the Cing spirit in a place called Golebang also.'

CH_CTW_BBC_POL_111720_3_Cing_Lan

(191)   *ŋa=ko*   *miŋ,*   *Bir Bahadur Chepang.*
        1SG=GEN   name   Bir Bahadur Chepang

        'My name, Bir Bahadur Chepang.'

CH_CTW_BBC_PID_011520_1_Being_Shaman

(192)   *tʌɾʌ*   *ŋi=ko*   *Polkim*   *dekʰin*   *leʔ=ti,*
        but   1PL=GEN   Polkim   from   take_buy=SEQ1

        'But having taken (all the villages) starting with our Polkim village…'

CH_CTW_BBC_POL_102420_3_Chepang_Kings

(193)   *bʌtɕtɕapʌn*   *dekʰi=le*   *maha*   *pe=to*   *sʌjli*
        childhood   since=DIS   very   be.good=NMZ:ADV2   speaking.style

        Pambung=ko   *Chepang=lʌm=kusi*   *mu=na.*
        Pambung=GEN   Chepang=PL=COM   COP=NPST

        'Since their early age, the Chepangs of Pambung have a genuinely nice speaking
        style.'

CH_MKW_BBC_SIL_032820_2_Pambung

The possessor marked with =*ko* can narrow down the sub-type of a semantic

category, as shown in (194) with a type of porridge, in (195) with a type of story, and in

(196) with a type of meat. It can also express the locational origin of a person, as in (197), or else it be used to talk about a quantity over a period of time, as in (198).

(194) *tɕui=ko*   *ahm*   *dʑe=gaɾ=ja=paj*
       rice.uncooked=GEN   porridge   eat=DES=COND=DIS

       bʌdʑaɾ=sʌj=le   *waʔn=sa*   *pʌɾ=na.*
       town=ABL=DIS   take.away=NMZ1   have.to_fall=NPST

       'If one wants to eat rice porridge, this is from town that one has to get it.'
       CH_MKW_SRNDC_SIL_081818_Life

(195) *Nak=ko*   *juin-ɾaj*   *ta=ti*   *bʌj=na=ŋ*   *da,*   *lʌw !*
       Nak=GEN   story   tell.story=SEQ1   give=NPST=1   PART   PART

       'I'm going to tell (you) the story of the Nak, hey, let's do it!'
       CH_MKW_SC_SIL_122619_6_Agreement

(196) *bʰuɾa=kʰa*   *dʑe=o,*   *dʌge*   *sʌmmʌ*   *dʑe=o*   *na=lʌ,*
       child=LOC2   eat=PERF   now   until   eat=PERF   COP=NEG

       juin=ko   *maj.*
       bat=GEN   meat

       'In my childhood, I had eaten (some), until now, I haven't eaten any, bat's meat.'
       CH_MKW_SC_SIL_120619_2_E_1

(197) *ʌ*   *i*   *Syamrang=ko*   *Ganamani*   *ɾadʑa*   *ɾʌ (…)*
       uh,   PROX   Syamrang=GEN   Ganamani   king   and

       'Uh, this king Ganamuni of Syamrang and (…)'
       CH_CTW_BBC_POL_102520_1_Polkim

(198) *ŋa*   *bʌɾsʌ=ko*   *jat*   *pʌltʌ*   *ʌ,*
       1SG   year=GEN   one   time   uh

ŋa=i	*puʣa*	*ʥahŋ=na=ŋ.*
1SG=ERG	ritual	do_make=NPST=1

'I do a worshiping ceremony (for them) once a year.'
CH_CTW_BBC_PID_011520_5_Underworld

The possessively modified head may not be overtly expressed when recoverable contextually, leaving just the *=ko* marked phrase. This is illustrated in (199) where the construction denotes the spirit that leaves the body of the Chepang people after death.

(199)	*sʌppʌj=ko=le*	*gljuhŋ=na,*	*Broso=ko=ma*
	all=GEN=DIS	come.out=NPST	Broso=GEN=ADD

	gljuhŋ=na,	*Gaurung=ko=ma*	*gljuhŋ=na, (...)*
	come.out=NPST	Gaurung=GEN=ADD	come.out=NPST

'That of everyone comes out, that of the Broso comes out too, that of the Gaurung comes out too, (…)'
CH_CTW_BBC_PID_011520_6_Tiger_Spirit

Finally, the morpheme *=ko* can encliticize to nominalized clauses formed with the nominalizer morpheme *=o*, as in (200).

(200)	*ʌɾu=ko*	*budʰi*	*aʔl=o=ko*	*ʥaɾi*
	other=GEN	wife	take.away=NMZ:REL=GEN	adultery.fine

	pajk=sa	*jat=lakʰ*	*ɾupja.*
	pay=NMZ1	one=lakh	rupee

'It was one lakh rupee to pay the adultary fine of the one who takes away someone else's wife.'
CH_CTW_SRP_GUN_102620_4_Local_justice

3.4.4.2. Locational *=kʌ*

The derivational locational morpheme *=kʌ* is a noun-based nominalizer, i.e., it occurs attaches to nouns that refer to a human or an animal, or any living entity, to denote the usual living place of the referent, as in (201). Metaphorically, a *=kʌ* construction can also convey the meaning of 'amongst' or 'at' when the noun refers to a community, as in (202), or a species or group, as in (203), or to a location associated with a figure representing an ideology, as in (204). It can also merely denote a living entity as a location itself, as in (205) and (206).

A *=kʌ* construction functions as a locational adverbial or oblique argument, which can further be encliticized with locational case markers. This is shown with the locative markers *=haŋ* and *=kʰa* in (207) and (208), respectively, with the ablative marker *=sʌj*, as in (209), and the allative marker *=taŋ*, as in (210).

(201) *o=kʌ=haŋ*   *koko-ɾo*   *te=laŋ*   *al=na=ŋ, (...)*
       DIST=NMZ:LOC=LOC1   proso.millet   ask.for=PUR   go=NPST=1

'I'll go to her/his place to ask for Proso Millet,'

CH_MKW_STC_SIL_120619_2_E_2

(202) *Chepang=kʌ=haŋ*   *mu=o*   *pande=lʌm,*
       Chepang=NMZ:LOC=LOC1   COP=NMZ:REL   shaman=PL

'The shamans that are amongst the Chepangs,'

CH_MKW_BBC_SIL_032920_2_My grandfather

(203) *tibiliŋ=kʌ=kʰa,*   *jadʑo*   *maha   pe=to*
       butterfly=NMZ:LOC=LOC2   one-CL   very   be.nice=NMZ:ADV2

       du=to   *tibiliŋ, (...)*
       be.red=NMZ:ADV2   butterfly

'Amongst the butterflies, (there is) one very nice red butterfly, (…)'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(204) *ŋa   Prabhu=kʌ=haŋ*   *pok=alaŋ.*
       1SG   lord_god=NMZ:LOC=LOC1   enter=1.PST

342

'I became Christian.'

lit. 'I entered in the lord's place.'

CH_MKW_BMB_BAN_103119_1_Church

(205)   *ŋa=ko*   *buda=kʌ=sʌj*   *la=bʌtiko,*
       1SG=GEN   husband=NMZ:LOC=ABL   grab=SEQ2

'Having grabbed it (letter of acknowledgment of debt) from my husband,'

CH_MKW_SC_SIL_010220_1_Life

(206)   *ʌni*   *o*   *bʌŋsʌs=lʌm=kʌ=sʌj=ma*   *Christian*   *dʰʌrmʌ=ko*
       then   DIST   ancestor=PL=NMZ:LOC=ABL=ADD   Christian   religion=GEN

       bare=haŋ   *budʑʰʌ=dʌj*   *mu=o*   *rʌjsʌ,*   *o-mʌj.*

       subject=LOC1   understand=PROG   COP=PERF   COP.MIR   DIST-PL.H

'And it turns out that they are learning about the Christian religion from those ancestors.'

CH_MKW_BBC_SIL_032920_3_Origin_Christianity

(207)   *(...), pande=kʌ=haŋ*   *al=sa*   *pʌr=ʌ=to.*
       shaman=NMZ:LOC=LOC1   go=NMZ1   have.to_fall=LN=REM.PST

'(…), one had to go the shaman's.'

CH_CTW_SRP_GUN_102620_5_Animism_to_Christianity

(208)   *dukʰʌ*   *tɕejew=ka=nʌ,*   *pʰeɾi*   *ŋa=ko*
       pain   see_find=2/3.PST=DIR/TR   again   1SG=GEN

       majti=kʌ=kʰa=le   *dah=ti*   *mu=na, (...)*

       maternal.home=NMZ:LOC=LOC2=DIS   reach=SEQ1   COP=NPST

'She found sorrow, then she is reaching my mum's house, (…)'

CH_MKW_PMRC_SIL_081818_1_Life

(209)   *baba*   *ɾʌ*   *kantɕʰi-ama=kʌ=sʌj*
       father   and   mother.in.law=NMZ:LOC=ABL

bʌʥe=kʌ=haŋ

grandmother=NMZ:LOC=LOC1

tɕʰut=ʌ=ti   *mu=o*   *bela=haŋ, (...)*

separate=LN=SEQ1   COP=NMZ:REL   time=LOC1

'At the time when you left your father and mother-in-law's home and lived at your grandmother's, (…)'

CH_MKW_SPMC_LC_SIL_100921_3_Conversation

(210) *didi,*   *waŋ=ʌ,*   *ŋa=kʌ=taŋ, (...)*
       elder.sister   come=2SG.IMP.INTR   1SG=NMZ:LOC=ALL

'Elder sister, come, to my place, (…)'

CH_CTW_BBC_POL_111720_5_Two Sisters

In combination with the verb *pahj-* 'leave,' a *=kʌ* construction specifically means 'get married'[43]. Note that no locational case marker is used in this construction. This is illustrated in (211) and (212).

(211) *ʥetʰi-tɕoʔ=ma*   *bʰaw=kʌ*   *pahj=waj=a.*
       elder.F-child=ADD   husband=NMZ:LOC   leave=CERT=PST

'My elder daughter too got married.'

lit. "My elder daughter too left to her husband's place.'

CH_MKW_SC_SIL_010220_1_Life

(212) *ŋa*   *bʰaw=kʌ*   *pahj=ma=kʰa=le,*   *baba*   *si=a.*
       1SG   husband=NMZ:LOC   leave=NEG=LOC2=DIS   father   die=PST

---

[43] It is current that a woman leaves her house to go live at her husband's place after marriage.

'My father died just before I got married to my husband.'

lit. 'My father died just before I left to my husband's place.'

CH_MKW_RC_JMC_SIL_120119_Conversation

When the morpheme *=kʌ* is encliticized with the locative case marker *=haŋ*, it results in a construction that recalls phonologically and functionally a similar construction that exists in Nepali. The Nepali forms कहाँ <kahāṃ> also denotes the usual living place of the referent, as shown in (213).

It seems unlikely that Chepang borrowed this construction. The morpheme *=haŋ* is not the only locational case marker that can attach to *=kʌ*. As seen, the locative morpheme *=kʰa*, ablative *=sʌj*, or else allative *=taŋ*, can also occur in a *=kʌ* construction, as in (214) and (215), respectively. In addition, unlike Chepang, the Nepali कहाँ <kahāṃ> construction only denotes the living place of the referent.

Nepali

(213) म   साथिकहाँ   जान्छु।
     ma   *sāthi=kahã̃~kã̃*   *jān=chu.*

     1SG   friend=NMZ:LOC   go_leave=1SG.NPST

     'I'm going to my friend's.'

     (Schmidt 1994: 92)

(214) *baba=kʌ=kʰa*   *dah=laŋ*   *al=a*   *pʌtɕʰi, (…)*
     father=NMZ:LOC=LOC2   reach=PUR   go=NMZ2   after

     'After having gone to arrive at the father's place.'

     CH_MKW_SC_SIL_122619_1_Nakko_Co'

(215) *dzʌnta=kʌ=sʌj*   *utʰaw=o*   *lebʰi=taŋ,*   *ʌ,*
     people=NMZ:LOC=ABL   raise=NMZ:REL   tax=ATT   uh

tjaw	*ɾadʑa=kaj*	*budʑʰaw=n=i=to,*
upwards	king=DAT	pay=DIR/TR=PL=REM.PST

tahŋ=o		*ɾadʑa=kaj.*
be.huge_be.like=NMZ:REL		king=DAT

'The tax raised from people's home, uh, they would pay it to the king of upwards, a powerful king.'

CH_MKW_SCBKC_SIL_081918_2_Chepang_king

Finally, the morpheme =*kʌ* occurs as well with nominalized verb to form a temporal subordinate construction, as illustrated in (216).

In this construction, the adverbial manner morpheme =*tʌ* attaches to the verb as a nominalizer and the locational derivational nominalizer morpheme =*kʌ* turns this form into a location, i.e., the place where the process of the verb takes place, which semantically metaphorically relates to time since it is a verbal process, such as: *ljuhŋ=tʌ=kʌ* 'at the place of getting out' is understood as 'when getting out.'

This is further evidence that the morpheme =*kʌ* is a native locational derivational morpheme and that its use has not developed in a construction calqued from Nepali.

(216)
i=tʌ,	*i=kʰa=sʌj*	*ljuhŋ=tʌ=kʌ*
PROX=NMZ:ADV1	PROX=LOC1=ABL	get.out=NMZ:ADV1=NMZ:LOC

kaʔm-kʰola	*dah=o,*	*oɾalo.*
downwards-river	reach=PERF	downhill

'It's like this, when you get out from here and have reached the downwards river, it's downhill / it goes down.'

CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

3.4.5. Case markers

This section is dedicated to the description of Chepang case markers and how they relate to verbal argument structure. Before I describe each of them, I briefly discuss the notions of syntactic relations and alignment.

Case markers express the syntactic functions of arguments, specifying their syntactic roles or relations to the verb or predicate in a clause. Case marking specifies, through morphosyntactic flagging, the roles that arguments play in relation to a predicate in a particular event or situation. The nature of these syntactic relations may be grammatical or syntactic, i.e., core argument, or adverbial, i.e., adverbial, or oblique argument. Adverbial or oblique phrases can also be marked through the presence of relator nouns or postpositions (§ 3.4.6). In Chepang, syntactic relations are also indicated through argument indexation on the verb (§ 5.8.3).

Chepang case markers are all enclitics, i.e., they attach a noun-phrase. Chepang has nine enclitic case markers: absolutive (§ 3.4.5.1), ergative (§ 3.4.5.3), dative (§ 3.4.5.2), instrumental (§ 3.4.5.4), comitative (§ 3.4.5.5), two locatives (§ 3.4.5.6), ablative (§ 3.4.5.7), and allative (§ 3.4.5.8). They are presented in Table 112.

Table 112. Enclitic case markers

case	gloss	form
absolutive	unmarked	
ergative	ERG	$=i$
dative	DAT	$=kaj$
instrumental	INST	$=i$
comitative	COM	$=kusi$
locative 1	LOC1	$=haŋ$
locative 2	LOC2	$=kʰa$
ablative	ABL	$=sʌj$
allative	ALL	$=taŋ$

Core arguments are the expression of a predicate in a clause. A predicate consists of a verb or the combination of a non-verbal form and a copula (in the case of non-verbal predication) whose semantics comprises participants that take on certain roles for its

process to be realized, consisting of the semantic valency and syntactic argument structure of the verb. In Chepang, core arguments may be overtly expressed, and if absent, they may be recoverable from context, and other means, such as argument indexation on the verb. Core arguments are marked with the absolutive case, the ergative case, and the dative case. Oblique arguments remain optional, adding different kinds of information that are not essential for the process to be semantically complete, such as tool, participants' company, or location. Oblique arguments are marked with the instrumental case, the comitative case, the two locative cases, the ablative case and the allative case.

To explore Chepang syntactic relations and alignment patterns of core arguments, I use the generalized participants' semantic roles or primitives of S, A, P, T, and R, that traditionally go back to Comrie (1978; 1989), Dixon (1979; 1994), and Croft (1990).

Since the type of predicate core arguments described here includes a variety of intransitive, transitive and ditransitive predicate constructions, these primitives are used as follows:

- S for the Single argument of an intransitive predicate;
- T for the Theme argument, refering to the third argument of a ditransitive predicate besides A and P, or A and R;
- A for the most Agent-like argument of a transitive or ditransitive predicate;
- P for the most Patient-like argument of a transitive or ditransitive predicate;
- R for the Recipient-like argument of a ditransitive predicate.

Chepang can be characterized as exhibiting an ergative-absolutive alignment system when it comes to case marking: the A argument is marked with the ergative case, while S and P are marked differently, i.e., left unmarked in the absolutive case. In some constructions it can also show a tripartite alignment, where all S, A and P are all marked differently.

With regard to the marking of P, T, and R arguments, Chepang features indirective, secundative, and neutral alignments (Dryer 1986; Dryer 1997).

Indirective alignment characterizes a pattern where the P argument of a transitive verb and the T argument of a ditransitive verb are marked the same way and the R

argument differently. This applies in Chepang when the P argument of a transitive verb and the T argument of a ditransitive verb are left unmarked in the absolutive case, while the R argument of a ditransitive verb is marked with the dative case.

Secundative alignment is when both the P of a transitive verb and the R of a ditransitive verb are marked the same way, while the T argument of a ditransitive is marked differently. In Chepang, both the P argument of a transitive verb and the R argument of a ditransitive verb can be marked with the dative case, while the T argument of a ditransitive verb can be left unmarked in the absolutive case.

In a neutral alignment, P, T, and R arguments are all marked the same way. In Chepang, the P argument of a transitive verb and the T and R arguments of a ditransitive verb may be all marked with the dative case. In such configuration, the P, T, and R may all be considered O arguments, i.e., the Other argument besides A of a transitive or ditransitive verb, since they may all be marked with the dative case or indexed the same way on the verb.

These alignment patterns are generalizations based on transitive and ditransitive constructions compared with one another. What remains constant over these three alignment patterns is the treatment of the R argument, always marked with the dative case, while splits are observed in the treatment of P and T arguments. Such splits are known as Differential Object Marking (Moravcsik 1978; Bossong 1985, 1998; Lazard 2001), tied to the referent's degree of animacy, definiteness, and pragmatic saliency. These triggering factors are described in the literature under the terms of referential prominence, discourse prominence, animacy hierarchy, or else empathy hierarchy (Aissen 1999; 2003; Shibatani 2006; Bornkessel-Schlesewsky & Schlesewsky 2009; Lockwood & Macaulay 2012; Malchukov 2008; Seržant & Witzlack-Makarevich 2018; Haspelmath 2021).

Besides the treatment of P, T, and R, differential case marking triggered by pragmatic forces is also attested with S arguments which may be marked like A arguments of (di)transitive verbs. This is known as Differential Agent Marking (Barðdal & Chelliah 2009; McGregor 2010; McGregor & Verstaete 2010; Fauconnier 2011) and common in TH languages (LaPolla 1995; Chelliah 2009; DeLancey 2011).

In addition to transitive and ditransitive constructions, experiencer and stimulus arguments take dative or absolutive markers. For intance, the S argument of an intransitive verb may be marked with the dative case in dative experiencer constructions borrowed from Nepali; or else the stimulus of another type of native experiencer construction may be marked with the dative case while the experiencer is left unmarked in the absolutive case.

Clearly, when it comes to accounting for all the types of case alignments or argument structures attested amongst all type of verbs and constructions, in addition to humanness, animacy and definiteness, or pragmatic forces that may entail changes in the marking of the arguments, the descriptive concepts provided by the typology of alignment systems seem insufficient.

I summarize in Table 113 the possible marking of core-arguments as described in the following sub-sections.

Table 113. Case-marking and alignments

referent	human definite indefinite	animate definite	inanimate definite	animate indefinite	inanimate indefinite
argument					
A	ERG	ERG	ERG	ERG	ERG
S	Ø/DAT	Ø	Ø	Ø	Ø
P	DAT/ABS	DAT	DAT	DAT	ABS/DAT
T	ABS/DAT	ABS/DAT	ABS	ABS/DAT	ABS
R	DAT	DAT	DAT	DAT	DAT
partly attested alignment patterns					
ergative-absolutive	A ≠ S, P				
tripartite	A ≠ S ≠ P				
indirective	R ≠ P, T				
secundative	T ≠ P, R				
neutral	R = P = T				

To understand the mechanisms at play behind the marking of predicates' arguments, I first considered verbal indexation. The reason for that is that Chepang verbal indexation is the primary morphosyntactic cue to the understanding of the argument

structure of a predicate and the semantic and syntactic valency potential of a verb. In addition to verbal indexation, derivational morphology is the second important morphosyntactic cue that confirms the semantic and syntactic valency potential of a verb.

These criteria help in understanding the type of constructions that can be used with a verb. By type of construction, I mean the possibly expressible argument structures: the number of arguments, their semantic or syntactic roles, the semantic nature of their participants (human, animate, inanimate), and what are the arguments that can occur or must occur, i.e., be overtly expressed; arguments that can occur may also not be expressed if recoverable through context or verbal indexation. It further helps to

Then, as a last step, clear observations can be made on the type of morphological endings that occur on the verbs; indeed, it exists pairs, triplets or more members of series that show traces of old derivational morphology; understanding these series sheds light on the type of verbs originally used in absence of derivational morphology and on the development of these verbs that led to a multiplicity of construction patterns observed in in the language spoken today.

In the following sub-sections, I describe case marking found on both core and oblique arguments, expressed through enclitics, in addition to postpositions and relator nouns.

The types of verbal constructions described here, which feature core arguments, along with example verbs, are summarized in Table 114.

Table 114. Core arguments constructions

	verb type	construction	verb	meaning
intr 1	intr dyn/st S	S ABS	*al-*	'go'

	verb type	construction	verb	meaning
			eʔn-	'sleep'
			si-	'die'
			waŋ-	'come'
			bʰam-	'be white'
			bɾaw-	'be big'
	intr SEXP	SEXP ABS Stim DAT	ɾʌj-	'be afraid, fear'
			kɾis-	'be disgusted'
	intr SEXP	SEXP DAT (Stim ABS)	waŋ-	'come'
			tʰaha mu-	'know'
	intr refl S	S ABS	boɾ-	'coil self up'
transitive & ditransitive	tr	A ERG P ABS A ERG P DAT	lat- aʔl- ɾa- ʣe- waʔn- leʔ- pʰe- lʰo- ʣahŋ- kʌn- boŋ- sat-	'carry' 'take away' 'cut (weed), saw' 'eat' 'bring' 'take, buy' 'leave, abandon' 'chop off' 'do, make' 'look at' 'look for' 'kill'
	ditr	A ERG T ABS R DAT A ERG T DAT R DAT	bʌj- te- kas- tan-	'give' 'ask for, beg' 'feed' 'show'
specials	lexicalized intr dyn *or* tr P-labile	S ABS [N ABS] Vintr A ABS P ABS	ti laʔ- ɾew maʔŋ- pes peʔ- eʔn bus- tuk kos-	'swim' 'dream' 'fart' 'be/feel sleepy' 'be filled up'
	lexicalized intr dyn *or* tr A-labile	S ABS [N ABS] Vintr A ABS (P ABS)	noʔ- ges- has- wah-	'speak' 'play' 'vomit' 'walk, go'
	verb type	**construction**	**verb**	**meaning**

lexicalized intr refl PSSR	PSSR ABS [N ABS] Vintr	*kɾjuŋ-*	'fold (arm, leg)'
		ŋoltum gɾjus-	'sit on knees'
		~grus-	
labile refl PSSR	PSSR ABS [N ABS] Vintr	*dʑis-*	'wash (body, body parts)'
	A ERG P DAT [N ABS] Vtr	*nuhl-~nuh-*	'rub (body, body parts)'
		mjaŋ tis-	'plait hair'
		mjaŋ rʌʔj-	'wash hair'
		mjaŋ sjal-	'tidy hair'
		pit-	'pinch'
lexicalized intr (refl) PSSR	PSSR ABS [N ABS] Vintr	*dʑik-*	'hurt'
	PSSR GEN [N ABS] Vintr	*lʰis-*	'come off (nail, cuticle)'
	PSSR ABS [N ABS] Vintr	*dut kʰas-*	'leak (breast milk)'
	(generic)	*klju-~klu-*	'fall off (hair, hairs)'
lexicalized intr (refl) PSSR labile	PSSR ABS [N ABS] Vintr	*tuk bljʌw-*	'be nauseated'
	PSSR DAT [N ABS] Vintr	*lʰuŋ blʌn-*	'have indigestion'
		sos-	'be itchy'

3.4.5.1.   Absolutive (unmarked)

The absolutive case is traditionally associated with the marking of the S argument of an intransitive verb and the P argument of a transitive verb in languages where the A argument of a transitive verb is marked differently, with the ergative case. This is illustrated in (217) to (219).

Chepang cannot be restrictively defined as displaying an ergative-absolutive alignment system. This alignment merely corresponds to the comparison of two specific types of constructions: an intransitive construction where the S argument is marked with the absolutive, as in (217), and a transitive construction where the A is marked with the ergative and the P with the absolutive, as in (220). However, as we will see, the P argument may as well be marked with the dative case, as in (219), and, though rare, the ergative case marker may be absent, as in (221), leaving the A argument in the absolutive case.

(217)   *ni=ko*   *tʰaw=haŋ,*   *dʑaʔ*   *waŋ=na.*
        1PL=GEN   place=LOC1   tiger   come=NPST

'In our area, tigers come.'

CH_MKW_PMRC_LAM_081618_3_Kalitar

(218) *gundʌɾi*   *nap=ti*   *kapi-kʌlʌm*   *kʌn=na=tɕ=u.*
mat   spread=SEQ1   copy-pen   look=NPST=1/3DU=3O/DIR

'Spreading the mat, they two look at their school stuffs.'

CH_MKW_BBC_SIL_042420_6_My two daughters

(219) *pʌhila pʌhila*   *ama=i*   *ŋi=kaj*   *sjas=u=to.*
before before   mother=ERG   1PL=DAT   raise=3O/DIR=REM.PST

'A long time ago, mother raised us.'

CH_MKW_SC_SIL_010220_2_Life

(220) *ʌhʌ,*   *naŋ=i*   *ama*   *sat=te=aka=n,*
EXPR   2SG=ERG   mother   kill=2=2/3.PST=DIR/TR

hʌw=jʌ !
younger.brother=VOC

'What have you done, you killed mother, younger brother!'

CH_MKW_DBC_MAI_1_020320_The two brothers

(221) *mʌdʌt*   *te=ti,*   *ŋi*   *tɕʰak-dʑaʔ*
help   ask.for=SEQ1   1PL   dead.tiger.spirit

sat=ŋʌ=su=to,   *pande*
kill=1=1PL.EXCL=REM.PST   shaman

tɕʰak-dʑaʔ   *sat=n=i=to.*
dead.tiger.spirit   kill=DIR/TR=PL=REM.PST

'Asking for help (of other spirits, gurus, gods), we used to kill the Tiger spirit of the dead, the shaman used to kill the Tiger spirit of the dead.'

CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

3.4.5.1.1.   S of intransitive verbs

In Chepang, the S argument of an intransitive verb is left unmarked, in the absolute case. This is further illustrated with S arguments referring to a human, as with *ŋa* '1SG' in (222), to an animate non-human, as with *tokɾak* 'toad' in (223), and to an inanimate, as with *siŋ* 'tree, wood' in (224). Some stative verbs may also function as a verbal predicate and have their S argument marked with the absolutive case, as in (225) and (226).

(222)   *"ŋa*   *al=ŋʌ=lʌ*   *ama,*   *al=ŋʌ=lʌ*   *ama" =taŋ*
         1SG   go=1=NEG   mother   go=1=NEG   mother=ATT

         to=o   *kʰe=to.*
         tell_say=PERF   COP=REM.PST

         '"I won't go, mother, I won't go, mother!" she had told her.'
         CH_MKW_SC_SIL_122619_1_Nakko_Co'

(223)   *tokɾak*   *dʑos=ti*   *al=a.*
         toad   jump=SEQ1   go=PST

         'The toad went away jumping.'
         CH_MKW_SC_SIL_120619_4_E_2

(224)   *dʑʌŋʌl=haŋ*   *siŋ=le*   *na=lʌ.*
         jungle=LOC1   tree.wood=DIS   COP=NEG

         'There is no wood in the jungle.'
         CH_MKW_SC_SIL_122619_2_Dhobini_rani

(225)   *ɾʌ*   *Dagu*   *ɾadʑa=ko*   *kliʔ*   *bʰam=a.*
         and   Dagu   king=GEN   shit   be.white=PST

         'And the shit of the king Dagu whitened.'
         CH_CTW_BBC_POL_102420_3_Chepang_Kings

(226)   *i=paj*   *i=tɕjuk*   *mi=to*   *kʰe=to,*
         PROX=DIS   PROX=QTY   be.small=NMZ:ADV2   COP=REM.PST

(...)   dʌŋi   bɾaw=a.

now   be.big=PST

'She was this small, now she got big.'

CH_MKW_CMC_BC_SIL_120619_2_Conversation_Dhirang

3.4.5.1.2.   Noun of intransitive verbal compounds

Some verbs can occur with two core arguments left unmarked in the absolutive case. In such constructions, a P argument may either be optionally overtly expressed, as shown in (227) and (228) with *has-* 'vomit,' and in (229) and (230) with *noʔ-* 'speak, talk,' or be required, as in (231) with *ti laʔ-* 'swim (< *'grab water')' and in (232) and (233) with *eʔn bus-* 'feel sleepy (< *'carry sleep')' by contrast with the S argument which may or may not occur if its referent is recoverable through context or verbal argument indexation.

Such verbs can be analyzed as ambitransitive or transitive labile verbs, since they feature two core arguments and can be used intransitively or transitively; however, their transitive use is merely the result of the presence of two arguments, since the verbal morphology remains intransitive. The presence of transitive morphology on such intransitive verbs is possible but in a causative derivational construction, as in (234).

Since they are intransitive verbs and feature ancient derivational consonants (§ 5.2), they could also be analyzed as verbal compounds or lexicalized verb; as a result of their historical origin and development, such types of derivations often led to the expression of intransitive anticausative constructions with an SP argument.

(227) *ʌni   wak wak   sjaw=na,   has=na   kja !*
       then   ONO ONO   become=NPST vomit=NPST   PART

'Then one becomes nauseated, one vomits if you will!'

CH_CTW_BBC_PID_011520_3_Witches

(228) *o   wiʔ   has=a.*
       DIST   blood   vomit=PST

'He vomited blood.'

CH_CTW_SPC_POL_E

(229) *gu=tɕuk*   *budi=kusi=ma*   *noʔ=alaŋ,*   *ʌ,*
       INT=QTY   wife=COM=ADD   speak=1.PST   uh,

  "ljam=le   *tʰaha*   *na=lʌ*   *buda,*
  path=DIS   know   COP=NEG   husband

  naŋ   *ga=haŋ=te*   *al=te=na ?"*
  2SG   INT=LOC1=2   go=2=NPST

'How much did I speak with my wife too, uh, "You don't know the way husband, where are you gonna go?"'

CH_MKW_SC_PB_BGR_101619_1_Conversation

(230) *ʌ*   *kunʌj* *manta=lʌm* *Chepang*   *kuɾa*   *pʰe=ti*
       uh   some   person=PL   Chepang   language   abandon=SEQ1

  Nepali   *kuɾa*   *noʔ=ka=i.*
  Nepali   language   speak=2/3.PST=PL

'Some people abandonned the Chepang language and spoke the Nepali language.'

CH_MKW_SMC_AYS_1_112819_Conversation_with_Bipana

(231) *o*   *ti*   *laʔ=a.*
       DIST   water   seize=PST

'He swam.'

CH_CTW_SPC_POL_E

(232) *pe=ma=to*   *eʔn*   *bus=ŋʌ=to,*   *ŋa=paj.*
       be.good=NEG=NMZ:ADV2   sleep   carry=1=REM.PST   1SG=DIS

'I used to be so sleepy.'

CH_MKW_STC_SIL_120619_4_E_2

(233) *noʔ=ʌ=ma*   *ane,*   *eʔn*   *bus=te=a ?*
       speak=2SG.IMP.INTR=DIS   PART   sleep   carry=2=PST

'Come on, speak please, are you sleepy?'

(234) doh=ma   dʌj=lʌ,   aʔl=ti,   dʑʌmma   pʰidʑaw=ti,
      what=ADD   happen=NEG   take.away=SEQ1   all   spread=SEQ1

      dʑʌmma   noʔ=tak=pa=n=i.
      all   speak=CAUS=DIR/TR=PL

'There is no problem, having taken away (the recordings), having spread them all,
may (people) make (the recordings) be listened to.'

Finally, the movement verb *wah-* 'walk, go,' may occur with A and P arguments as well marked in the absolutive. The P arguments found in this construction are *ljam* 'path, way' and *kʰola* 'river,' which both represent paths along which one walks. This is illustrated in (235) and (236).

(235) dʑaʔ=le   lʌj=ko   ljam   wah=na,
      tiger=DIS   SLF.INTS=GEN   path   walk=NPST

'The tiger walks its own path,'

(236) kʰolʌ=i   kʰola=le   wah=sa   pʌr=na.
      river=INTS   river=DIS   walk=NMZ1   have.to_fall=NPST

'One has to walk along / walk the river.'

3.4.5.1.3.   S of intransitive reflexive verbs

The S argument of an intransitive reflexive verb is left unmarked in the absolute case. With such type of reflexive verbs, reflexive morphology may occur. This is illustrated in (237) with the verb *boɾ-* 'coil self up.'

(237) ʌ   o=haŋ=le,   sjaʔn,
       uh   DIST=LOC1=DIS   insect

       o=haŋ=le   *boɾ=na=sʌ.*

       DIST=LOC1=DIS   coil.self.up=NPST=REFL

       'Uh, there, the insect, there it coils itself up.'
       CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

3.4.5.1.4.   S of dative experiencer intransitive verb

The dative experiencer constructions illustrated in (238) to (241) are calqued from Nepali, as shown in (242) and (243). These constructions feature the intransitive verb *waŋ-* 'come' and the non-verbal predicate *tʰaha-* 'knowledge' in combination with the copula *mu-*. They express the meaning of the verb 'know.' Within such constructions, the Stimulus may be overtly expressed, left unmarked in the absolute case.

(238) *(…) tʌɾʌ*   *dʑʌnawʌɾ*   *sat=sa*   *ni=kaj*   *waŋ=lʌ.*
       but   wild.animal   kill=NMZ1   1PL=DAT   come=NEG

       '(…), but we don't know how to kill wild animals.'
       CH_MKW_BMB_BAN_100617_4_Jungle_animals

(239) *i*   *sʌppʌj=kaj*   *waŋ=jak=lʌ,*   *ekat*   *pande=kaj*
       PROX   all=DAT   come=REM.PST=NEG   some   shaman=DAT

       mʌtɾʌj   *waŋ=to,*

       only   come=REM.PST

       'Not all would know that, only some shamans would.'
       CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(240) | *baba* | *ke=o* | | *git* | *ŋa=kaj* | *tʰaha* | *mʌ=na.* |
|---|---|---|---|---|---|---|
| father | sing=NMZ:REL | | song | 1SG=DAT | knowledge | COP=NPST |

'I know the songs that my father sang.'

CH_MKW_SC_SIL_010220_3_Life

(241) | *o* | *wa* | *degi,* | *niŋ=kaj* | *tʰaha* | *mu=na* | *ki ?* |
|---|---|---|---|---|---|---|
| DIST | bird | now, | 2PL=DAT | knowledge | COP=NPST | or |

'Now, do you know that bird?'

CH_MKW_PSC_MAI_012720_Local_History_1

(242) | ऊस्लाइ | नेपालि | आउँछ। | |
|---|---|---|---|
| *us=lāi* | *Nepāli* | *āũcha,* | |
| 3SG=DAT | Nepali | come.3SG.NPST | |

'He knows Nepali.'

(Schmidt 1994: 33)

Nepali

(243) | मलाइ | थाह | छैन। |
|---|---|---|
| *ma=lāi* | *thāha* | *chaina* |
| 1SG=DAT | knowledge | COP.NEG.NPST |

'I don't know.'

(Schmidt 1994: 298)

3.4.5.1.5. S and possessed noun of reflexive verbs

Four additional types of intransitive reflexive constructions may also feature the presence of two core arguments: S and a possessed noun left unmarked in the absolutive case.

In all four types, the S argument semantically functions as the possessor of the referent of the possessed noun. Such intransitive reflexive verbs express processes characterized by the interactions between a person and their own body or body parts or someone else's.

The first type of intransitive reflexive verb displays a construction where both S and the possessed noun can be expressed, or the S can be absent if its referent is recoverable through context or verbal argument indexation, as is the case S and A arguments in general. The presence of the possessed noun is required. The verb takes reflexive morphology. The possessor S argument has a semantic role of agent since its referent volitionally acts in the process. This is illustrated in (244) and (245).

(244) *ŋa   ŋoltum   gɾjus=alaŋ=sʌ.*
       1SG   knee   sit.on=1.PST=REFL

       'I sat on my knees.'

       CH_MKW_PC_SIL_E

(245) *o   kɾut   kɾjuŋ=ka=sʌ.*
       DIST   hand.arm   wash.hair=2/3.PST=REFL

       'S/he folded her/his arm.'

       CH_MKW_PC_SIL_E

The second type of reflexive verb can be characterized as ambitransitive or labile, i.e., displaying reflexive lability: it can function as an intransitive reflexive and a transitive verb, as shown in (246) to (247). The possessor S argument has a semantic role of agent since its referent volitionally acts in the process.

In the intransitive reflexive construction, reflexive morphology may occur, as in (246), but speakers may as well drop it, as in (248). When the construction is transitive, the possessor S argument becomes the A argument and is marked with the ergative case, as in (247).

When such reflexive verb is derived with causative morphology, the additional P argument, semantic causee, is marked with the dative and functions as the underlying S argument of the intransitive reflexive verb, as in (249). When a body part is not specified, the noun *ti* 'water' is used, left unmarked in the absolutive case, and holding a semantic role of instrument. This is illustrated in (247).

(246)　ʌ　　　　kɾut-sut　　　dʑis=sʌ,　　　　　　mʰotoŋ-sotoŋ　dʑis=sʌ.
　　　　　uh　　　hand-ECHO　　wash=2SG.IMP.REFL　mouth-ECHO　　wash=2SG.IMP.REFL

　　　　‘Uh, wash your hands, wash your mouth.’

　　　　CH_MKW_BMB_BAN_090118_2_Verb

(247)　waŋ=ʌ　　　　　　　ama,　　　ŋa=i　　　　ti　　　dʑis=tɕe=na=ŋ.
　　　　come=2SG.IMP.INTR　mother　　1SG=ERG　　water　wash=1>2=NPST=1

　　　　‘Come mother, I’m gonna wash you.’

　　　　CH_MKW_DBC_MAI_1_020320_The two brothers

(248)　ŋa　　　kʰen　　dʑis=alaŋ.
　　　　1SG　　face　　wash=1.PST

　　　　‘I washed my face.’

　　　　CH_MKW_PC_SIL_E

(249)　ŋa=i　　　　hʌw=kaj　　　　　　　kɾut　　dʑis=tak=alaŋ.
　　　　1SG=ERG　　younger.brother=DAT　hand　　wash=CAUS=1.pst

　　　　‘I made younger brother wash his hands (himself).’

　　　　CH_MKW_PC_SIL_E

The third type of intransitive reflexive verb may or may not feature reflexive morphology on the verb and can express an S argument and the possessed noun. The possessor S argument has a semantic role of patient since its referent is affected by the process.

Two types of constructions are attested where: (a) both the possessor S and possessed noun are overtly expressed, and the verb agrees with the S and optionally takes reflexive morphology, as in (250) and (251) with the verb dʑik- ‘hurt’; (b) the possessed noun is the only argument overtly expressed; it controls verb agreement and the verb does not take reflexive morphology; in such a construction, either the possessor is understood through other means, as in (252) and (253), or the construction conveys an impersonal or generic interpretation, as shown in (254) to (256).

Note that with the verb *dʑik-* 'hurt,' the absence of possessed noun entails the meaning of 'be hurt emotionally, upset (not physically),' as in (257); such meaning can also be conveyed through the presence of the noun *mʌn* 'mind, heart,' in a construction calqued from Nepali, which uses मन दुख्नु <man dukhnu> 'have one's feelings hurt, be offended, upset' (Schmidt 1994: 506) as in (258). In both cases, the reflexive marker is not present on the verb.

(250) *ŋa   dom   dʑik=alaŋ=sʌ.*
       1SG   leg   hurt=1.PST=REFL

       'My leg hurts.'
       CH_CTW_SPC_POL_E

(251) *naŋ   dom   dʑik=te=a ?*
       2SG   leg   hurt=2=PST

       'Does your leg hurt?'
       CH_MKW_PC_SIL_E

(252) *ŋa=ko   dom   dʑik=na.*
       1SG=GEN   leg   hurt=NPST

       'My leg hurts.'
       CH_MKW_1_18_BMB_BAN_100717_4_E

(253) *dom   dʑik=a,   ŋa=ko=paj.*
       leg   hurt=PST   1SG=GEN=DIS

       'My leg hurts.'
       CH_MKW_PC_SIL_E

(254) *lok=to   wah=ja=kaj,   dom   dʑik=na,*
       be.far=NMZ:ADV2   walk=COND=DAT   leg   hurt=NPST

'If one/you walk far, one's/your legs hurt.'

CH_MKW_PC_SIL_E

(255) *bʌlʌ   bʌlʌ   ʥan   ʥik=na,*
little   little   body   hurt=NPST

'The body hurts a little,'

CH_MKW_BMB_BAN_102717_1_Asparagus

(256) *wiʔ   dum=ja=kaj,   talaŋ   ʥik=na.*
blood   clot=COND=DAT   head   hurt=NPST

'If the blood clots, one's/your head hurts.'

CH_MKW_STC_SIL_120619_2_E_2_Archive

(257) *ŋa   ʥik=alaŋ.*
1SG   hurt=1.PST

'I'm hurt.'

CH_CTW_SPC_POL_E

(258) *ŋa   mʌn   ʥik=alaŋ.*
1SG   mind   hurt=1.PST

'I'm hurt.'

CH_CTW_SPC_POL_E

The fourth type of intransitive reflexive construction can be characterized by the fact that the reflexive verb is ambitransitive or labile, i.e., displaying reflexive lability. By contrast with the second type of intransitive reflexive construction (also labile), the possessor S argument becomes a P argument in the transitive construction and is marked with the dative case. In the reflexive intransitive construction, the possessor S argument has a semantic role of patient since its referent is affected by the process.

In the intransitive reflexive construction, the possessor S and the possessed noun are overtly expressed and left unmarked in the absolutive case. The possessor S argument can be absent if its referent is recoverable through context or verbal argument indexation,

but the possessed noun is required; the verb agrees with the S and optionally takes reflexive morphology. This is illustrated in (259) to (261) with the verbs *tuk bljʌw-* 'be nauseated,' and *sos-* 'be itchy.'

In the transitive construction, the possessor S argument is marked with the dative case while the possessed noun is left unmarked in the absolutive case, and the verb takes inverse morphology, as in (262) with the verb *sos-* 'be itchy.'

(259) *ŋa   tuk   bljʌw=alaŋ=sʌ.*
       1SG   stomach   be.nauseated=1.PST=REFL

       'I am nauseated.'
       CH_MKW_PC_SIL_E

(260) *ŋa   talaŋ   sos=alaŋ=sʌ.*
       1SG   head   be.itchy=1.PST=REFL

       'My head is itchy.'
       CH_MKW_PC_SIL_E

(261) *o   tuk   bljʌw=a.*
       DIST   stomach   be.nauseated=PST

       'He is nauseated.'
       CH_MKW_PC_SIL_E

(262) *tɕoʔ=kaj   talaŋ   sos=ak=tʌ=i.*
       child=DAT   head   be.itchy=PST=INV=3>3SG

       'The child has her/his head itching her/him.'
       CH_CTW_SPC_POL_E

3.4.5.1.6.   P of transitive verbs

In transitive constructions, the P argument may be left unmarked in the absolute case. In such case, the referent of the P argument is generally animate and indefinite, as with *wa* 'hen' in (263) and *juʔ* 'mouse' in (264), or inanimate and indefinite, as with *siŋ* 'wood' in (265), *nelaw* 'nettle' in (266) and *kapi-kʌlʌm* 'school materials' in (267). As described in § 3.4.5.2, when the referent is inanimate and definite, the P argument of a transitive verb may also be marked with the dative case, as with *loʔ* 'leaf' in (268).

(263) *manta*   *al=na=i,*   *wa*   *lat=ti*   *lat=ti,*
       person   go=NPST=PL   hen   carry=SEQ1   carry=SEQ1

       pudʑʌw=laŋ   *al=na=i.*
       worship=PUR   go=NPST=PL

       'People go, carrying hens, they go worship.'
       CH_MKW_SC_SIL_120619_4_E_2

(264) *ane*   *tɕjew=ja,*   *juʔ=ma*   *dʑe=na=u.*
       so   see_find=COND   mouse=ADD   eat=NPST=3O/DIR

       'So if it (the cat) finds (some), it eats mice too.'
       CH_MKW_STC_SIL_120619_2_E_2

(265) *Dhading*   *Chainpur*   *al=ti,*   *siŋ*   *ɾa=laŋ*
       Dhading   Chainpur   go=SEQ1   tree_wood   cut_saw=PUR

       al=sa.
       go=NMZ1

       'You go to Chainpur, Dhading, and you go cut wood.'
       CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

(266) *ŋa=i*   *nelaw*   *dʑe=a=na=ŋ,*   *sjahŋ-tɕitsʌj*
       1SG=ERG   nettle   eat=EPIS=NPST=1   tomorrow-after.tomorrow

kahw=ti	*dʑe=laŋ*	*al=naŋ*	*mʰʌɾ=ti*	*mʌ=na=ŋ.*
pick=SEQ1	eat=PUR	go=NPST	think=SEQ1	COP=NPST=1

'I'm definitely gonna eat nettle, I'm thinking of going to pick them and eat them in a day or two.'

CH_MKW_SC_SIL_120619_1_E

(267)
gundʌɾi	*nap=ti*	*kapi-kʌlʌm*	*kʌn=na=tɕe=u.*
mat	spread=SEQ1	copy-pen	look=NPST=1/3DU=3O/DIR

'Spreading the mat, they two look at their school stuffs.'

CH_MKW_BBC_SIL_042420_6_My two daughters

(268)
sja?n=i=kʰe	*o*	*ɾajo=ko*	*lo?=kaj*
insect=ERG=DIS	DIST	mustard.greens=GEN	leave=DAT

lat=ti	*a?l=ka=nʌ.*
carry=SEQ1	take.away=2/3.PST=DIR/TR

'The caterpillar carried and took away those mustard greens leaves'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

3.4.5.1.7.   T of ditransitive verbs

The T argument of a ditransitive verb is often left unmarked in the absolutive case, regardless of animacy and definiteness, by contrast with P arguments of transitive verbs. The following examples are constructions formed with the ditransitive verbs *bʌj-* 'give,' *te-* 'ask for, beg,' and *kas-* 'feed.'

T arguments expressing inanimate indefinite referents are illustrated with *bʌl* 'strength' in (269) and *ahm* 'rice, porridge' in (270), inanimate definite referents are illustrated with *ni?ŋ* 'poison' and *dʌbaj* 'medecine' in (271), an animate indefinite referent is illustrated in (272) with *sja?n* 'insect,' and animate definite referent is illustrated in (273) with *kwi* 'dog,' human definite referents are illustrated in (274) with *tɕo?* 'child, son' and in (275) with *tɕo?-djaŋ* 'daughter,' and a human indefinite referent is illustrated in (276) with *tɕo?* 'child, son.'

(269)   *siŋ,*   *naŋ=i=ma,*   *ŋa=kaj*   *bʌl*   *bʌj=tɕi.*
           tree   2SG=ERG=ADD   1SG=DAT   strength   give=2SG>1SG

           'Tree, you too, give me strength.'

           CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(270)   *ama=i*   *dʑʌmmʌj=kaj=le*   *ahm*   *kʰaŋ=ti*   *bʌj=ka=nʌ.*
           mother=ERG   all=DAT=DIS   rice   cook=SEQ1   give=2/3.PST=DIR/TR

           'Mother cooked and gave rice to everyone.'

           CH_MKW_SC_SIL_120619_4_E_2

(271)   *o*   *niʔŋ*   *kas=a*   *kʰe=ja=kaj*
           DIST   poison   feed=NMZ2   COP=COND=DAT

           o=ko=ma   *dʌbʌj*   *bʌj=na=ŋ.*
           DIST=GEN=ADD   medecine   give=NPST=1

           'In the case where such poison is fed (to someone), I give the medecine for it too.'

           CH_CTW_BBC_PID_011520_10_Shaman_healing

(272)   *guɾu=i*   *pande=kaj*   *sjaʔn*   *kas=ka=n.*
           teacher=ERG   shaman=DAT   insect   feed=2/3.PST=DIR/TR

           'The guru fed the shaman with insects.'

           CH_CTW_SPC_POL_E

(273)   *o=i*   *ŋa=kaj*   *i*   *kwi*   *bʌj=ka=ta=ŋ.*
           DIST=ERG   1SG=DAT   PROX   dog   give=2/3.PST=INV=1

           'S/he gave me this dog.'

           CH_CTW_SPC_POL_E

(274)   *naŋ=ko*   *tɕoʔ*   *bʌj=tɕi=ma*   *Sayrang Mama Gryay !*
           2SG=GEN   child   give=2SG>1SG=DIS   Sayrang Mama Gryay

'Please, give me your child Sayrang Mama Gryay!'

CH_CTW_JBC_BHR_102420_1_Cing_Lan

(275) *"oho,* *tɕoʔ-deaŋ* *ŋa=kaj* *te=a=ta=ŋ..."* *mʰʌɾ=a.*
EXPR daughter 1SG=DAT ask.for=PST=INV=1 think=PST

'"Oh my, he asked me for my daughter (to marry)." I thought.'

CH_MKW_CPR_BAN_102817_1_Mit_Co'

(276) *ŋa=kaj* *jat=dʑjo* *tɕoʔ* *bʌj=sa* *pʌɾ=a.*
1SG=DAT one=CL1 child give=NMZ1 have.to_fall=PST

'You have to give me one child.'

CH_CTW_JMC_PYK_101920_Cing_Lan

3.4.5.2. Dative *=kaj*

The dative case marker *=kaj* consistently encliticizes to the Stimulus argument of experiencer intransitive verbs and to the R argument of ditransitive verbs.

When it comes to the marking of the P argument of a transitive verb and the T argument of a ditransitive verb, as briefly mentioned in § 3.4.5.1.6 and § 3.4.5.1.7, a split based on animacy, definiteness and pragmatic saliency is observed.

3.4.5.2.1. Stimulus of experiencer intransitive verbs

The Stimulus arguments of experiencer intransitive verbs that express negative feelings such as *rʌj-* 'be scared, fear' and *kris-* 'be disgusted' are marked with the dative case, while the S argument, which holds a semantic role of experiencer, is left unmarked in the absolutive case. The referent of such oblique arguments may be inanimate, as with *ljam* 'path' in (277), animate, as with *juʔ* 'mouse' in (278) and (279), or human, as with *manta* 'person, people' in (280).

(277) *o* *ljam=kaj=le* *rʌj=na=ŋ,* *didi.*
DIST path=DAT=DIS be.scared=NPST=1 elder.sister

'I fear that path, elder sister.'

CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(278)   *o*   *juʔ=kaj*   *rʌj=na.*
         DIST   mouse=DAT   be.scared=NPST

'S/he fears mice.'

CH_CTW_SPC_POL_E

(279)   *i*   *manta*   *sʌjŋ=o*   *maj=kaj*   *kris=na.*
         PROX person   smell.bad=NMZ:REL meat=DAT   be.disgusted=NPST

'This person is disgusted by smelly meat.'

CH_CTW_SPC_POL_E

(280)   *i*   *manta=kaj*   *kris=na=ŋ.*
         PROX person=DAT   be.disgusted=NPST=1

'I am disgusted by this person.'

CH_CTW_SPC_POL_E

3.4.5.2.2.   P of transitive verbs and T of ditransitive verbs

The dative case morpheme *=kaj* can mark the P argument of a transitive verb.

P arguments marked with the dative case usually express human referents and animate and inanimate definite referents. This is illustrated in (281) to (284) with human definite referents, in (285) with a human indefinite referent, in (286) with an animate definite referent, and in (287) with an inanimate definite referent.

(281)   *ane*   *didi,*   *naŋ=i*   *ŋa=kaj*   *pʰe=ljam.*
         so   elder.sister 2SG=ERG   1SG=DAT   abandon_leave=IMP.NEG

'So, elder sister, don't leave me.'

CH_MKW_SC_SIL_122619_2_Dhobini rani

(282)   *ŋa=i*   *naŋ=kaj*   *pʰe=ŋʌ=lʌ.*
         1SG=ERG   2SG=DAT   leave.behind=1=NEG

'I won't leave you.'

CH_MKW_BMB_BAN_090118_7_Chepang_marriage

(283) *ten=ko*   *pan*   *din=bʰitɾa=haŋ,*   *budʰa*   *waŋ=na,*
today=GEN   five   day=inside=LOC1   husband   come=NPST

naŋ=kaj   *leʔ=laŋ,*   *biha*   *dʑahŋ=laŋ.*
2SG=DAT   take_buy=PUR   marriage   do_make=PUR

'Within the next five days, (your) husband will come, to take you, to marry.'

CH_MKW_SC_SIL_122619_2_Dhobini rani

(284) *o*   *ɾadʑa=kaj*   *sat=ka=n=i.*
DIST   king=DAT   kill=2/3.PST=DIR/TR=PL

'They killed that king.'

CH_CTW_SBC_BBC_GUN_012120_Chepang_Kings

(285) *dʑaʔ=i*   *manta=kaj*   *dʑe=na=u*   *da !*
tiger=ERG   person=DAT   eat=NPST=3O/DIR   PART

'The tiger eats people, hey!'

CH_MKW_SC_SIL_120619_1_E

(286) *o=sʌj*   *ŋi,*   *o*   *juin=kaj*   *o=ko=le*   *awadʑ=haŋ,*
DIST=ABL   1PL   DIST   bat=DAT   DIST=GEN=DIS   sound=LOC1

ʌ,   *tjuʔl=na=ŋ=su*   *ŋi,*   *o=kaj.*
uh,   imitate=NPST=1=1PL.EXCL   1PL   DIST=DAT

'From there, we imitate the sound of those bats, we imitate them.'

lit. 'From there, we imitate those bats in the sound that is theirs, we imitate them.'

CH_MKW_PNC_SIL_081818_2_Bat_hunting

(287) *naŋ=ko*   *ɾam-siʔ*   *klʌhj=dʑe=pa,*
2SG=GEN   sickle   break.with.hands=EPIS=OPT

lek-si-siŋ=kaj   *lʰo=dʑe=ljam.*

flame-tree-tree=DAT   chop.off=EPIS=IMP.NEG

'May your sickle break, don't you dare chop off the flame tree.'

CH_CTW_JMC_PYK_101920_Cing Lan

As described in § 3.4.5.1.7, T arguments of ditransitive verbs are left unmarked in the absolutive case. This is shown again with *niʔŋ* 'poison' in (288) and *tɕoʔ* 'child, son' in (289).

(288)   *manta=i*   *niʔŋ*   *kas=na=n=i*   *kja !*
         person=ERG   poison   feed=NPST=DIR/TR=PL   PART

       'People feed (people) poison, be aware of that!'

       CH_CTW_BBC_PID_011520_10_Shaman_healing

(289)   *"niŋdʑi=paj*   *tɕiŋ-lan=kaj*   *bʌj=o*   *rʌjsʌ,*
         2DU=DIS   Cing-spirit=DAT   give=PERF   COP.MIR

       tɕoʔ=paj"=tʌ   *rja=ti=taŋ,*

       child=DIS=REP   say=SEQ1=ATT

       'And he said: "It looks like you two have given to the Cing spirit, the child,"'

       CH_CTW_BBC_GUN_102620_1_Cing_Lan

The distribution of case marking for the P argument of a transitive verb and for the T argument of a ditransitive verb shows that: (1) P arguments are marked with the dative case when expressing human referents in addition to definite animate and inanimate referents; (2) regardless of humanness, animacy and definiteness, T arguments are left unmarked in the absolutive case. This is summarized in Table 115.

In § 3.4.5.2.3, I will briefly show that pragmatic saliency may trigger changes in the case marking patterns observed for S, P and T arguments.

Table 115. Case marking for P and T arguments of transitive and ditransitive verbs

referent	human	animate	inanimate	animate	inanimate
	definite	definite	definite	indefinite	indefinite
	indefinite				
argument					
P	DAT	DAT	DAT	ABS	ABS
T	ABS	ABS	ABS	ABS	ABS

3.4.5.2.3. Pragmatic saliency

As described in § 3.4.5.2.2, the P argument of a transitive verb is marked with the dative case when the referent expressed is human, animate and definite, and inanimate and definite, while indefinite animate and inanimate are left unmarked in the absolutive case, and the T argument of a ditransitive verb is left umarked in the absolutive case, regardless of the referent being human, animate, inanimate, definite or indefinite.

Examples of indefinite inanimate marked with the dative case may be found, as shown in (290). Inanimate referents such as *siŋ* 'tree,' *baŋ* 'stone,' or *saʔ* 'soil, earth' are considered as deities and thus anthropomorphized. Two other examples are given in (291) and (292). They are pragmatically more salient in the context of spiritual worship.

(290)   *ŋi*   *siŋ=kaj*   *puʣ=ʌ=na=ŋ=su.*
          1PL   tree_wood=DAT   worship=LN=NPST=1=1PL.EXCL

         'We worship trees.'

         CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(291)   *ʌni*   *ŋi*   *bʰumi putrʌ*   *hʌ*   *dahj=ti*
          then   1PL   Bhumi son   COP   say=SEQ1

         siŋ=kaj   *ʣogaw=sa,*   *baŋ=kaj*   *ʣogaw=sa,*
         tree=DAT   protect=NMZ1 stone=DAT   protect=NMZ1

         baŋ=kaj,   *baŋ=kaj*   *ʣʰor=sa*
         stone=DAT   stone=DAT   greet=NMZ1

ʌni   siŋ=kaj=ma   ʥʰoɾ=sa.
then   tree=DAT=ADD   greet=NMZ1

'And because we are the sons of the Goddess Bhumi, we protect the trees, we protect the stones, and for the stones, we greet the stones, and we greet the trees too.'

CH_MKW_SCBKC_SIL_081918_3_Chepang_culture

(292) ŋi   siŋ=kaj   puʥ=ʌ=na=ŋ=su,   ŋi   baŋ=kaj
       1PL   tree=dat   worship=LN=NPST=1=1PL.EXCL   1PL   stone=DAT

       puʥ=ʌ=na=ŋ=su,   saʔ=kaj
       worship=LN=NPST=1=1PL.EXCL   earth=DAT

       puʥ=ʌ=na=ŋ=su.
       worship=LN=NPST=1=1PL.EXCL

'We worship the trees, we worship the stones, we worship the earth.'

CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

The absence of dative case on the P arguments that refer to an indefinite human and indefinite animate in (293) is also pragmatically salient. In this case, killing people or the tiger spirit is presented as something banal, since it was the usual routine of the shaman, referent of the A argument. By contrast, the P argument that refers to trees, an indefinite inanimate, as for the audience, killing a tree with shamanic powers is rather unusual and hence marked with the dative case.

(293) manta   sat=sa=ma   kʰaj=to,   siŋ=kaj
       person   kill=NMZ1=ADD   be.able=REM.PST tree=DAT

       sat=sa=ma   kʰaj=to,
       kill=NMZ1=ADD   be.able=REM.PST

       tɕʰak-ʥaʔ   sat=sa=ma   kʰaj=to,
       dead.tiger.spirit kill=NMZ1=ADD   be.able=REM.PST

374

'(The shaman) could also kill people, he could also kill trees, he could also kill the Tiger spirit of the dead,'
CH_CTW_RLC_JIM_101920_Language and Culture

In (294), the same kind of pragmatic saliency occurs, where the P argument that refers to an indefinite human is left unmarked in the absolutive case. The fact that the tiger spirit may kill people is not unexpected if the shaman let it wander around.

(294) *sat=ma=ja=kaj,*   *manta*   *ʥe=na=u.*
     kill=NEG=COND=DAT person   eat=NPST=3O/DIR

'If (the shaman) doesn't kill (the Tiger spirit of the dead), it kills people.'
CH_CTW_BBC_PID_011520_9_Tiger_Spirit

Some examples of T arguments of ditransitive verbs unusually marked with the dative case are also found, as in (295). What is pragmatically salient here is the fact that giving the girl to this particular man to be husband goes against her wish.

(295) ʌ,   *baba* *ɾʌ*   *ama=i,*   ʌ,   *keti=kaj,*
     uh   father and   mother=ERG   uh   girl=DAT

     o   *ʥʰak=ma=l=o*   *buda=kaj*   *bʌj=ka=n=i.*
     DIST   like=NEG=COP=NMZ:REL   husband=DAT give=2/3.PST=DIR/TR=PL

'Uh, the father and mother, uh, they gave their daughter to the husband she didn't love.'
CH_MKW_SC_SIL_122619_2_Dhobini_rani

Another example of T argument of a ditransitive verb marked with the dative is given in (296). The Cing spirit has already asked for the two eldest daughters, that he killed and ate, and is not done yet, determined to ask for one more daughter.

(296) *al=na=ŋ*   *djah,*   *pʰeɾi* *sajli-ʨoʔ=kaj*   *te=na=ŋ.*
     go=NPST=1   now   again   third.eldest.F-child=DAT ask.for=NPST=1

'I'll go now, I'm going to ask for the third eldest daughter.'

Finally, examples of S arguments of intransitive verbs may show the presence of ergative morphology, as in (297) with the intransitive verb *dahj-* 'say.'

(297) *naw=le*   *Chepang,*   *aw*   *manta=le*   *Chepang,*
      name=DIS   Chepang   CAT   person=DIS   Chepang

       naw=le   *Chepang*   *dʌj,*   *ŋa=i*   *dahj=o*
      name=DIS   Chepang   DIS   1SG=ERG   say=NMZ:REL

       i=tɕjuk   *le.*
      PROX=QTY   COP

'The name is Chepang, this person here is Chepang, the name is Chepang hey, what I said is this much.'

Further research is needed to better understand the different pragmatic motivations behind the attested differential object marking in Chepang, and whether they fall within referential or discourse prominence (Haspelmath 2021).

3.4.5.2.4. R of ditransitive verbs

The R argument of ditransitive verbs are marked with the dative case. This is illustrated in (298) to (303) with the verbs *bʌj-* 'give,' *te-* 'ask for, beg,' *kas-* 'feed,' and *tan-* 'show.' The referents of the R arguments of such ditransitive verbs are either human or animate.

(298) *i*   *keta=kaj*   *ʌwtʰi*   *bʌj=na=ŋ,*   *ten=paj !*
      PROX   boy=DAT   ring   give=NPST=1   today=DIS

'I'm going to give a ring to that boy, today!'
CH_MKW_SC_SIL_122619_2_Dhobini_rani

(299) *lʌw*   *djah=paj*   *pahj=i,*   *djah=paj*   *bʰʌjsi=kaj=ma*
    well   now=DIS   leave=1PL.IMP.INTR   now=DIS   buffalo=DAT=ADD

    ti   *bʌj=sa*   *pʌɾ=na.*
    water   give=NMZ1   have.to_fall=NPST

'Well now, let's leave, now I have to give water to the buffaloes too.'
CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(300) *o*   *ɾuiŋ=ko*   *gʰas=tɕʌhe*   *kim=haŋ*
    DIST   bamboo=GEN   weed=DIS   house=LOC1

    mʌ=o   *kʰʌsi=kaj*   *bʌj=ka=n=i.*
    COP=NMZ:REL   castrated.goat=DAT   give=2/3.PST=DIR/TR=PL

'They gave that bamboo weed to the castrated goats that are at home.'
CH_CTW_BBC_POL_111720_6_Jogi

(301) *o=kaj*   *pʌjsa*   *te=ti*   *kim*
    DIST=DAT   money   ask.for=SEQ1   house

    dzahŋ=sa   *pʌɾ=a.*
    do_make=NMZ1   have.to_fall=PST

'Asking him for money, we had to build a house.'
CH_MKW_SC_SIL_010220_1_Life

(302) *e,*   *jo-ɾo*   *aʔl=ti*   *ploʔm-juin=kaj*
    EXPR   Indian.Butter.tree-flower   take.away=SEQ1   grandson-bat=DAT

    pan   *din*   *sʌmmʌn=taŋ*   *kas=ka=nʌ.*
    five   day   during=ATT   feed=2/3.PST=DIR/TR

'Taking away Indian Butter tree's flowers, (the grandmother) fed her grandson-bat for five days.'
CH_CTW_KMC_TAP_102520_2_The bat and the crab

(303) *naŋ=kaj*   *ŋa=i,*   *ŋa=ko*   *kim*   *tan=ʨe=na=ŋ !*
      2SG=DAT   1SG=ERG   1SG=GEN   house   show=1>2=NPST=1

'I'm gonna show my house to you!'

CH_MKW_19-20_SCBKC_SIL_021920_51_E

3.4.5.2.5.   Other types of dative arguments

The dative case occurs as well on oblique arguments used as point of reference. It can occur in a metalinguistic context, as in (304) and (305), when giving one's personal opinion on something, as in (306), or to express a purpose, as in (307).

(304) *kʌtʰa=kaj*   *"juin-ɾaj,"*   *Chepang*   *bʰasa=sʌj*
      story=DAT   story   Chepang   language=ABL

      "juin-ɾaj"   *to=na=n=i.*

      story   tell_say=NPST=DIR/TR=PL

'For "story," it's "yuin-rāy," in the Chepang language, they say "yuin-rāy."'

CH_CTW_YMC_TAP_102420_3_Agreement

(305) *ʌ̃*   *mak-saj=kaj*   *"bʰaɾla=ko*   *tata."*
      yeah   bauhinia_vahlii-fruit_seed=DAT   bharlā=GEN   tātā

'Yeah, for "mak-sāy," it's "bauhinia vahlii fruit".'

CH_MKW_RC_JMC_SIL_120119_Conversation

(306) *i=ʨuk=ko*   *keta,*   *ŋa=kaj*   *sjaw=na.*
      PROX=QTY=GEN   boy   1SG=DAT   happen=NPST

'A boy of this kind is fine with me.'

CH_MKW_BMB_BAN_103119_9_Mother and daughter

(307) *i*   *pɾʌbʰu=ko*   *kam=kaj,*   *ŋa=i*   *nʌɾamɾo*
      PROX   god=GEN   work=DAT   1SG=ERG   bad

      to=ŋʌ=lʌ.

tell_say=1.NEG=NEG

'For the work of this god, I don't say bad things.'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian

3.4.5.3.   Ergative =*i*

The ergative case marks the A argument of a transitive or ditransitive verb. Ergative case is typically found in an ergative-absolutive alignment system where the A argument is marked differently from the S and P arguments.

The ergative case is marked with the enclitic morpheme =*i*. The same form is used for the instrumental case (§ 3.4.5.4).

The ergative case is mostly used with A arguments whose referents are animate, as illustrated with the transitive verb *gʌm-* 'put, keep' and *dʑahŋ-* 'do, make,' in (308) and (309) respectively, and the ditransitive verb *bʌj-* 'give' in (310).

The presence of the ergative case is nevertheless not restricted to animate referents, as shown in (311) with the noun *baŋ* 'stone.'

(308) *ŋa=i*   *i*   *tuʔm*   *gʌm=o,*   *nis*   *bʌrsʌ*   *sjaw=a.*
       1SG=ERG   PROX   bee   put_keep=PERF   two   years   become=PST

'It's been two years since I set up these bees.'

CH_MKW_MRNDC_SIL_081818_1_Bee keeping

(309) *dʰere*   *dʑʌso*   *[Chepang=lʌm]ᴺᴾ =i esu-kriɕ=haŋ*
       very   often   Chepang=PL=ERG   Jesus   Christ=LOC1

       biswas   *dʑahŋ=o.*
       faith   do=PERF

'Very often, the Chepangs believe in Jesus-Christ.'

CH_MKW_BBC_SIL_032920_2_My grandfather

(310) *ten=paj*   *ŋa=i=paj*   *lʌj=ko*   *sun=ko*   *ʌwtʰi*
       today=DIS   1SG=ERG=DIS   SLF.INTS=GEN gold=GEN   ring

       aʔl=ti   *o=kaj*   *bʌj=na=ŋ.*

379

take.away=SEQ1   DIST=DAT give=NPST=1

'Today, I will bring and give her my own golden ring.'

CH_MKW_SC_SIL_122619_2_Dhobini rani

(311) *ʌni* *baŋ=i*   *tɕjap=ti*   *sat=na=u.*
    then stone=ERG fall.on_crush=SEQ1 kill=NPST=3O/DIR

'Then, the stone falls on (the porcupine) and kills it.'

CH_MKW_PMRC_LAM_081618_4_Porcupine

3.4.5.4.   Instrumental =*i*

  The instrumental case marks the argument that expresses the tool with which an action is carried out. The instrumental is marked with the enclitic morpheme =*i*, which is phonologically identical to the ergative case (§ 3.4.5.3). The instrumental case is illustrated in (312) and (313). The ergative case is illustrated in (314) for comparison.

  It is possible that the ergative and instrumental historically share the same origin since ergative marker can also be found with inanimate nouns, as shown in (315).

(312) *kʰoɾ=sʌj* *kɾut=i* *metɕʰja* *tɕum=ka=n,*
    shelter=ABL hand=INST goat catch=2/3.PST=DIR/TR

    bʌhila   *metɕʰja.*
    sterile   goat

    'S/he caught a goat with her/his hands from the shelter, a sterile goat.'

CH_MKW_SC_SIL_010220_3_Life

(313) *kahw=bʌteko* *ɾama=i* *plak=sa.*
    pick=SEQ2   sickle=ERG split=NMZ1

    'Having picked it (Bauhinia vahlii), we split it with the sickle.'

CH_MKW_RC_JMC_SIL_120119_Conversation

(314) *(...) manta=lʌm=i* *hum=ti=tɕʌhe,*   *o-mi=kaj=tɕʌhe*

person=PL=ERG   gather.together=SEQ1=DIS   DIST-PL.H=DAT=DIS

tɕum=ti   *aʔl=ti*   *pʌkrʌ=ti*   *nelaw=i*
catch=SEQ1   take.away=SEQ1   arrest=SEQ1   nettle=INST

sɾap=n=i=to.
flog=DIR/TR=PL=REM.PST

'People would gather together, catch the person, take them away, arrest them, and flog them with nettle.'
CH_CTW_SRP_GUN_102620_4_Local justice

(315) *siŋ=i*   *tɕjap=ti*   *sat=aka=n,*
        tree=ERG   press_crush=SEQ1   kill=2/3.PST=DIR/TR

        ŋa=ko   *majli-tɕoʔ-djaŋ.*
        1SG=GEN   second.eldest.F-daughter

        'The tree crushed her and killed her, my second eldest daughter.'
        CH_MKW_KKBP_LPK_101917_1_Life

3.4.5.5.   Comitative =*kusi*

The comitative case =*kusi* marks an oblique argument that expresses the company of whom or what is carrying out the process expressed by the verb.

This is illustrated in (316) to (319) with people as referents, and in (320) with a noun referring to a species of yam accompanying a dish.

(316) *ŋi=ko*   *Chairang*   *pastʌr=ko*   *kim*   *kaʔm=taŋ=ko*
        1PL=GEN   Chairang   pastor=GEN   house downwards=ALL=GEN

        bʌdʑe=kusi   *ta=na=ŋ,*
        old.woman=COM   tell.story=NPST=1

        'I tell stories with the old woman from around downwards the house of our pastor from Chairang.'
        CH_MKW_SC_SIL_010220_3_Life_Archive

(317) ba-ama,   ʨoʔ=lʌm=kusi,   sʌbʌj=kusi=le
       parents   child=PL=COM   all=COM=DIS

       ŋa   Chepang   bʰasa=le   noʔ=na=ŋʌ.
       1SG   Chepang   language=DIS speak=NPST=1

       'I speak Chepang with everybody, with my parents, my children.'
       CH_MKW_TRC_DAM_112819_Conversation_with_Bipana

(318) o   neta   ŋi=kusi=le   kam   ʣahŋ=na=u.
       DIST   politician   1PL=COM=DIS   work   do=NPST=3O/DIR

       'That politician works with us.'
       CH_MKW_BBC_SIL_032820_2_Pambung

(319) Chepang   noʔ=ma=l=o=kusi   kʰʌsʌnta=le   noʔ=na=ŋ.
       Chepang   speak=NEG=COP=NMZ:REL=COM   Nepali=DIS   speak=NPST=1

       'With those who don't speak Chepang, I speak Nepali.'
       CH_MKW_TRC_DAM_112819_Conversation_with_Bipana

(320) ʌ,   ʌni   ahm   kʰaŋ=sa=kaj   ʌ,   gu-kjan=kusi !
       uh   then   porridge   cook=NMZ1=DAT   uh   yam.sp-dish=COM

       'Uh, so, to cook porridge, uh, it's with a yam (*Dioscoreaceae*) dish!'
       CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

The comitative morpheme =*kusi* attaches to nouns referring to persons, or objects, but it is also found attached to abstract nouns, as in (321), or nominalized verbs, as in (322), to form manner adverbials, describing the way the process is realized. This type of construction developed through calque from Nepali which employs the morpheme संग <saṃga> as a comitative marker, as shown in (323) and (324).

(321) mʌʣʣa=kusi sati=haŋ   tar=ʌ=a   pʌʨʰi
       joy=COM   oil=LOC1   fry=LN=NMZ2   after

       o=ko   maj=kʰe   kʰʌsi=ko

DIST=GEN	meat=DIS	castrated.goat=GEN

maj=le   *hʌ*   *tahŋ=to*   *sjaw=na.*

meat=DIS   COP   be.huge_be.like=NMZ:ADV2   become=NPST

'After frying it nicely in oil, its meat becomes like that of a castrated goat.'

CH_MKW_PMRC_LAM_081618_4_Porcupine

(322) *maha* *pe=to=kusi*   *dʑe=ti=le*   *mu=na=ŋ.*
       very   be.nice=NMZ:ADV2=COM   eat=SEQ1=DIS   COP=NPST=1

'I'm eating very well.'

CH_MKW_BMB_BAN_090118_8_Marriage_Archive

Nepali

(323) मज्जासँग
      majjā=sãnga

joy.satisfaction=COM

'with satisfaction, enjoyably'

Nepali

(324) राम्रोसँग
      rāmro=sãnga

nice=COM

'well, nicely'

The type of verbs found with comitative oblique arguments can be intransitive, as illustrated in (325) to (329), transitive, as in (330), or else ditransitive, as in (331).

(325) *o*   *bela=haŋ,*   *klaʔ=ti=le*   *ane,* *o=le,*
       DIST   moment=LOC1   fall=SEQ1=DIS   PART DIST=DIS

       kantɕʰi-ama=kusi   *mu=o.*
       step.mother=COM   COP_stay_live=PERF

383

'At that time, I went down, it's that, I lived with my step-mother.'

CH_MKW_SPMC_LC_SIL_100921_3_Conversation

(326) sjaʔn=ma   o   bop=lʌm=kusi   eʔn=a.
insect=ADD   DIST   snail=PL=COM   sleep=PST

'The insect too slept with the snails.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(327) ba=kusi   kaj=ŋʌ=to.
father=COM   get.worked.up=1=REM.PST

'I used to get worked up with my father.'

CH_MKW_BMB_BAN_090118_8_Marriage

(328) ŋa=kusi   waŋ=dʑe=ljam,   (...)
1SG=COM   come=EPIS=IMP.NEG

'Don't you dare coming with me,'

CH_MKW_SC_SIL_122619_2_Dhobini_rani

(329) naŋ   dʌjti   ŋa=kusi   noʔ=te=lʌ   ane ?
2SG   why   1SG=COM   speak=2=NEG   PART

'Why don't you speak with me then?'

CH_MKW_SC_SIL_122619_2_Dhobini_rani

(330) (...) NGO=kusi   kam   dʑahŋ=ŋʌ=to.
NGO=COM   work   do_make=1=REM.PST

'I used to work with an NGO.'

CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(331) meʔn=kusi   bʌj=o   ja,   meʔn   waj=ti
hair=COM   give=PERF   or   hair   throw.away=SEQ1

kʰaŋ=o   ja ?
cook=PERF   or

'Have you given (bats to her) with the hair, or have you thrown the hair away and cooked (them)?'

CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

While a comitative argument does not trigger verbal indexation, some examples are found when 1st person singular is involved, where the S or A argument is marked as dual on the verb in presence of a comitative argument. This is illustrated in (332) and in (333) with an S argument, and in (334) with an A argument.

(332) | *talim* | *al=alaŋ,* | *ŋa* | *Butwal* | *al=alaŋ=tɕʌ* |
| --- | --- | --- | --- | --- |
| training | go=1.PST | 1SG | Butwal | go=1.PST=1/3DU |

badʑja	*Bhabikan=kusi,*	*ek*	*nʌmbʌr.*
elder.M	Bhabikan=COM	one	number

'I went to the training, we two went to Butwal with elder Bhabikan, in Butwal ward number 1.'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian_Archive

(333) | *si=te=tɕja,* | *si=sa* | *kʰe=lʌ,* | *pahj=ti* | *waŋ=ʌ,* |
| --- | --- | --- | --- | --- |
| die=2=IRR | die=NMZ1 | COP=NEG | leave=SEQ1 | come=2SG.IMP.INTR |

ŋa=kusi	*mu=tɕʌ*		*asam,*	*bolaw=alaŋ*	*ŋa=i.*
1SG=COM	COP=1DU.IMP.INTR		PART	call=1.PST	1SG=ERG

'Don't die, we don't want you to die, leave and come, let's the two of us stay with me, if it's like this, I asked her to come.'

CH_MKW_SC_SIL_010220_3_Life

(334) | *ŋa=kusi=paj* | *pɾem* | *dʑahŋ=tɕu,* |
| --- | --- | --- |
| 1SG=COM=DIS | love | do_make=1DU.IMP.TR |

kantɕʰi	*Syangja*	*al=tɕʌ*	*da !*

385

| younger.one.F | Syangja | go=1DU.IMP.INTR | PART |

'Let the two of make love with me, let the two of us go to Syangja, little one.'
CH_MKW_SC_SIL_122619_3.2_Song

As seen, the morpheme =*kusi* marks the comitative case on arguments that are considered oblique, since their presence is not necessary to the predicate.

However, two constructions formed with the transitive verb *rʌj*- 'fear, be scared' and the ditransitive verb *te*- 'beg, ask for, request,' feature the presence of a core argument marked with the comitative morpheme =*kusi*. These constructions are likely calqued from Nepali since the native constructions are still in use, and the same core argument is marked with the dative case morpheme =*kaj*. The two constructions are illustrated in (335) to (340), where one can compare the construction calqued from Nepali with the native Chepang construction and the Nepali construction.

(335) | *dʑun* | *ki* | *tɕʰak-dʑaʔ* | *sjak-dʑaʔ=kusi* | *ŋi* |
| whichever | or | dead.tiger.spirit | tiger.spirit=COM | 1PL |

| *rʌj=ti* | *mʌ=sa* | *kuɾa* | *kʰe=lʌ* | *ki* | *tʌ ?* |
| fear=SEQ1 | COP=NMZ1 | thing | COP=NEG | or | PART |

'It is not about us fearing either the Tiger spirit of the dead or the Tiger spirit of the alive, or is it?
CH_CTW_RLC_JIM_101920_Language and Culture

(336) | *ŋa=paj* | *rʌj=a=na=ŋ* | *ru=kaj=paj,* |
| 1SG=DIS | fear=EPIS=NPST=1 | snake=DAT=DIS |

| *ne=na=u* | *ni !* |
| bite=NPST=3O/DIR | PART |

'I definitely fear snakes, be sure that they bite!'
CH_MKW_SC_SIL_120619_4_E_2
Nepali

(337) | मलाई | भुतभन्दा | चोरसँग | ज्यादा | डर | लाग्छ। |
| *ma=lāi* | *bhut=bhandā* | *cor=sãga* | *jyādā* | *ḍar* | *lāg=cha.* |

1SG=DAT	ghost=COMP	thief=COM	much	fear	apply=COP.3SG

'I fear thieves more than ghosts.'

(Schmidt 1994: 265)

(338) ŋa=ko   aj   te=laŋ   al=a,
      1SG=GEN   mother.in.law   ask.for=PUR   go=PST

      ŋa=ko   ama-ba=kusi   dʌj !
      1SG=GEN   mother-father=COM PART

'My mother-in-law went to request (me for marriying her son) to my parents!'

CH_MKW_BMB_BAN_090118_8_Marriage

(339) ʌɾu=kaj   gun   te=laŋ   waŋ=o
      other=DAT   vertue_property   ask.for=PUR   come=NMZ:REL

      manta=kaj   gun   bʌj=u=to,
      person=DAT   vertue_property   give=3O/DIR=REM.PST

      ʌm=tak=u=to.
      heal=CAUS=3O/DIR=REM.PST

'He would give the (shamanic) knowledge to those who come to request shamanic knowledge to others, he would heal them.'

CH_MKW_BBC_SIL_032920_2_My grandfather

Nepali

(340) दाइले   कसैसँग   अनुमति   मागेन।
      dāi=le   kasai=sãga   anumati   māg=e=na.
      elder.brother=ERG   anyone=COM   permission ask.for=3SG.PST=NEG

'Elder brother didn't ask anyone's permission.'

(Schmidt 1994: 13)

Finally, comitative arguments may also express possession, as shown in (341) and (342). This construction, also found in Nepali, may also have developed through calque, as shown in (343).

(341) ŋa=kusi   pʌjsa   na=ma=l=o   bela=haŋ,
      1SG=COM   money   COP=NEG=COP=NMZ:REL   moment.time=LOC1

      'When I don't have money,'

      CH_MKW_PMRC_SIL_081818_1_Life

(342) ŋa=kusi   pɾʌman   mʌ=na.
      1SG=COM   proof   COP=NPST

      'I have proof.'

      CH_CTW_BBC_PID_011520_7_Chepang_Raute

      Nepali

(343) मासँग   पैसा   छैन।
      ma=sãga   paisā   chaina.

      1SG=COM   money   COP.3SG.NEG

      'I don't have money.'

      (Schmidt 1994: 393)

3.4.5.6.   Locative =*haŋ* and =*kʰa*

Two locative markers are attested: a morpheme =*haŋ* and a morpheme =*kʰa*. They mark the argument that indicates the stative locational position or goal direction (in space or time) of a referent, event or action. The morphemes =*haŋ* and =*kʰa* are respectively illustrated in (344) and (345) with the proximal pronoun *i*.

(344) naŋ=ko   didi=lʌm=ko   bɾeh   i=haŋ   mu=na.
      2SG=GEN   elder.sister=PL=GEN   finger PROX=LOC1   COP=NPST

      'Your elder sisters' fingers are here.'

      CH_CTW_JMC_PYK_101920_Cing_Lan

(345) ŋaʔ   i=kʰa   mu=na   ɾʌjsʌ !
      fish   prox=LOC2   COP=NPST   COP.MIR

'Fish are here!'
CH_CTW_SP_POL_111420_Ream_Tokrak

In terms of space, both morphemes function as general locatives, in the sense that the position they express can be inessive (inside something) and superessive (on the surface of something), or the direction allative (movement to, towards, into, onto something).

For instance, the located referent in (346) and (347), could either be situated inside (inessive) or on top of (superessive) the house. It could also be located next to (adessive) the house since the meaning of *kim* 'house' does not just imply the building itself but also the land on which it is built.

(346) *lʌw   ama,   naŋ   kim=haŋ   mu=ʌ.*
      well   mother   2SG   house=LOC1   COP=2SG.IMP.INTR

  'Well, mother, you stay at home.'
  CH_MKW_DBC_MAI_2_020320_Newa_Dung

(347) *ŋa   pahj=na=ŋ,   kim=kʰa   mu=ʌ.*
      1SG   leave_go.home=NPST=1   house=LOC2   COP=2SG.IMP.INTR

  'I'm going to leave, stay at home.'
  CH_MKW_BMB_BAN_090118_8_Marriage

The following examples of arguments marked with *=haŋ* and *=kʰa* illustrate inessive positions and dynamic motions in (348) to (353), and superessive positions and dynamic motions in (354) to (359). Examples of allative directions are given in (360) and (361).

(348) *dʑabi=haŋ   laŋ   lat=ti   al=na=ŋ=su.*
      net.bag=LOC1 net   carry=SEQ1   go=NPST=1=1PL.EXCL

  'We go carrying the net in the net bag.'

(349) ʌ,   jʌ-sati=kʰa   kʰaŋ=ti,
       uh   Indian.butter.tree.oil=LOC2   cook=SEQ1

'Uh, having cooked (the bats) in Indian butter tree oil,'

CH_MKW_PNC_SIL_081818_2_Bat_hunting

(350) o   gʰaŋ=haŋ=taŋ   dʰampʌ=i=taŋ   r̥oŋ=o.
       DIST   hole=LOC1=ATT   stick=INST=ATT   impale=PERF

'He has impaled (the crab) in/into that hole with the stick.'

CH_CTW_RC_KCR_101920_2_Myth_Origin

(351) o   sjaʔn   gʰaŋ=kʰa   klaʔ=na.
       DIST   insect   hole=LOC2   fall=NPST

'That insect falls in/into the hole.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(352) kim=haŋ   pok=laŋ   pahj=a.
       house=LOC1   enter=PUR   leave_go.home=PST

'He left to enter in/into the house.'

CH_CTW_BBC_POL_111720_1_Cing_Lan

(353) bʌn=kʰa   pok=a.
       jungle=LOC2   enter=PST

'He entered in/into the jungle.'

CH_MKW_BMB_BAN_090118_10_Cak_Ja'

(354) baŋ=kʰa=paj=taŋ (...)   naj   mʌtrʌj=taŋ
       stone=LOC2=DIS=ATT   clothes   only=ATT

       ʨjew=na=n=i.

       see_find=NPST=DIR/TR=PL

'On the stone, (…) they only find the clothes.'

(355) *jat=ʤjaŋ*   *ɾu*   *baŋ=haŋ*   *mu=to.*
one=CL1   snake   stone=LOC1   COP=REM.PST

'One snake was on the stone.'

CH_MKW_1_21-22_CPR_BAN_100817_1_E

(356) *o=haŋ*   *al=ti,*   *eʔn=alaŋ,*   *tebul=haŋ.*
DIST=LOC1   go=SEQ1   sleep=1.PST   table=LOC1

'Having gone there, I lied down, on the table.'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(357) *bʌn=kʰa=taŋ*   *al=na,*   *pʰeɾi*   *baŋ=kʰa=taŋ*   *tɕjuŋ=taŋ=na.*
jungle=LOC2=ATT   go=NPST   again   stone=LOC2=ATT   sit=ATT=NPST

'She goes to the jungle, and again she sits on the stone.'

CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

(358) *saʔ=haŋ*   *dah=alaŋ.*
earth_ground=LOC1   reach=1.PST

'I reached the ground.'

CH_MKW_1_BMB_BAN_100517_3_E

(359) *ŋa*   *ɾʌh=alaŋ,*   *saʔ=kʰa.*
1SG   fall=1.PST   earth_ground=LOC2

'I fell, on the ground.'

CH_MKW_1_BMB_BAN_100517_3_E

(360) *ŋa=ko*   *kim=haŋ*   *al=i.*
1SG=GEN   house=LOC1   go=1PL.IMP.INTR

'Let's go to my house.'

CH_MKW_CPR_BAN_102817_1_Mit_Co'

(361) ʌni   sahu=ko   kim=kʰa   al=a.
       then   employer=GEN   house=LOC2   go=PST

'Then, he went to the employer's house.'

CH_MKW_SMYMP_CHI_080118_2_Food_ressources

With proper nouns expressing toponyms, i.e., names of countries, cities, towns and villages, the locative markers may be absent, as shown in (362) and (363) with dynamic motion, and in (364) with stative location.

This is common cross-linguistically in languages displaying differential place marking (Malchukov & De Swart 2008; Creissels 2009; Stolz, Lestrade & Stolz 2014; Schlossberg 2017; Haspelmath 2019) since toponyms semantically refer to locations.

(362) kantɕʰi-tɕoʔ   India   pahj=o   mʌ=na.
       younger.F-child   India   leave=PERF   COP=NPST

'My younger daughter has left to India.'

CH_MKW_SC_SIL_010220_1_Life

(363) uh   Bharatpur   al=alaŋ,   kʰaj=ta=ŋ=lʌ.
       REM   Bharatpur   go=1.PST   be.able=INV=1=NEG

'I went to, this one there, Bharatpur, it wasn't possible to me.'

CH_MKW_PSC_MAI_012620_2_Becoming_Christian

(364) bʰaj   si=o   bela=haŋ,   ŋa
       younger.brother   die=NMZ:REL   moment=LOC1   1SG

India   kʰe=ŋʌ=to.
India   COP=1=REM.PST

'At the time of the death of my younger brother, I was in India.'

CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

It is unclear how functionally different these two morphemes are and in particular when it comes to expressing location in space. They are used interchangeably to express both a stative locational position and the goal direction of a dynamic motion.

The type of verb employed does not seem to trigger the presence of a morpheme over the other. This is illustrated in Table 116, with co-occurrences of verbs expressing dynamic motion and stative location attested with the noun *kim* 'house' marked either with =*haŋ* or =*kʰa*. Further, the presence of =*haŋ* or =*kʰa* does not a priori seem to correspond to any variational difference either, since both can occur in a single speaker's speech.

The speakers' intuition is that there is no meaning difference between the two markers. Caughley (1982: 60) characterizes =*haŋ* as a "general locative" and =*kʰa* as an "inessive/approximate locative *in*, *at*," but does not provide any description.

Table 116. Types of verbs attested with *kim* 'house'

	verbs	*kim=haŋ* (203 tokens)	*kim=kʰa* (85 tokens)
dynamic	*al-* 'go'	✓	✓
	aʔl- 'take away'	✓	✓
	dah- 'reach'	✓	✓
	dak- 'deliver'	✓	✓
	djahn- 'return'	✓	
	lat- 'carry'	✓	
	pahj- 'leave, go home'	✓	✓
	pok- 'enter'	✓	
	waŋ- 'come'	✓	✓
	wah- 'walk, go'	✓	
	waʔn- 'bring'	✓	✓
stative	*utpʌdʌn dzahŋ-* 'produce (crops)'		✓
	dze- 'eat'	✓	✓
	eʔn- 'sleep'	✓	✓
	gʌm- 'put, keep'	✓	✓
	kʰaŋ- 'cook'	✓	
	mu- COP	✓	✓

na=lʌ COP=NEG	✓	✓
bʰasa noʔ- 'speak language'		✓
biha sjaw- 'marry'	✓	
rʰam- 'graze, make graze'		✓

In terms of frequency[44], the number of occurrences of the morpheme *=haŋ* is higher than *=kʰa* in our corpus: 6,790 tokens for *=haŋ* against 2,260 for *=kʰa*. When looking at co-occurrence between the roots and their locational markers, I find two main differences.

First, the morpheme *=haŋ* occurs more significantly with abstract nouns than the morpheme *=kʰa*. This is illustrated with a few examples in Table 117. Note that most abstract nouns are borrowed from Nepali. Despite the frequency results found in our corpus, when these examples are discussed with the speakers, they suggest that using *=kʰa* with these abstract nouns is also possible. Counterexamples also exist, such as *mʌn* 'mind, opinion, feelings,' which can also be found attached with the morpheme *=kʰa*, as in (365).

(365) *naŋ=ko*   *mʌn=kʰa*   *mu=o*   *tɕidz,* *kura=lʌm (...)*
       2SG=GEN   mind=LOC2   COP=NMZ:REL thing   thing=SML

'The things that are in your mind (...)'
CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

Table 117. Higher occurrence frequency of *=haŋ* with abstract nouns

abstract nouns	***=haŋ* (6790 tokens)**	***=kʰa* (2260 tokens)**

baɾe 'topic, subject' (<N.)	86	-
bʰasa 'language' (<N.)	42	-
sʌnsaɾ 'earth, world' (<N.)	44	-
ɾup 'aspect, shape' (<N.)	38	-
kuɾa 'language, thing' (<N.)	34	-
nepali 'Nepali language' (<N.)	30	-
biswas 'religion, belief' (<N.)	21	-
kʰʌndʌ 'part, section' (<N.)	24	-
ʌwʌstʰa 'situation, setting' (<N.)	27	-
dukʰʌ 'pain' (<N.)	24	-
pɾʌbʰu 'lord, god' (<N.)	21	-
bitɕaɾ 'thought, opinion' (<N.)	18	-
tʰok 'thing' (<N.)	18	-
pɾitʰibʰi 'nature' (<N.)	15	-
dzʌndzati 'indigenous people' (<N.)	12	-
nam 'name' (<N.)	10	-
kʌltɕaɾ 'culture' (<N. <E.)	9	-
kʌɾmʌ 'destiny' (<N.)	9	-
umeɾ 'age' (<N.)	9	-

Second, when expressing location in time, *=kʰa* is found combining with deictic temporal adverbs that specifically refer to the past, as illustrated in Table 118, while both *=haŋ* and *=kʰa* occur with nouns expressing non-deictic time, as shown in Table 119.

Table 118. Higher occurrence frequency of *=kʰa* with past deictic temporal adverbs

deictic temporal adverbs	**=haŋ (6790 tokens)**	**=kʰa (2260 tokens)**

mis 'earlier'	-	24
pʌhjla 'before, in the past'	-	18
kjamnʌm 'three days ago'	-	4
tɕitnʌm 'two days ago'	-	3
joh 'yesterday'	-	2
ʌnʌm 'recently'	-	2
teh 'last year'	-	2
uhile 'long ago'	-	2

Table 119. Occurrence frequency of *=haŋ* and *=kʰa* with non-deictic temporal adverbs

non-deictic temporal nouns	*=haŋ* (6790 tokens)	*=kʰa* (2260 tokens)
bela 'time, moment' (<N.)	552	204
bʌrsʌ 'year' (<N.)	85	9
sʌmʌje 'time' (<N.)	72	3
din 'day' (<N.)	70	4
pala 'generation, time of' (<N.)	36	2
sal 'year' (<N.)	30	-
bʌkʰʌt 'moment' (<N.)	28	-
suɾu 'beginning' (<N.)	22	-
ʌntim 'end, last' (<N.)	18	-
mʌhjna 'month' (<N.)	15	3

With deictic temporal adverbs that refer to the future, the morpheme *=kʰa* is not present; either there is no specific morphology, as in (366) with *sjahŋ* 'tomorrow,' or these adverbs are formed with the ablative morpheme *=sʌj*, as in (367).

(366)　*sjahŋ*　　　*ŋi=ko*　　　*lipi*　　　*hʌɾaw=dʰʌj*
　　　　tomorrow　　1PL=GEN　　script　　lose=PROG

　　　　al=o　　　*mu=na,*　　　*hʌjnʌ ?*

　　　　go=PERF　　COP=NPST　　no

'Tomorrow, our script will have gone being lost, don't you think?'
CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

(367) ʨit=sʌj   doh   ʥe=sa ?
       two.days.later=ABL   what   eat=NMZ1

       'What to eat after tomorrow?'
       CH_MKW_SC_BGR_101619_1_Life

The use of =haŋ and =kʰa can situate an event at one point in time, as in (368) with a deictic temporal adverb referring to the past, or in (369) and (370) with non-deictic temporal nouns.

(368) teh=kʰa   ŋa=ko   meʨʰja   ʥaʔ=i
       last.year=LOC2   1SG=GEN   goat   tiger=ERG

       ʥe=o   kʰe=to.
       eat=PERF   COP=REM.PST

       'Last year, a tiger had eaten my goat.'
       CH_MKW_PMRC_LAM_081618_3_Kalitar

(369) 72   sal=ko   tʰjakkʌj   13   gʌte   bʌjsakʰ   mʌhjna=haŋ,
       2072 year=GEN   precisely   13   day   Apr.-May   month=LOC1

       ŋako   baba=i=ma   sʌnsaɾ=sʌj   bida
       1SG=GEN   father=ERG=ADD   world=ABL   leave

       leʔ=ka=nʌ.
       take_buy=2/3.PST=DIR/TR

       'In the month of April-May of the year 2072, precisely the 13th, my father passed.'
       CH_MKW_BBC_SIL_032920_2_My grandfather

(370) wa=na,   magʰ   mʌhjna=kʰa.
       rain=NPST   Jan.-Feb.   month=LOC2

       'It rains, in the months of January-February.'

Deictic temporal adverbs that refer to the past may also occur without the morpheme =kʰa, as one can see comparing (371) and (372). This is also the case for nouns expressing non-deictic time, as shown in (373), where nor =haŋ, nor =kʰa occur. The genitive morpheme =ko is also attested attached to deictic temporal adverbs that refer to the past, as illustrated in (374). This construction originates from the possessive modification function of =ko that applied not only to a noun but also to a clause.

(371) ŋa=i   mis   Nepali   to=bʌti,
       1SG=ERG   earlier   Nepali   tell_say=SEQ2

       dzʰʌn ŋa=kaj   Chepang   kuɾa   hot=ti   mu=o.
       more   1SG=DAT   Chepang   language   ask=SEQ1   COP=PERF

       'After I talked in Nepali earlier, she is asking me more in Chepang.'
       CH_MKW_BLC_SIL_113019_1_Shaman_Song

(372) ŋa=i   mis=kʰa   lʰak=o=ma   kʰe=to.
       1SG=ERG   earlier=LOC2   tell_recount=PERF=ADD   COP=REM.PST

       'It was what I had told too earlier.'
       CH_CTW_BBC_POL_111720_3_Cing_Lan

(373) o   "kljaɾ-kljaɾ"=tʌ   lekʰʌ   lekʰʌ=pʌti   bag=ʌ=na,
       DIST   klyar-klyar=REP   ridge   ridge=side   scream=LN=NPST

       tɕʌjt   bʌjsakʰ   mʌhjna   mu=a.
       Mar.-Apr.   Apr.-May   month   COP=PST

       'It (the bird) screams klyar-klyar on the side of the ridge, it stayed in the months of March, April, May.'
       CH_CTW_BBC_POL_111720_6_Jogi

(374) ʌ,   o=le   waŋ=na,   o   uhile=ko

uh	DIST=DIS	come=NPST	DIST	long.ago=GEN

ŋa=ko	*baba=le*	*ke=o.*
1SG=GEN	father=DIS	sing=PERF

'Uh, that is the one that comes (to me), that is what my father had sung of long ago.'

CH_MKW_SC_SIL_122619_5_Song

The use of *=haŋ* and *=kʰa* with nouns expressing non-deictic time can refer to a period of time which corresponds to the time length of the event, as illustrated in (375) to (377), and to a period of time after which the event takes place, as shown in (378) to (381).

(375)
ma,	*din=haŋ,*	*atʰ*	*gʰʌntʌ*	*kam*	*dʑahŋ=na=ŋ*	*ma=ba.*
yes,	day=LOC1	eight	hours	work	do=NPST=1	PART=PART

'Yes, in/during a day, I work eight hours for sure.'

CH_MKW_SC_SIL_120619_2_E_4

(376)
ŋa	*egʰara*	*bʌrsʌ=haŋ,*	*pura*	*gajt*	*dʑahŋ=ti*	*wah=o*	*kja.*
1SG	eleven	year=LOC1	full	guide	do=SEQ1	walk=PERF	PART

'The thing is, I was fully living as a guide for eleven years.'

CH_CTW_BBC_PID_011520_7_Chepang_Raute

(377)
maj,	*u*	*bela=kʰa=paj,*	*dʌsʌj=kʰa*	*mʌtrʌj*	*dʑe=to.*
meat	REM	moment=LOC2=DIS	Dasain=LOC2	only	eat=REM.PST

'Meat, at that time, would only be eaten during the Dasain festival.'

CH_MKW_RC_JMC_SIL_120119_Conversation

(378)
pan	*din=haŋ*	*ti*	*wa=to.*
five	day=LOC1	water	rain=REM.PST

'It rained after five days.'

CH_MKW_SCBKC_SIL_081918_3_Chepang_culture

(379)   *tɕʰʌ*   *mʌhjna=haŋ*   *ŋa*   *djahn=ti*   *waŋ=na=ŋ.*
        six   month=LOC1   1SG   return=SEQ1   come=NPST=1

        'I'll come back in/after six months.'

        CH_MKW_1_33-37_BMB_BAN_101017_E

(380)   *tɕoʔ*   *dzʌrmʌ=ti*   *atʰ*   *din=kʰa=paj*   *dah=laŋ=le*
        child   be.born=SEQ1 eight   day=LOC2=DIS   reach=PUR=DIS

        waŋ=a.
        come=PST

        'The child was born and arrived home eight days later.'

        CH_MKW_SC_SIL_010220_1_Life

(381)   *biha*   *dzahŋ=ti=paj,*   *tʰupro* *bʌrsʌ=kʰa*   *si=a.*
        marriage   do=SEQ1=DIS   a.lot   year=LOC2   die=PST

        'Having got married, he died a lot of years later.'

        CH_MKW_SPMC_LC_SIL_100921_3_Conversation

        Interrogative pronouns referring to location and time are respectively formed with the morphemes =*haŋ* and =*kʰa*. They are illustrated in (382) and (383).

(382)   *ga=haŋ*   *al=a,*   *o*   *radza ?*
        INT=LOC1   go=PST   DIST   king

        'Where did he go, that king?'

        CH_CTW_SBC_BBC_GUN_012120_Chepang_Kings

(383)   *ŋa=ko*   *tokrak*   *kwa,*   *gʌ=lʌ=kʰa*
        1SG=GEN   frog.*sp*   bond.friend   INT=when=LOC2

        dah=na   *ane ?*
        reach=NPST   PART

        'My frog bond friend, when will s/he arrive then?'

        CH_CTW_SP_POL_111420_Ream_Tokrak

A nominal stem marked with =*haŋ* or =*kʰa* can also occur combining with the ablative case marker =*sʌj* to indicate source location, although this construction does not show any difference in meaning compared to the sole presence of the ablative when it comes to encliticizing nouns (§ 3.4.5.7).

Constructions combining the distal demonstrative pronoun *o* with =*haŋ* or =*kʰa* followed by the ablative =*sʌj*, can also express temporal and causal meanings as discourse connectors (§ 3.4.5.7).

Finally, =*haŋ* and =*kʰa* can both attach to nominalized clauses to form adverbial simultaneous subordinate clauses: =*haŋ* combines with =*to* nominalized clauses, as in (384) and (385); =*kʰa* combines with =*o* nominalized clauses, as in (386) and (387). In addition, =*kʰa* combines with the negation marker =*ma* to form antecedent subordinate clauses, as in (387).

(384) *ŋi*   *dʑega=tɕʌhe*   *dʑe=to=haŋ,*   *ŋi=ko=tɕʌhe*
       1PL   food=DIS   eat=NMZ:ADV2=LOC1 1PL=GEN=DIS

       pet   *dʰad=ʌ=na.*
       stomach   have.stomach.ache=LN=NPST

       'When eating food specifically, our stomach hurts.'
       CH_CTW_BBC_PID_011520_3_Witches

(385) *jat=dʑo*   *aʔl=to=haŋ*   *pãtɕ* *bʌjhni*
       one=CL1   take.away=NMZ:ADV2=LOC1 five   younger.sister

       tɕaɾ *bʌhjni*   *tʰop=na=i.*
       four   younger.sister   be.enough=NPST=PL

       'When one (a man) takes away (a woman), five younger sisters, four younger sisters are enough (for him).'
       CH_CTW_JMC_PYK_101920_Cing_Lan

(386) ŋa   mʌj=o=kʰa=le   Chepang   gãw=haŋ=le
     1SG   be.small=NMZ:REL=LOC2=DIS   Chepang   village=LOC1=DIS

     mu=o   kʰe=to.
     COP=PERF   COP=REM.PST

     'I was living in a Chepang village when I was little.'
     CH_MKW_KMC_SK_082918_6_Childhood

(387) i=ma   batɕa   bʌj=o=kʰa   mʌtrʌj   waŋ=to,
     PROX=ADD   vision give=NMZ:REL=LOC2   only   come=REM.PST

     sjan=ti   dʑʰja=o=kʰe   sjaw=lʌ.
     teach_learn=SEQ1   shaman.practice=NMZ:REL=DIS become=NEG

     'This (shaman's power) also would come (to the shaman) only when given
     visions, it doesn't work for those who, having learned, practice ceremonies.'
     CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(388) naŋ=i   ʌ,   biha   dʑahŋ=ma=kʰa,   naŋ
     2SG=ERG   uh   marriage   do=NEG=LOC2   2SG

     doh   kam   dʑahŋ=te=u=to ?
     what   work   do=2=3O/DIR=REM.PST

     'Before you got married, what work did you do?'
     CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

The origin of the morpheme =kʰa is uncertain. The morpheme =haŋ is cognate
with the Bhujel locative marker =haŋ and Magar locative marker =aŋ, while the
morpheme =kʰa does not have any known cognate in Bhujel or Magar. The morpheme
=kʰa is hence likely a Chepang innovation.

The uses and morphosyntactic distribution of =haŋ and =kʰa show two interesting
functional differences: the morpheme =haŋ attaches to nominal stems or nominalized
constructions, while the morpheme =kʰa has further the ability to mark adverbs in
addition to having an adverbializing function.

First, as seen, the morpheme =kʰa can occur with deictic temporal adverbs that refer to the past, while =haŋ does not, marking temporal non-deictic nouns; second, when forming adverbial simultaneous subordinate clauses, the morpheme =kʰa attaches directly to nominalized clauses formed with the nominalizer =o, while =haŋ attaches to nominalized clauses formed with the nominalizer =to.

In fact, the nominalizer =to can historically be analyzed as the combination of the manner adverbializer =tʌ and the nominalizer =o. The manner adverbializer =tʌ can form manner adverbs based on demonstrative pronouns, as in (389), and manner adverbs based on onomatopoeic roots, as in (390) in addition to negative manner adverbial subordinate clauses, as in (391).

That is, for =haŋ to create an adverbial simultaneous subordinate clause, the adverbializing function of the construction must have first been carried out by the adverbializer morpheme =tʌ, and not =haŋ, since this latter only attaches to nominal forms, hence needing the presence of the nominalizer =o in such construction.

The combination of =tʌ and =o, which led to the form =to, is mainly used when stative intransitive verbs hold the position of adjectival predicates, i.e., structurally nominalized adverbials, as in (393), or adverbials, as in (394), while the mere presence of the nominalizer =o serves the function of creating adjectivals off of stative intransitive verbs when modifying a nominal, as in (395), structurally a nominalized verb.

(389) o=tʌ   lʰok=ma=ja   sjaw=to.
       DIST=NMZ:ADV1   send=NEG=COND   become=REM.PST

       'It would be fine if (she) didn't reject (her) like that.'
       CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

(390) kɾʌːːːːːm=tʌ   dʑʌjk=ti=taŋ   dʑe=ak=ta=i
       kraaaaam=NMZ:ADV1   bite=SEQ1=ATT   eat=PST=INV=3>3sg

       o   dʑetʰi-tɕoʔ=kaj=ma.
       DIST   elder.F-child=DAT=ADD

       '(The Cing spirit) bit and ate her greedily, that elder daughter too.'
       CH_CTW_BBC_GUN_102620_1_Cing_Lan

(391) *naŋ=kaj*   *ŋaj=ma=tʌ,*   *djahn=ti*
2SG=DAT   meet=NEG=NMZ:ADV1   return=SEQ1

pahj=ti   *waŋ=o.*
leave=SEQ1   come=PERF

'I've come back without meeting you.'

CH_MKW_SC_SIL_120619_2_E_2

(392) *si=sa*   *kʰaj=ma=tʌ*   *dukʰʌ*   *ʨejew=ti*
die=NMZ1   be.able=NEG=NMZ:ADV1   sorrow   see_find=SEQ1

mu=o.
COP=PERF

'Not being able to die, I have been finding pain.'

CH_MKW_SMYMP_CHI_080118_4_Seasons

(393) *degi*   *pe=to=le*   *mu=na=ŋ.*
now   be.good=NMZ:ADV2=DIS   COP=NPST=1

'Now, I'm well/good.'

lit. 'Now, I am who is being good.'

CH_MKW_BMB_BAN_090118_8_Marriage

(394) *tʌɾʌ*   *ʌɾu*   *kaʔm=taŋ=ko*   *juba=lʌm,*
but   other   downwards=ALL=GEN   youngster=PL

Chepang   *bʰasa*   *pe=to*   *ʨiʔ=n=i=li.*
Chepang   language   be.good=NMZ:ADV2   know=DIR/TR=PL=NEG

'But the other youngsters of the downwards area, they don't know well the Chepang language.'

CH_MKW_PNC_SIL_081818_4_Chepang_Culture_Language

(395) *gʌ=ʨjuk*   *pe=o*   *boka*   *kʰe=to !*
INT=QTY   be.good=NMZ:REL   castrated.goat   COP=REM.PST

'What a nice castrated goat it was!'

3.4.5.7.   Ablative =*sʌj*

The ablative case =*sʌj* marks an argument that expresses the source location from where the process expressed by a verb takes place. This is illustrated in (396).

(396) *i*   *Maisirang*   *waŋ=o*   *dʑʌti*   *puɾkʰja=lʌm=kʰe,*
       PROX   Maisirang   come=NMZ:REL   as.much.as   ancestor=PL=DIS

       ʌ,   *Pambung=sʌj*   *waŋ=o.*
       uh   Pambung=ABL   come=PERF

       'As much ancestors as those who have come to this Maisirang, uh, they have come from Pambung.'
       CH_MKW_DKC_CBC_MAI_020220_Local_History

Beyond location, the morpheme =*sʌj* is also used metaphorically to express the origin of things, as shown in (397), the means by or through which a process is realized, as in (398) and (399), to enumerate periods of time chronologically, as in (400), or else to subtract, as in (401).

(397) *o*   *sjaʔn=sʌj*   *tibiliŋ*   *sjaw=ti,*
       DIST   insect=ABL   butterfly   become=SEQ1

       'Having become a butterfly from that insect,'
       CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(398) *hʌhʌ,* *Bhwi* *diʔŋ=sʌj*   *o=kaj*   *tɕejew=o !*
       hʌhʌ   Bhwi   god=ABL   DIST=DAT   see_find=PERF

       'Haha, one has found it through the god Bhwi!'
       CH_MKW_BLC_SIL_113019_1_Shaman_Song

(399) *"juin-ɾaj"*   *to=na,*   *kʌtʰa=kaj*   *"juin-ɾaj,"*
       story   tell_say=NPST   story=DAT   story

       Chepang   *bʰasa=sʌj*   *juinɾaj*   *to=na=n=i.*
       Chepang   language=ABL   story   tell_say=NPST=DIR/TR=2/3PL

'One says "Yuin-rāy," for 'story,' "Yuin-rāy," they say "Yuin-rāy" by means of the Chepang language.'

CH_CTW_YMC_TAP_102420_3_Agreement

(400) *ʌni*   *Prithivi Narayan Shaha=sʌj (…),*   *Tribhuvan Shaha,*
       then   Prithivi Narayan Shah=ABL   Tribhuvan Shah=ABL

       Tribhuvan Shaha=sʌj,   *Mahendra Shaha,*   *Mahendra Shaha=sʌj,*
       Tribhuvan Shah=ABL   Mahendra Shah   Mahendra Shah=ABL

       Birendra Shaha (…)
       Birendra Shah

'Then from Prithivi Narayan Shah (…), Tribhuvan Shah, from Tribhuvan Shah, Mahendra Shah, from Mahendra Shah, Birendra Shah (...)'

CH_CTW_BBC_POL_102520_1_Polkim

(401) *degi*   *gu=tɕjuk*   *bʌrsʌ*   *sjaw=a,*   *ektɕalis*   *sal=sʌj ?*
       now   INT=QTY   year   become=PST forty-one   year=ABL

'Now, how many years it's been, from the year 2041?'

CH_MKW_BLC_SIL_113019_2_Becoming_Shaman

The ablative case marker =sʌj is also attested combining with the locative case markers =haŋ and =kʰa.

Three examples of a noun marked with both =haŋ and =sʌj are found in our corpus. They come from the varieties spoken in Manahari (MAN-7, RAK-8); they are presented in (402), (403) and (404). They do not seem to present any difference in meaning with regard to the absence of the morpheme =haŋ, since they too merely indicate the source location of the process expressed.

(402) *pʰeɾi, bʰiɾ=haŋ   tʌn=alaŋ,   bʰiɾ=haŋ=sʌj=ma   ɾʌh=alaŋ.*
again cliff=LOC1   run=1.PST   cliff=LOC1=ABL=DIS fall=1.PST

'What is more, I ran on the cliff, I also fell from the cliff.'

CH_MKW_ARRKC_CHI_102919_1_Life

(403) *Dhan Lal   pastʌɾ kja,   ʌ,   u=ko,   Bartani=ko.*
Dhan Lal   pastor PART   uh   CAT=GEN   Bartani=GEN.

Bartani=haŋ=sʌj   i=haŋ   waŋ=a.
Bartani=LOC1=ABL   PROX=LOC1   come=PST

'The pastor Dhan Lal hey, uh, of there, of Bartaniko. He came here from Bartani.'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(404) *i   moʔm-tɕoʔ   bʌdzaɾ=haŋ=sʌj   kim=kʰa*
PROX   girl   town=LOC1=ABL   house=LOC2

waŋ=ti=le   mu=na.
come=SEQ1=DIS   COP=NPST

'This girl is coming home from town.'

CH_MKW_1_73-79_CPR_BAN_102417_2_Verb

In fact, the combination of *=haŋ* (§ 3.4.5.6) and *=sʌj* (§ 3.4.5.7) is mainly found attached to the demonstrative pronouns: the proximal *i*, distal *o*, and remote *u*. These constructions formerly correspond to the formation of the locational adverbs *ihaŋ* 'here,' *ohaŋ* 'there' and *uhaŋ* 'over there' then encliticized with the ablative morpheme *=sʌj* (§ 3.4.5.7), as in (405), (406) and (407).

Besides, the form *ohaŋsʌj* is also used as a discourse connector between events expressed in a temporal or causal manner, as in (408). This function has not developed with *ihaŋsʌj* and *uhaŋsʌj*.

(405)   *Kathmandu*   *lʌt=sa=kaj=le*   *i=haŋ=sʌj*
        Kathmandu   climb_go.up=NMZ1=DAT=DIS   PROX=LOC1=ABL

        pan   *din=taŋ*   *lagʌ=to.*
        five   day=ATT   apply=REM.PST

        'To go up to Kathmandu, it used to take five days from here.'
        CH_CTW_BBC_POL_102520_1_Polkim

(406)   *o=haŋ=sʌj*   *dah=ti,*   *ŋa=i*   *Nepal=haŋ*   *dah=ti,*
        DIST=LOC1=ABL   reach=SEQ1   1SG=ERG   Nepal=LOC1   reach=SEQ1

        ŋa=i   *sahu=kʌ=haŋ*   *dʑob*   *dʑahŋ=alaŋ*
        1SG=ERG   employer=NMZ:LOC=LOC1   job   do_make=1.PST

        Manahari=haŋ.
        Manahari=LOC1

        'Arriving from there (India), I, arriving in Nepal, I worked at an employer's in
        Manahari.'
        CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

(407)   *u=haŋ=sʌj,*   *ka?m=sʌj,*   *Chitwan=sʌj,*
        REM=LOC1=ABL   downwards=ABL   Chitwan=ABL

        Dhan Lal   *waŋ=a.*
        Dhan Lal   come=PST

        'Dhan Lal came from over there, from down there, from Chitwan.'
        CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(408)   *kantɕʰi-ama=paj,*   *ŋa=ko*   *palo,*   *"lʌw*   *lʌw*   *naŋ=kʰe*
        youner.F-mother=DIS 1sg=GEN   turn   well   well   2SG=DIS

        dʑe=ti   *mu=sa,"*   *ŋa=kʰe*   *bʌstu*   *rʰam=sa*
        eat=SEQ1   COP=NMZ1   1SG=DIS   cattle   make.graze=NMZ1

        mʌtrʌj,   *mʰʌr=ti,*   *o=haŋ=sʌj,*
        only   think=SEQ1   DIST=LOC1=ABL

ʌnum=ko	*sjahŋ*	*dekʰin=paj*
recently=GEN	tomorrow	since=DIS

bju-mʌkʌj	*tjal=ti,*	*dzʌŋgʌl=taŋ*	*aʔl=ti*
corn.grain	roast=SEQ1	jungle=ALL	take.away=SEQ1

dze=o,	*ŋa=i.*
eat=PERF	1SG=ERG

'My stepmother, (it was) my turn, "Come on, with you, it's about eating and resting," and thinking that for me it was only making the cattle graze, from there/because of that, the following days, roasting corn grains and taking (them) away to the jungle, I would eat (them there).'

CH_MKW_SPMC_LC_SIL_100921_3_Conversation

Like with *=haŋ*, the morpheme *=sʌj* is found attached to the forms *ikʰa* 'here,' *okʰa* 'there' and *ukʰa* 'over there', as in (409), (410) and (411).

(409)
i=kʰa=sʌj	*Bharat*	*tɕjew=a=na.*
PROX=LOC2=ABL	India	see=EPIS=NPST

'From here, India is seen for sure / one sees India for sure.'

CH_MKW_SC_SIL_120619_2_E_4

(410)
jom	*waŋ=a,*	*o=kʰa=sʌj=le*	*tʰa=ti*	*waŋ=a.*
bear	come=PST	DIST=LOC2=ABL=DIS	appear=SEQ1	come=PST

'A bear came, it appeared from there and arrived.'

CH_MKW_1_28-31_BMB_BAN_100917_1_E

(411)
pʰeɾi	*u=kʰa=sʌj=ma,*	*Yukdhung*	*waŋ=le=na*	*kja !*
again	REM=LOC2=ABL=ADD	Yukdhung	come=DIS=NPST	PART

'Again, from over there as well, Yukdhung comes, hey!'

CH_MKW_BRC_CYO_120119_Yukdhung

The form *okʰasʌj* 'from there' is also attested as a discourse connector expressing a temporal or causal meaning, as illustrated in (412) and (413).

(412) *mobajl*   *pʌsʌl=haŋ*   *kam*   *dʑahŋ=alaŋ,*
     mobile   shop=LOC1   work   do_make=1.PST

     mʌhjna=ko   *tin*   *hʌdʑaɾ.*
     month=GEN   three   thousand

     o=kʰa=sʌj,   *o*   *dʑob*   *pʰe=ti,*
     DIST=LOC2=ABL   DIST   job   leave.behind=SEQ1

     'I worked in a mobile store, for 3,000 NPRs ($30) per month. From there/Because of that, leaving that job,'

     CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

(413) *ane*   *o*   *ŋa=ko*   *kim=kʰe*   *Sidam.*   *ek*   *nʌmbʌɾ.*
     so   DIST   1SG=GEN   house=DIS   Sidam   one   number

     ane   *o=kʰa=sʌj,*   *noʔ=sa,*   *i*   *Silinge*   *dada*   *ɾʌ,*
     so   DIST=LOC2=ABL   speak=NMZ1   PROX   Silinge   hill   and

     Silinge   *dada*   *Makal Damar,*
     Silinge   hill   Makal Damar

     ane   *Sidam gaw*   *noʔ=sa*   *ja=dʑo*   *pʰʌɾʌk.*
     so   Sidam village   speak=NMZ1   one=CL1   different

     ja=dʑo=le.
     one=CL1=DIS

     'So, that house of mine is in Sidam. In Sidam number 1. So, because of that, speaking, in this Silinge hill and, in this Silinge hill and in Makal Damar and in the Sidam village, speaking is similar. It's the same.'

     CH_MKW_RC_JMC_SIL_120119_Conversation

Only two examples are found in our corpus with a noun marked by both =*kʰa* and =*sʌj*. They are presented in (414) and (415). Example (414) does not seem to show any difference in meaning in absence of =*kʰa*. The construction in (415), however, expresses a temporal or causal meaning rather than a locational one. While this is not a construction accepted by the speakers it was discussed with, it nevertheless recalls the temporal or causal meaning expressed when =*kʰa* and =*sʌj* encliticize to the distal demonstrative pronoun *o* to function as a discourse connector, or to the Nepali borrowed noun *karʌn* which means 'cause, reason,' illustrated in (416) and (417). In the case of example (415), the construction may simply be the result of a metaphoric analogical implication of a causal meaning carried out by the expression 'through the phone.'

(414) *ŋi*   *Chepang=lʌm=kʰe,*   *pʌhjla=kʰa*   *dʑʌŋgʌl=kʰa=sʌj*   *waŋ=o,*
       1PL   Chepang=PL=DIS   before=LOC2   jungle=LOC2=ABL   come=PERF

       dʑʌŋgʌl=haŋ=paj   *mʌ=jak=tʌh=i=li*   *ni,*   *ma*   *ŋi* ?
       jungle=LOC1=DIS   COP=REM.PST=1PL.INCL=PL=NEG   PART   yes   1PL

       'As for us, Chepangs, in the past, we had come from the jungle, we didn't live in the jungle before, right, did we?
       CH_CTW_SRP_GUN_102620_3_Gundi_settlement

(415) *pʰon=kʰa=sʌj,*   *dʰerʌj*   *tɕendʑ*   *sjaw=laŋ*   *waŋ=a*   *kja.*
       phone=LOC2=ABL   a.lot   change   become=PUR   come=PST   PART

       'Since the phone / because of the phone, a lot of change came to become.'
       CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

(416) *i=haŋ*   *ʌgadi*   *kam*   *bʌdaw=sa*
       PROX=LOC1   forward   work   increase=NMZ1

       kʰaj=ma=l=o   *karʌn=kʰa=sʌj,*   *o=haŋ=sʌj*   *ŋa*
       be.able=NEG=COP=NMZ:REL   cause=LOC2=ABL   DIST=LOC1=ABL   1SG

       mobajls elektroniks   *sik=ʌ=laŋ*   *Kathmandu*   *al=alaŋ.*
       mobiles electronics   learn=LN=PUR   Kathmandu   go=1.PST

'Because of the fact that here the modern work can't increase, from there, I went to Kathmandu to learn mobiles electronics.'
CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

(417) *ŋi=ko*   *Chepang*   *bʰasa*   *lop*   *sjaw=ti*
       1PL=GEN   Chepang   language   lost   become=SEQ1

       al=o   *mu=na.*
       go=PERF   COP=NPST

       o   *kaɾʌn=kʰa=sʌj*   *pɾaj-ʣʌsto*   *ŋa*
       DIST   cause=LOC2=ABL   usually-like   1SG

       i=taŋ   *lʌj=ko*   *bʰasa=le*   *noʔ=na=ŋ.*
       PROX=ALL   SLF.INTS=GEN   language=DIS   speak=NPST=1

'Our Chepang language has started to become lost. Because of that, I usually speak my own language around here.'
CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

Finally, the morpheme *=sʌj* also functions as a comparative marker. It attaches to the argument whose referent is compared with the referent of another argument. This is illustrated in (418).

(418) *ŋa=sʌj=paj*   *hʌw*   *manta, (...)*
       1SG=CMPR=DIS   younger.brother   person

'He is younger than me, (…)'
CH_MKW_SC_SIL_010220_1_Life

3.4.5.8.   Allative *=taŋ*

The allative case marker *=taŋ* encliticizes to an argument that expresses the goal of a motion, as illustrated in (419) with the verb *al-* 'go' and in (420) with the verb *pahj-* 'leave.' The presence of the allative morpheme *=taŋ* co-occurs with verbs expressing

dynamic motions, such as those reported in Table 120, along with examples of nominal stems with which the morpheme =taŋ attaches.

(419) *nam*   *tɕiʔ=ma=tʌ*   *naŋ=kusi*   *Manahari=taŋ*
       name   know=NEG=NMZ:ADV1   2SG=COM   Manahari=ALL

       brʌk=le   *al=ŋʌ=lʌ.*
       together=DIS   go=1=NEG

       'Without knowing your name, I won't go together with you to Manahari.'
       CH_CTW_SC_SP_POL_102620_2_Raksi_lo'hang_Song

(420) *Tongra*   *kʰola=taŋ*   *bol*   *kʰel=laŋ*   *pahj=na=i.*
       Tongra   river=ALL   ball   play=PUR   leave=NPST=PL

       'They leave to the Tongra river to play football.'
       CH_MKW_MRC_LPK_080918_3_Chepang_Language

Table 120. Dynamic verbs and nominal stems attested with allative case

allative case		
dynamic verbs	**nominal stems**	
al- 'go'	*bʌn* 'jungle'	*kaʔm* 'downwards'
dah- 'reach'	*bʌʥaɾ* 'town'	*ljam* 'path'
djahn- 'return'	*kim* 'house'	*ti* 'water (sense of well)'
lʌt- 'carry'	*gaw* 'village'	*tjaw* 'upwards'
pahj- 'leave'	*kʰola* 'river'	

The morpheme =taŋ can also co-occur with the verb *pok-* 'enter, fall in.' In that case, it carries an inessive (inside something) meaning. This is illustrated in (421) with the noun *kim* 'house,' and in (422) with the noun *ti* 'water.'

(421) *o*   *buda=ko*   *kim=taŋ=le*   *pok=laŋ*   *pahj=alaŋ.*
       DIST   husband=GEN   house=ALL=DIS   enter=PUR   leave=1.PST

'I left to enter the house of that husband.'
CH_MKW_SC_SIL_010220_1_Life

(422) *plukʨjuŋ,*   *ti=taŋ*   *pok=a !*
ONO   water=ALL   enter=PST

'Plukcyung, she fell into the water!'
CH_MKW_DBC_MAI_2_020320_Newa_Dung

When the morpheme =*taŋ* occurs with stative verbs, like the locational copula *mu-*, it expresses a circumessive meaning (around something), i.e., a vague location. This is illustrated in (423).

(423) *ʥʌmmʌj*   *bʌʥaɾ=taŋ=le*   *mʌ=i=to.*
all   town=ALL=DIS   COP=PL=REM.PST

'They all were around town.'
CH_MKW_SC_SIL_120619_2_E_3

Finally, the allative morpheme =*taŋ* should not be confounded with the discourse attentional marker =*taŋ* with which it can also co-occur, as in (424).

(424) *o=sʌj=paj*   *o*   "*lʌw* *mʌh=lʌm*   *bʌh=lʌm*
DIST=ABL=DIS DIST   well   sister.in.law=ASS   brother.in.law=ASS

i=kʰa   *mu=na*   *gwej* *djaw=nu"=taŋ*   *ʨʰjan=ti*
PROX=LOC2   COP=NPST   yam dig.out=2PL.IMP.TR=ATT   show=SEQ1

pʰe=bʌti=taŋ   *ʌɾko*   *dada=taŋ=taŋ*   *ʥjal=na.*
leave=SEQ2=ATT   other   hill=ALL=ATT   be.on.the.way=NPST

'Then, he (said) "Well, sister-in-law, brother-in-law, and family, there is yam here, dig it out," and having showed them, after he left them, he is on his way to another hill.'
CH_MKW_JMC_SIL_120619_1_Barbalyak

3.4.6. Relator nouns and postpositions

Twenty-five nominal postpositional elements are attested in Chepang: six native relator nouns, two relator nouns borrowed from Nepali, and seventeen postpositions borrowed from Nepali. Relator nouns and postpositions are presented in Table 121.

I describe native relator nouns in § 3.4.6.1 and relator nouns and postpositions borrowed from Nepali in § 3.4.6.3.

In addition to these, the morpheme *ljam* 'path,' which developed the function of a nominal directional and ablative case marker from its morphosyntactic position of semantic head in a determinative or descriptive compound is described in § 3.4.6.2.

In the following descriptions, in addition to data from text, I use examples collected through the questionnaire developed in *Topological relations picture series* (Bowerman & Pederson 1992).

Table 121. Relator nouns and postpositions

	ref. frame	locational (space / time)	
relator nouns	absolute	*kaʔmʌ~kaʔm*	'downwards' (hill, river)
		tjawʌ~tjaw	'upwards' (hill, river)
	intrinsic	*tajli*	'under, below'
		mʌlgʌ~eleŋ~laŋkʰa	'above'
		tuʔŋ	'foot (of tree)'
		dʑuba~tɕjo	'top (of tree, stone, bamboo)'
	absolutive intrinsic	*lon*	'down edge (of field)'
postposition	intrinsic	*ʌgaɖi* (<N.)	'in the front (of), before'
		pʌtɕʰaɖi (<N.)	'in the back (of), after'
		bʰitrʌ (<N.)	'inside (of)'
		bahjɾʌ (<N.)	'outside (of)'
	relative intrinsic	*waɾi* (<N.)	'across (*origo* dir.)'
		paɾi (<N.)	'across (*origo* opposite dir.)'
		waɾi-paɾi~paɾi-waɾi (<N.)	'around, all sides'
	intrinsic	*muni* (<N.)	'under, below'

		matʰi (<N.)	'above'
	intrinsic	*sʌmmʌ(n)* (<N.)	'up to, until'
		dʰekʰi(n) (<N.)	'from, since'
	intrinsic	*maʥʰʌ* (<N.)	'middle (of), amongst'
	intrinsic	*biʨ~biʨɛʨʌ* (<N.)	'middle (of)'
		other types	
relator nouns		*lagi* (<N.)	'for'
		nimti (<N.)	'for'
postposition		*mʌddʰe* (<N.)	'amongst'
		ʌnusaɾ (<N.)	'according to'
		bahek (<N.)	'except'
		bina (<N.)	'without'

3.4.6.1.   Native relator nouns

The seven native relator nouns are all dedicated to the expression of location in space. They can be divided into two types of binary spatial relation systems with regard to their semantic and morphosyntactic behaviors, serving two types of frame of reference (Levinson 2003): absolute and intrinsic.

- Absolute frame of reference is conveyed by the two relator nouns *kaʔmʌ~kaʔm* 'downwards' and *tjawʌ~tjaw* 'upwards,' whose landmarks are the hill or the river; they are used as adverbs (in absence of case marking), can be modified by a genitive construction, can function as nominal modifiers, and show a development towards becoming postpositions (§ 3.4.6.1.1).

- Intrinsic frame of reference is conveyed by the five relator nouns *tajli* 'under, below,' *mʌlgʌ~eleŋ~laŋkʰa*[45] 'above,' *tuʔŋ* 'foot (of tree),' and *ʨejo* 'top (of tree),' and *lon* 'down edge (of field).' They do not function as adverbs in absence of locational case

---

[45] In the Lothar variety, *mʌlgʌ* and *eleŋ* are both used in RAK-6, while the morpheme *eleŋ* and *laŋkʰa* are used in RAP-13 and RAP-11.

markers, occur when modified by a genitive construction, and one of them show a development towards becoming postposition (§ 3.4.6.1.2).

3.4.6.1.1. Absolute frame of reference: hill or river

The morphemes *kaʔmʌ~kaʔm* 'downwards' and *tjawʌ~tjaw* 'upwards' situate the referent or Figure (Talmy 1972; 1983; Levinson 2003) with or without a point of reference or Ground (Talmy 1972; 1983; Levinson 2003) in an absolute frame of reference (Levinson 2003), whose landmarks, i.e., in the sense of secondary object of reference which functions as a deictic center (Talmy 1972; Talmy 1983), are either the hill or the river.

A Figure can be situated up or down the hill, or up or down the river; in the case of the river, the direction from where the river takes its source corresponds to up the river or upstream, and the direction to where the river ends its course or mouth of the river corresponds to down the river or downstream. This absolute frame of reference may relate to the traditional environment of the Chepang people who have lived in the hills and along rivers. In such system, although absolute, the position of the speaker at the time of speech intrinsically plays a role in the projection of the Figure in a location, and the function of this position is similar to the *origo* (speaker's position) of a system based on a relative frame of reference. When the speaker goes downhill to the river, the landmark may remain the hill but may also become the river; or else when the speaker is located in a place where there is no hill, the sole landmark becomes the river. The origo plays thus a role in the speaker's choice of landmark, since the river or the hill is not like cardinal coordinates, which remain absolute, but more like the speaker's front, back, left or right side, which changes the projected location as they move.

Such absolute frame of reference is very common in TH languages given the topography of the Himalayas; it is found through different grammatical expressions, such as demonstratives, adverbs, relator nouns, adpositions, case marking, verbs, or directional marking on the verb (DeLancey 1985; Bickel 1994; Bickel 1997; Ebert 1999; Dirksmeyer 2008; Post 2019; Genetti et al. 2021).

Chepang absolute frame of reference is illustrated in Figure 74. The symbol <
indicates the direction of the river source and river mouth.

Figure 74. Absolute frame of reference: hill or river
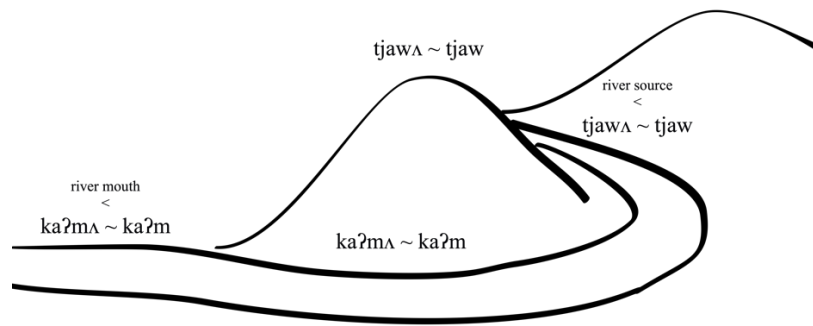
The morphemes *ka?mʌ~ka?m* 'downwards' and *tjawʌ~tjaw* 'upwards' function as
relator nouns. Within the absolute frame of reference they form, the Figure may be
located in the following ways: (a) against a Ground; (b) in absence of Ground; (c) or else
the Ground may be a labeled angle of the frame of reference, i.e., the relator noun itself.

These three types of contextual situations are expressed through different
morphosyntactic constructions, as follows:

(a) When the Ground functions as a point of reference for the location of the
Figure, these relator nouns are modified by a genitive construction, with or without the
presence of locational case markers, as shown in (425) to (428).

(b) When the Figure is located in absence of Ground, these relator nouns function
as adverbs; they can occur with the presence of locational case markers, as in (429) and
(430), or be marked with locational case markers, as shown in (431) to (433).

(c) When the labeled angles of the absolute frame of reference function as
Grounds, these relator nouns function as possessor modifiers, as in (434) and (443), or as
modifiers in a construction similar to a determinative or descriptive compound, i.e.,
preceding the modified head noun, as in (436) and (437).

418

(425) *kim=ko* *kaʔmʌ* *al=ʌ.*
house=GEN downwards go=2SG.IMP.INTR

'Go downwards the house.'

CH_MKW_PC_SIL_E

(426) *mʌndiɾ=ko* *tjaw=kʰe,* *sjaʔm=lʌm* *mu=na=i,*
temple=GEN upwards=DIS Tamang=PL COP=NPST=PL

Tamang=lʌm *mu=na=i.*
Tamang=PL COP=NPST=PL

'Upwards the temple, live the Tamangs, live the Tamangs.'

CH_MKW_MRNDC_SIL_081818_2_Chepang_Language_Culture

(427) *ŋa=ko* *kim=ko* *kaʔm=haŋ,*
1SG=GEN house=GEN downwards=LOC1

ja=ʥo *ɾak-si-siŋ* *mu=na.*
one=CL1 sal-tree-tree COP=NPST

'There is a Sal tree downwards my house.'

CH_MKW_PC_SIL_E

(428) *i* *tɕuɾtɕ=ko* *tjaw=kʰa,* *didi=ko*
PROX church=GEN upwards=LOC2 elder.sister=GEN

kim *mu=na.*
house COP=NPST

'Upwards this church, there is elder sister's house.'

CH_MKW_PC_SIL_E

(429) *kaʔmʌ,* *Besi Tol* *pʌɾ=na.*
downwards Besi Tol have.to_fall=NPST

'Besi Tol is situated downwards.'

CH_MKW_PMRC_SIL_081818_1_Life

(430) *o*   *Maisirang=sʌj*   *dah=o*   *ja=ʥjo*
       DIST   Maisirang=ABL   reach=NMZ:REL   one=CL1

       keti   *tjawʌ*   *hʌ.*
       girl   upwards   COP

       'There is one girl upwards who arrived from that Maisirang.'
       CH_MKW_RC_JMC_SIL_120119_Conversation

(431) *naŋ=kʰe*   *kaʔm=haŋ,*   *o*   *ti=kaj*
       2SG=DIS   downwards=LOC1   DIST   water=DAT

       taʔ=sa   *pʌɾ=na.*
       block=NMZ1   have.to_fall=NPST

       'You have to block that water downwards.'
       CH_MKW_PBPC_CHI_110219_3_Canoe

(432) *bʰʌjsi*   *mu=na,*   *tʰoɾ=ma*   *kaʔm=kʰa=le.*
       buffalo   COP=NPST   cow=ADD   downwards=LOC2=DIS

       'There are the buffalos, the goats too, downwards.'
       CH_MKW_SC_DAM_112819_Conversation_with_Bipana

(433) *suk=o*   *kʰe=lʌ*   *hʌj,*   *ane,*
       plant=perf   cop=neg   part   so

       ŋa=i   *kaʔm=sʌj*   *waʔn=o.*
       1SG=ERG   downwards=ABL   bring=PERF

       'I didn't plant (the rice) hey, so I brought it from downwards.'
       CH_MKW_RC_JMC_SIL_120119_Conversation

(434) *kaʔm=ko*   *ɾak-si-siŋ*   *tahŋ=to*   *mu=na.*
       downwards=GEN   sal-tree-tree   be.huge_be.like=NMZ:ADV2   COP=NPST

       'The Sal tree of downwards is huge.'
       CH_MKW_PC_SIL_E

(435) tjaw=ko   gʰas   boŋ=bʌtiko   bʌlʌ-tɕo
      upwards=GEN grass   look.for=SEQ2   a.little

      siŋ=ma   lat=ti   waʔn=ti,
      wood=ADD   carry=SEQ1   bring=SEQ1

      'Having looked for grass upwards, I carried and brought back some wood too,'
      CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

(436) ʌ,   ja=dʑo   kaʔm-dʑetʰa-tɕoʔ=ko   got=haŋ
      uh   one=CL1   downwards-elder.M-child=gen shelter=LOC1

      gʌm=ti   mu=na=ŋ.
      put_keep=SEQ1   COP=NPST=1

      'Uh, I'm keeping one in the shelter of my elder son of downwards.'
      CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

(437) jat   suɾka dwi   suɾka tjaw-jam   wahɾ=o,
      one   leaf   two   leaf   upwards-rice   sow=PERF

      'I sowed one or two leaves of the upwards rice seeds,'
      CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

In addition, examples are found where these relator nouns follow a noun with no presence of genitive marker, showing the development of a morphosyntactic behavior similar to postpositions (d). In such examples, the relator nouns do not occur without locational case markers. This is illustrated in (438) and (439).

(438) ʌni   Aisarang   kaʔm=kʰa   mʌ=na=i.
      so   Aisirang   downwards=LOC2   COP=NPST=PL

      'So (the in-laws) live downwards Aisirang.'
      CH_MKW_SC_DAM_112819_Conversation_with_Bipana

(439) Manahari   kʰola tjaw=taŋ   pahj=sa   pʌɾ=na.
      Manahari   river upwards=ALL   leave_go=NMZ1   have.to_fall=NPST

'We have to go upwards the Manahari river.'
CH_MKW_MRC_LPK_080918_1_Life

While the landmarks of the morphemes *kaʔmʌ~kaʔm* 'downwards' and *tjawʌ~tjaw* 'upwards' are the hill or the river, the morpheme *kaʔmʌ~kaʔm* may refer to the underworld, realm of divinities underneath the earth, as in (440), which is also called *pʌtal*, as in (442), and the morpheme *tjawʌ~tjaw* to the direction of the abode of the gods above in the sky, as in (442) and (443), or simply to the sky, as in (444).

(440) *sat*   *tʌla=taŋ*   *sjaw=na,*   *kaʔmʌ=tɕʌhe,*
       seven   level=ATT   become=NPST   downwards=DIS

       'There are seven levels, in the underworld,'
       CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(441) *pʌtal=haŋ=tɕʌhene*   *nʌw*   *tʌla*   *sjaw=na,*
       underworld=LOC1=DIS   nine   level   become=NPST

       akas=haŋ=tɕʌhene   *sat*   *tʌla*   *sjaw=na.*
       sky=LOC1=DIS   seven   level   become=NPST

       'In the underworld, there are nine levels, and in the sky, there are seven levels.'
       CH_CTW_BBC_PID_011520_2_Underworld

(442) *- ŋi=ko*   *i=haŋ*   *mu=lʌ=le*   *dʌj,*
       1PL=GEN   PROX=LOC1   COP=NEG=DIS   PART

       si=a   *pʌtɕʰi*   *tjawʌ,*
       die=NMZ2   after   upwards

       - ga=haŋ ?
       INT=LOC1

       - sʌrgʌ=haŋ.
       abode.of.the.gods=LOC1

       '- Our (soul) does not stay here, hey, after dying, (it goes) upwards,

- Where?

- In the abode of the gods.'

(443) ŋa=kaj   ʌ,   i   tjaw=ko   prʌbʰu=i=ma
      1SG=DAT   uh   PROX   upwards=GEN lord=ERG=ADD

      ŋa=kaj   at   bʌj=pa=ta=ŋ.
      1SG=DAT   hand   give=OPT=INV=1

'To me, uh, may this lord of up there too give me a hand.'

(444) tjaw=sʌj,   sʌrgʌ=sʌj   waŋ=ma=tʌ
      upwards=ABL   abode.of.the.gods=ABL   come=NEG=NMZ:ADV1

      pʰere   mʌkʌj sjaw=lʌ.
      again   corn   become=NEG

'Without (water) coming from the sky, yet there is no corn.'

The morphemes *kaʔmʌ~kaʔm* 'downwards' and *tjawʌ~tjaw* 'upwards' likely originally come from nouns since they can function as labelled angles modifying nouns in genitive constructions or in determinative or descriptive types of compound. This is illustrated again in an example where (b) and (c) construction types (see above) are expressed using *kaʔm~kaʔmʌ* 'downwards,' and where the relator noun shows the morphosyntactic behavior of a postposition ((d) construction type).

Such relator nouns may occur outside a genitive construction when used as adverbs (b), where it may syntactically fall between a nominal head and a predicate, such as the construction {noun – relator noun – predicate}. This construction can trigger the reanalysis of the adverbial relator noun as a postposition following a noun without the need of being modified by a genitive construction, especially when the noun that the relator noun follows does not only function as the Figure and the S or A argument of the verb, as in (446), by contrast with (445), but is also a pronoun, which is deictic and hence

in this construction represents also the Ground against which the Figure is situated in the locational adverbial expression, such as: 'Do you go downwards (of yourself) too?'

(445) *ane* *kaʔm (b),* *u* *kaʔm-ljam (c)* *pʌti=ko* *o,*
  so downwards rem downwards-path side=GEN DIST

  ŋi=ko   *Chairang* *pastʌr=ko* *kim* *kaʔm=taŋ=ko (c, d)*
  1PL=GEN Chairang pastor=GEN house downwards=ALL=GEN

  bʌd͡ze=kusi *ta=na=ŋ,*
  grandmother=COM tell.story=NPST=1

'So downwards, there, I tell stories with the grandmother of around downwards the house of our pastor from Chairang of the side of the downwards path.'
CH_MKW_SC_SIL_010220_3_Life

(446) *ʌni* *naŋ=kʰe* *sʌdʰʌj* *i=taŋ=le* *mu=te=na* *ja,*
  so 2SG=DIS always PROX=ALL=DIS COP=2=NPST or

  naŋ *kaʔm=taŋ=ma* *al=te=na* *ja ?*
  2SG downwards=ALL=ADD go=2=NPST or

'So, do you always stay here, or do you go downwards too?'
CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

3.4.6.1.2. Intrinsic frame of reference

The second type of native relator nouns comprises five morphemes: *tajli* 'under, below,' *mʌlgʌ~eleŋ~laŋkʰa* 'above,' *tuʔŋ* 'foot (of tree),' *d͡zjuba~tɕjo* 'top (of tree, stone, bamboo, arrow),' and *lon* 'down edge (of field).' They indicate the location of the referent or Figure with regard to a point of reference or Ground, within an intrinsic frame of reference.

Morphosyntactically, by contrast with the first type of relator nouns, they do not function as adverbs without locational case markers and are not used as nominal modifiers in determinative or descriptive type compounds. They primarily occur as

relator nouns modified by a genitive construction and marked with locational case markers, as shown in (447) to (459). Note that these relator nouns do not distinguish locations that entail or do not entail surface contact.

The relator nouns *tuʔŋ* 'foot (of tree),' *dʑuba~tɕjo* 'top (of tree, stone, bamboo, arrow),' and *lon* 'down edge (of field)' are different from the two relator nouns *tajli* 'under, below' and *mʌlgʌ~eleŋ~laŋkʰa* 'above,' for which the Ground can be any point of reference; their intrinsic frame of reference is already determined by specific Grounds, i.e., either a tree, stone, bamboo, arrow, or the down side of a field, since they have preserved their original semantics[46] and are not attested with any other point of reference. Such relator nouns could be characterized as representing their own absolute frame of reference where the specific Ground functions as a landmark against which the Figure can be located: along a vertical axis for a tree, stone or bamboo, or on an horizontal axis along a downhill side for a field.

Note that in the Lothar variety of RAK-6 and in the Manahari variety of RAK-8 and MAN-7, both the morphemes *dʑuba*[47] and *tɕjo* 'top (of tree, stone, bamboo, arrow)' are attested, while the allomorph *tɕjo* 'top (of tree, stone, bamboo)' is primarily used in the Lothar variety of RAP-13. Examples of the morpheme *dʑuba~tɕjo* 'top (of tree, stone, bamboo)' are given in (453) to (457).

(447) *puɕi   tejbʌl=ko   tajli=haŋ   mu=na.*
       cat   table=GEN   under=LOC1   COP=NPST

'The cat is under the table.'
CH_MKW_BBC_SIL_032920_Topological_Questionnaire_31_6

---

[46] In the Handikhola variety (MAN-4), the morpheme *tɕjo* 'top (of tree, stone, bamboo)' is primarily used and has further developed the meaning of first (or second) floor of a house.

[47] The term *dʑuba* 'top (of tree, stone, bamboo)' is likely a compound whose first syllable is the same morpheme as in the noun *dʑuɾi*, also likely formed through compounding, which refers to the crown of the head where hair forms a hair whorl. The morpheme *tɕjo* is either the result of a simplified form that underwent clipping and devoicing, or of another origin.

(448) - *ga=haŋ*   *mu=na,*   *ŋa=ko*   *glas ? (...)*
        INT=LOC1   COP=NPST   1SG=GEN   glass

        - *ŋa=ko*   *tajli=kʰa*   *hʌ*
        1SG=GEN   under=LOC2   COP

        '- Where is my glass?

        - It's underneath me [the chair where the person is sitting].'

        CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

(449) *puɛi*   *tejbʌl=ko*   *tajli=sʌj*   *waŋ=o.*
        cat   table=GEN   under=ABL   come=PERF

        'The cat has come from under the table.'

        CH_MKW_BBC_SIL_032920_Topological_Questionnaire_31_4

(450) *tejbʌl=ko*   *mʌlgʌ=haŋ*   *naj*   *mu=na.*
        table=GEN   above=LOC1   clothe COP=NPST

        'There are clothes above/on/up the table.'

        CH_MKW_BBC_SIL_032920_Topological_Questionnaire_29_2

(451) *pahad=ko*   *mʌlgʌ=haŋ*   *mus*   *mu=na.*
        hill=GEN   above=LOC1   cloud COP=NPST

        'There is a cloud above/on/up the hill.'

        CH_MKW_BBC_SIL_032920_Topological_Questionnaire_36_1

(452) *ɾu*   *siŋ=ko*   *tʰuta=ko*   *mʌlgʌ=sʌj*   *al=dej*   *mu=na.*
        snake tree=GEN stump=GEN above=ABL go=PROG   COP=NPST

        'The snake is going from above the stump of the tree.'

        CH_MKW_BBC_SIL_032920_Topological_Questionnaire_43_1

(453) *ʌni*   *o*   *dwi=ta*   *manta=kʰe,*   *ʌ,*   *ŋi=ko*
        so   DIST two=CL person=DIS uh 1PL=GEN

        tɛjas=ko   *dʑuba=haŋ=taŋ*
        bamboo.bark=GEN   top=LOC1=ATT

lʌt=ti　　　　*mu=o*　　　　*rʌjsʌ.*

climb=SEQ1　　COP=PERF　　COP.MIR

'So it turned out that two people, uh, were climbing on the top of our bamboo bark.'

CH_CTW_RC_KCR_101920_2_Myth_Origin

(454)　*ane*　*siŋ=ko*　　　*ʥuba=kʰa*　　*lʌhn=ti=taŋ,*
　　　　so　　tree=GEN　　top=LOC2　　climb=SEQ1=ATT

　　　　kim　　*lat=ti=taŋ*　　　　*mu=na=ʨʌ.*
　　　　house　carry=SEQ1=ATT　　COP=NPST=1/3DU

'So, they two climbed to the top of the tree carrying the house and stayed there.'

CH_MKW_DBC_MAI_1_020320_The two brothers

(455)　*"ʨjobaŋ"=tʌ*　　　*dahj=ti*　　*to=ti,*
　　　　Chepang=REP　　　say=SEQ1　　tell_say=SEQ1

　　　　"baŋ=ko　　*ʥuba=sʌj*　　*waŋ=o*　　　　　　*ʥat*　　*ni=paj,"*
　　　　stone=GEN　　top=ABL　　come=NMZ:REL　　　group　1PL=DIS

　　　　i　　　*bʰram=haŋ*　　　*par=ʌ=sa*　　　　　　*kam*
　　　　PROX　mistake=LOC1　　fall.into_make=LN=NMZ1　work

　　　　ʥahŋ=sa　　　　*sjaw=lʌ.*
　　　　do_make=NMZ1　　happen=NEG

'They say "ʨjobaŋ," and (they say) "We are a group that comes from the top of the stones," falling into this mistake/creating confusion is not doing the work.'

CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(456)　*siŋ=ko*　　　*ʨjo=haŋ,*　　*wa*　　*sjah=na*
　　　　tree=GEN　　top=LOC1　　bird　dance=NPST

　　　　pʰiriri,　　　　*pahj=ʨʌ*　　　　　*kanʨʰi,*
　　　　ONO　　　　　leave=1DU.IMP.INTR　younger.one.F

'On top of the tree, the bird dances *phiriri*, let's the two of us go, little one,'

427

CH_MKW_MK_BAN_102417_Song_2

(457) *siŋ=ko*   *tɕjo=kʰa*   *al=ŋʌ=to.*
tree=GEN   top=LOC2   go=1=REM.PST

'I used to go to the top of trees.'
CH_MKW_1_BMB_BAN_100517_3_E

(458) *ʌni*   *sal*   *aʔl=ti*   *baba=i,*
so   placenta   take.away=SEQ1   father=ERG

siŋ=ko   *tuʔŋ=haŋ*   *djaw=aktiko*

tree=GEN   foot=LOC1   dig=SEQ2

saʔ=haŋ   *pahm=na,*   *kja !*
earth=LOC1   burry=NPST   PART

'Then, my father, having taken away the placenta, having dug at the foot of a tree, it is buried in the ground!'
CH_MKW_BMB_BAN_100617_5_Birth

(459) *siŋ=ko*   *tuʔŋ=kʰa=le*   *gʌm=sa*   *pʌr=na.*
tree=GEN   foot=LOC2=DIS   put=NMZ1   have.to_fall=NPST

'We have to put it at the foot of a tree.'
CH_CTW_BBC_PID_011520_5_Underworld

As a side anthropolinguistic note, it has been claimed by Caughley (1982; 2000) that the term Chepang, which is sometimes pronounced [tɕjobãŋ], could convey the etymological meaning of 'top of stone' and thus would reveal that the name of the Chepang people relates to their "mountainous" habitat. While Caughley (1982: 2) himself notes that is not a satisfactory explanation, such folk etymology has nevertheless spread greatly, and in particular through the publication in Nepal of *The people of the stones* by Rai (1985) and the Centre for Nepal & Asian Studies, Tribhuvan University.

This analysis is highly unlikely and often rejected by community members who consider that such statement is unfounded and harmful, participating in the derogatory

image of the Chepang people spread by other communities holding social, educational, and political powers, since it may be taken as a historical fact by the Chepang community and beyond – it has already inundated the imaginary collective and anthropological literature up to very recently: Rai (1985), Viel (2020: 230). Further, the term *Cyobang* or *Cyo'bang* was chosen as an official name to refer to the language and the people by Caughley (2000), which ultimately contributes in strengthening this idea.

The present analysis of the relator noun *dʑjuba~tɕjo* 'top (of tree, stone, bamboo, arrow),' shows evidence that it is indeed not the case, since, first, such type of relator noun does not morphosyntactically precede a noun to modify it as in a determinative or descriptive compound, i.e., a form **dʑjuba~tɕjo-baŋ* is not grammatically possible; further, as mentioned in footnote[47], the original morpheme for 'top (of tree, stone, bamboo, arrow),' is likely the compound *dʑjuba*, since another compound formed with *dʑju-* is attested, i.e., *dʑjuɾi* 'crown of head,' both referring to a top part; finally, the pronunciations of the term Chepang, in addition to [tɕepãŋ] and [tɕjobãŋ], may as well be [tɕjʌbãŋ], [tɕjebãŋ], [tɕjʌpãŋ], [tɕjopãŋ], [tɕjepãŋ], all of which finding an explanation in articulatory processes and morphophonological changes: palatalization following the affricate [tɕ], such as [tɕʲ], when preceding a front mid-close vowel [e]; anticipatory or regressive assimilation for the front mid-close vowel [e] which may change into [ʌ] or [o], assimilating a backness feature of the central low vowel [a] of [pãŋ], which is pronounced further back; and last, the voicing of the initial consonant [p] of the second syllable [pãŋ], such as [bãŋ], can be analyzed as a result of voicing assimilation, since [p] is pronounced between two vowels, which are voiced segments.

Another possible origin for the name Chepang, also put forward by Caughley (1982: 2–3), is discussed in § 1.4.

As mentioned in footnote[45], in the Lothar varieties, the morphemes *mʌlgʌ* and *eleŋ* 'above' are primarily used in RAK-6 while *eleŋ* and *laŋkʰa* 'above' are used in RAP-13 and RAP-11. The etymological origin of *mʌlgʌ* and *eleŋ* is unknown, but *laŋkʰa* likely comes from the noun *laŋka* 'sky,' which occurs in shamanic chants and discourse, as illustrated in (460) and (461).

In RAP-13, the morpheme *laŋkʰa* also means 'sky' but extended its meaning to 'the first (second) floor of the house,' as shown in (462) and (463). The change from the

syllable /ka/ to /kʰa/ likely occurred by analogy with the locative case marker =kʰa since *laŋkʰa* is used without the presence of a locative marker when functioning as an adverb; it may nevertheless be found with the locational case marker =*haŋ*. This is illustrated in (462) to (464).

Like the other relator nouns of its type, the relator noun *laŋkʰa* can be used to mean 'above' and occurs modified by a genitive construction, as in (465). By contrast with the other relator nouns, it is not found marked by locational case markers in this construction. In addition, the relator noun *laŋkʰa* 'above' developed the morphosyntactic behavior of a postposition, as shown in (466). This is the only relator noun of this type, along with *tajli* 'under, below,' described below, to function like a postposition.

(460) *"laŋka-dʑaʔ"=tʌ*   *o=tʌ*   *dahj=na=i,*
       sky-tiger=REP   DIST=NMZ:ADV1   say=NPST=PL

       joh-nʌm   *din=ko*   *joh-nʌm*   *baɾ=ko,*
       yesterday-CL   day=GEN   yesterday-CL   day=GEN

       '"Tiger of the sky," they say like that, in the days of the past, in ancient times,'
       CH_MKW_KMC_SK_110319_4_Shaman_Song

(461) *laŋka=sʌj*   *tuɾ=dʑe=mʌj,*   *lʌhʌɾa=sʌj*   *kɾjuk=dʑe=mʌj*
       sky=ABL   pull=EPIS=IMP.HON   creeper=ABL   catch=EPIS=IMP.HON

       'Pull (the spirit) from the sky, catch it from the creeper.'
       CH_MKW_BLC_SIL_113019_1_Shaman

(462) *dʑʰʌn*   *laŋkʰa dah=waj=taŋ=a,*   *akaʃ*   *dah=waj=taŋ=a,*
       furiously   sky   reach=CERT=ATT=PST   sky   reach=CERT=ATT=PST

       '(The Flame tree) furiously reached the sky, reached the sky,'
       CH_CTW_JBC_BHR_102420_1_Cing_Lan

(463) *ama=kʰe=ko*   *palo*   *laŋkʰa*   *lʌt=ti*   *ʌ,*
       mother=DIS=GEN   turn   first.floor   climb=SEQ1   uh

       dam   *hʌɾ=sa=taŋ*   *tʰal=ka=n,*
       coin   fill.up_measure=NMZ1=ATT   begin=2/3.PST=DIR/TR

       'It was the mother's turn, climbing up the first floor, uh, she started to fill up (a
       measuring recipient) with coins,'
       CH_CTW_BBC_POL_111720_5_Two Sisters

(464) *laŋkʰa=haŋ*   *njas=ja*   *pʌtɕʰi,*
       first.floor=LOC1   lay.down=NMZ2   after

       "lʌw   *moʔm,*   *njas=te=ka=tɕi,"*
       well   granddaughter   lay.down=2=2/3.PST=2SG>1SG

       'After she lied him down on the first floor, (he said) "Well, granddaughter, you
       laid me down,"'
       CH_CTW_BBC_POL_111720_5_Two Sisters

(465) *puɛi*   *gundʌɾi=ko*   *laŋkʰa*   *eʔn=ti*   *mu=na.*
       cat   straw.mat=GEN   above   sleep=SEQ1   COP=NPST

       'The cat is sleeping on the straw mat.'
       CH_CTW_SPC_POL_E

(466) *o*   *tʌɾwaɾ,*   *awsi-puɾne=ko*   *bela=haŋ,*
       DIST   sword   new.moon=GEN   moment=LOC1

       gʌ=tʌ   *tahŋ=to*   *mʌ=na=i,*
       INT=NMZ:ADV1   be.huge_be.like=NMZ:ADV2   COP=NPST=PL

       bʰiɾ=ʌ=ti,   *sʌɾlʌk=tʌ*   *bʰiɾ=ʌ=bʌtiko,*
       put.on=LN=SEQ1   straight= NMZ:ADV1   put.on=LN=SEQ2

       swat=tʌ   *ud=ʌ=ti,*
       all.at.once=NMZ:ADV1   fly=LN=SEQ1

siŋ	*laŋkʰa*	*lʌt=sa*	*kʰaj=to,*
tree	above	climb=NMZ1	be.able=REM.PST

bina		*pap,*	*pap=le*	*na=lʌ.*
without		wing	wing=DIS	COP=NEG

'(My great-grandfather and great-great-grandfather), during the new moon, how are they like, putting on that sword, putting straight on that sword, all at once flying, they would be able to climb up the trees, without wings, they had no wings.'

CH_CTW_RLC_JIM_101920_Language and Culture_Archive

The morpheme *lon* 'down edge (of field)' falls within an intrinsic frame of reference but semantically conveys an absolute location since the specific location of the edge of the field is downwards the hill. This is illustrated in (468).

(467)	*ɾaŋ=ko*	*lon=haŋ,*	*mʌkʌj*	*suk=alaŋ.*
	field=GEN	down.edge=LOC1	corn	plant=1.PST

'I planted corn on the down edge of the field.'
CH_MKW_PC_SIL_E

(468)	*ɾaŋ=ko*	*lon=kʰa,*	*wa*
	field=GEN	down.edge=LOC2	bird_hen

wah=ti	*mʌ=na.*
walk=SEQ1	COP=NPST

'Hens are walking along the down edge of the field.'
CH_MKW_PC_SIL_E

The morpheme *tajli* 'under, below' developed the morphosyntactic behavior of a postposition, following nouns without the presence of the genitive morpheme *=ko*, as in (469). This change also coincides with another change, a semantic extension that led the morpheme *tajli* 'under, below' to be used in place of the morpheme *kaʔmʌ~kaʔm* 'downwards' expanding its frame of reference from intrinsic to absolute. Since the

432

morpheme *kaʔmʌ~kaʔm* 'downwards' has developed a postpositional use, it is likely that the semantic extension of *tɑjli* 'under, below > downwards' took place in this particular setting of postpositional construction. This is illustrated in (470) and (472).

(469) | *puɕi* | *tejbʌl* | *tɑjli=haŋ* | *mu=na.* |
| cat | table | under=LOC1 | COP=NPST |

'The cat is under the table.'

CH_MKW_BBC_SIL_032920_Topological_Questionnaire_31_8

(470) | *- e,* | *bʰʌkʰʌɾ* | *duŋ=ti* | *mu=na,* |
| EXPR | just | sprout=SEQ1 | COP=NPST |

| *- i* | *kim* | *tɑjli=ma,* | *ma,* | *kim* | *tɑjli=ma,* |
| PROX | house | under=ADD | yes | house | under=ADD |

| *i* | *kim=ko* | *tjaw=ko=ma.* |
| PROX | house=GEN | upwards=GEN=ADD |

'- Okay, the (mustard) is sprouting just now,

- Downwards this house, yes, downwards the house too, and on the upwards side of this house too.'

CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

(471) | *kim* | *tɑjli* | *ɾuiŋ* | *suk=o.* |
| house | under | bamboo | plant=PERF |

'Downwards the house, some bamboo was planted.'

CH_CTW_BBC_POL_111720_6_Jogi

(472) | *ŋa* | *Simmu* | *tʰan* | *ʌ,* | *tɑjli* | *sʌmmʌn,* |
| 1SG | Simmu | place | uh | under | up.to |

| *ŋa* | *dʑʰja=ti* | | *dah=o* | *mʌ=na,* | *didi.* |
| 1SG | shaman.practice=SEQ1 | | reach=PERF | COP=NPST | elder.sister |

433

'I have reached up to downwards the village of Simmu practicing shamanic ceremonies, elder sister.'

CH_MKW_DLC_CYO_120119_Shaman_Life

(473) i   kim   tajli=haŋ   mʌj-saj   tʰam   mʌ=na.
       PROX   house under=LOC1   banana-fruit.seed   stem   COP=NPST

'There is a banana tree downwards this house.'

CH_CTW_SPC_POL_E

Another semantic extension has taken place, but here the other way around, the morpheme *tjaw* 'upwards,' which falls within an absolute frame of reference, as illustrated again in (474), replaces the use of *mʌlgʌ~eleŋ~laŋkʰa* 'above,' expanding its frame of reference to an intrinsic one, as shown in (475).

(474) - tjaw=sʌj   rʰam=ti   mu=o   rʌjsʌ,
       upwards=ABL   graze=SEQ1   COP=PERF   COP.MIR

       awas *bʌj=ti*   *mu=to,*   *bʰrut-bʰruttʌ,*
       noise   give=SEQ1   COP=REM.PST   ONO

       - gʌjda=i ?
       rhinoceros=ERG

       - ma   hʌj,   gʌjda=i !
       yes   PART   rhinoceros=ERG

'- It turns out it was grazing from upwards/up the river, it was making sound,
bhrut-bʰrutta,
- A rhinoceros?
- Yes of course, a rhinoceros!'

CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(475) i   kitap   tejbʌl=ko   tjaw=haŋ   mu=na.
       PROX   book   table=GEN   above=LOC1   COP=NPST

'This book is above the table.'

CH_MKW_PC_SIL_E

The relator nouns *tajli* 'under, below,' *mʌlgʌ~eleŋ~laŋkʰa* 'above,' *tuʔŋ* 'foot (of tree),' *dʑjuba~tɕjo* 'top (of tree, bamboo, stone, arrow),' and *lon* 'down edge (of field)' likely originate from nouns. The three morphemes *tuʔŋ* 'foot (of tree),' *dʑjuba~tɕjo* 'top (of tree, bamboo, stone, arrow),' and *lon* 'down edge (of field)' are still used as such, as illustrated in (476) to (480).

(476) | *siŋ=ko* | *tuʔŋ* | *boʔn=to* | *mu=na.* |
|---|---|---|---|
| tree=GEN | foot | be.big=NMZ:ADV2 | COP=NPST |

'The foot of the tree is big.'

CH_MKW_PC_SIL_E

(477) | *o* | *dʑjuba=i* | *tɕʰu=ja=ma* |
|---|---|---|
| DIST | top=ERG | touch=COND=ADD |

o	*han=to*	*sjaw=na.*
DIST	be.sharp=NMZ:ADV2	become=NPST

'Even if that upper part (of the arrow) touches (anything/anyone), it feels sharp.'

CH_MKW_SCBKC_SIL_081918_3_Chepang_culture

(478) | *u,* | *tɕjo=tɕʌhe* | *ʌ,* | *gʰui=ti* | *kim=ko* | *huŋ=haŋ=taŋ* |
|---|---|---|---|---|---|
| CAT | top=DIS | uh, | bend=SEQ1 | house=GEN | roof.ridge=LOC1=ATT |

tɕʰu=ti	*mʌ=to*	*kʰe=to.*
touch=SEQ1	COP=NMZ:ADV2	COP=REM.PST

'That one, the top (of the bamboo), uh, having bent, touched the roof ridge of the house and was sitting (there).'

CH_CTW_BBC_POL_111720_6_Jogi

(479) | *ni=ko* | *tɕjo=taŋ=tɕʌhene,* | *o=ŋ=sʌj* |
|---|---|---|
| 1PL=GEN | top=ATT=DIS | DIST=LOC1=ABL |

ni=ko=tɕʌhene,	*o,*	*do*	*do*	*sjaw=o*	*mʌ=na,*
1PL=GEN=DIS	DIST	what	what	become=PERF	COP=NPST

dzʌmma=le	*tɕjew=na.*
ALL=DIS	see_find=NPST

'As for our top/the top above us, from there, as per our (belief), that (thing), whatever has happened, everything is seen.'

CH_CTW_BBC_PID_011520_2_Underworld_Archive

(480)

ɾaŋ=ko	*lon*	*pe=to*	*mu=na.*
tree=GEN	down.edge	be.nice=NMZ:ADV2	COP=NPST

'The downside of the field is nice.'

CH_MKW_PC_SIL_E

Finally, note that the ablative case marker *=sʌj*, which also functions as a comparative marker (§ 3.4.5.7), as shown in (481), can serve the marking of the Ground against which the Figure is located. In such construction, the Ground can be anything, as in (483), or the labeled angle itself, as in (484).

Such construction has developed under the influence of Nepali which uses the comparative marker भन्दा <bhandā>, in (482), in a similar construction with some relator nouns and postpositions, as shown in (485).

(481)

naŋ,	*naŋ=ko*	*ba=sʌj*	*jas=to*	*mu=te=na !*
2SG	2SG=GEN	father=CMPR	tall=NMZ:ADV2	COP=2=NPST

'You are taller than your dad!'

CH_CTW_20-21_KC_SIL_082420_E

Nepali

(482)

बारुलाले	भन्दा	अरिंडालले	सारो	चिन्छ।
bārulā-le	*bhandā*	*aringāl-le*	*sāro*	*cil-cha.*
wasp-ERG	COMP	hornet-ERG	hard	sting-3SG.NPST

'A hornet stings worse than a wasp.'

(Schmidt 1994: 21)

(483) *i* *tahŋ=o*   *ɾak-si-siŋ=sʌj* *kaʔm=taŋ,*
PROX be.huge_be.like=NMZ:REL sal-tree-tree=CMPR downwards=ALL

didi=ko   *kim* *mu=na.*
elder.sister=GEN house COP=NPST

'Downwards this huge Sal tree, there is elder sister's house.'
CH_MKW_PC_SIL_E

(484) *o* *mʌkʌj* *biɾuwa=ko* *loʔ=ko* *tajli=sʌj,*
DIST corn shoot=GEN   leaf=GEN under=CMPR

ʌ, *sjaʔn=ko* *um* *mu=na.*
uh insect=GEN egg COP=NPST

'Under the leaf of that corn shoot, there are insects' eggs.'
CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

Nepali

(485) भौसी   बाखा   फुकाएर, घर＝भन्दा   माथि
bhaisi   *bākhrā*   *phukā-era,* *ghar=bhandā*   *māthi*
buffalo   goat   unfasten-SEQ house=CMPR   above

छिमेकीकोमा   बस्यौँ।
chimeki-ko-mā   *bas-yõm.*
neighbour=GEN=LOC live=1PL.PST

'After having unfastened the buffalos and goats, we lived at the neighbors' above
(our) house.'

Ukera (online News) - https://www.ukeraa.com/news/2021/06/26/7101

3.4.6.2. *ljam* 'path' compound to nominal directional to ablative case

The morpheme *ljam~ljamduŋ* is a noun meaning 'path, way,' as illustrated in (486) to (488). This morpheme is found as semantic head in a determinative or descriptive compound whose modifier can be a relator noun, as shown in (489) and (490) with *kaʔm* 'downwards,' a toponym, as in (491) and (492), or else the body part *kah* 'back,' as in (493). Such relator nouns and nominals function as Grounds for the location of the Figure within an absolute frame of reference.

(486) o   *ljam=kaj=le*   *rʌj=na=ŋ,*   *didi.*
       DIST   path=DAT=DIS   be.scared=NPST=1   elder.sister

'I'm scared of that path, elder sister.'

CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(487) *Ratan=kaj*   *ljam=haŋ*   *sat=a*   *pʌtɕʰi,*
       Ratan=DAT   path=LOC1   kill=NMZ2   after

       Ratan=ko   *talaŋ*   *lat=ti*   *mu=na=i,*
       Ratan=GEN   head   carry=SEQ1   COP=NPST=PL

       o   *sena=lʌm,*
       DIST   army=PL

'Having killed (the king) Ratan on the path, they are carrying Ratan's head, those soldiers,'

CH_MKW_PSC_MAI_012720_Local_History_1

(488) *bat*   *maɾ=ʌ=ti=taŋ*   *pahj=taŋ=ka=tɕʌ,*   *ljamduŋ=haŋ,*
       conversation   kill=LN=SEQ1=ATT   leave=ATT=2/3.PST=1/3DU   path=LOC1

'While talking, the two left on the path.'

CH_CTW_BBC_GUN_102620_1_Cing_Lan

(489) *sipahi-pʌltʌn=lʌm*   *kaʔm-ljami*   *gʌm=bʌte,*
       soldier-regiment=PL   downwards-path   put=SEQ2

munda-siŋ=taŋ	*ton=ti*	*lʰok=n=i=to.*
log-tree=ATT	fall=SEQ1	send=DIR/TR=PL=REM.PST

'Having put the soldiers, the people of the regiment on the downwards path, they would send logs of wood to fall (on them).'

CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

(490)
ane	*kaʔm,*	*u*	*kaʔm-ljam*	*pʌti=ko*	*o,*
so	downwards	REM	downwards-path	side=GEN	DIST

ŋi=ko	*Chairang*	*pastʌr=ko*	*kim*	*kaʔm=taŋ=ko*
1PL=GEN	Chairang	pastor=GEN	house	downwards=ALL=GEN

bʌʣe=kusi	*ta=na=ŋ,*
elder.F=COM	tell.story=NPST=1

'So downwards, there, I tell stories with the grandmother of around downwards the house of our pastor from Chairang of the side of the downwards path.'

CH_MKW_SC_SIL_010220_3_Life

(491)
Sisneri-ljami	*ljam*	*bʰʌnda*	*sjaw=ti,*
Sisneri-path	path	close	become=SEQ1

Palung-ljami	*gʰum=ʌ=ti*	*al=da=na.*
Palung-path	go.about=LN=SEQ1	go=PROG=NPST

'The path of the Sisneri path being close, we diverted going on the Palung path.'

CH_MKW_BMB_BAN_103119_6_Conversation_Wife and husband

(492)
nik=to	*maru*	*sʌrʌrʌ*
be.cool=NMZ:ADV2	wind	ONO

Pengcya-ljam=sʌj	*al=tɕʌ*	*Manahari !*
Pengcya-path=ABL	al=1DU.IMP.INTR	Manahari

'With the cool wind *sarara*, let's the two of us go from the path of Pengcya to Manahari!'

CH_CTW_SC_SP_POL_102620_2_Raksi_lo'hang_Song

(493) *o=ŋ=sʌj*   *bʌʥe=kaj=taŋ*   *kah=ljam=sʌj*
       DIST=LOC1=ABL   grandmother=DAT=ATT   back=DIR=ABL

       tɾʌjŋ=ti   *lʰok=u=to,*

       hit.with.stick=SEQ1   send=3O/DIR=REM.PST

       'He (the grandson bat) hit the grandmother from behind with the stick and sent
       her off.'

       CH_CTW_KMC_TAP_102520_2_The bat and the crab

In such construction, when marked with the ablative case morpheme *=sʌj*, it can
be reanalyzed as a nominal directional morpheme whose function is merely to indicate
that the motion expressed by the verb takes a path or direction towards the Ground, as in
(492) and (493). It is possible that such construction triggered the reanalysis of
=ljam~ljami as an ablative case marker, as illustrated in (494) to (496).

Note that when functioning as an ablative case, the proximal deictic pronoun *i* in
(494) does not entail the presence of the locational morpheme *=haŋ* or *=kʰa* to mean
'here,' as is the case with the ablative marker *=sʌj*, such as, *i=haŋ=sʌj* 'from here.' This
shows further evidence that the directional function of *=ljam~ljami* is a morphosyntactic
path towards its development as an ablative case, since its semantics intrinsically
indicates a path of motion, and hence the deictic morpheme to which it attaches expresses
a location, and not for instance a person.

This historical development from the noun *ljam* 'path' to a nominal directional
may have started at the level of Proto-Chepang-Bhujel (PCB), before pursuing its
development towards an ablative case marker separately in Chepang and Bhujel. Indeed,
while the cognate ablative case marker *-lyam* /ljam/ is used in Bhujel with nominals
expressing spatial locations (Regmi 2007: 169–170), the cognate ablative morpheme *-sei*
/sej/ is found in a sequential temporal subordinate construction (Regmi 2007: 167). In
addition, the function of *=ljam* as an ablative case marker is not attested in all the studied
varieties, but mainly in the Lothar varieties of RAP-13 and RAP-11, while the use of *ljam*
in a determinative or descriptive compound is attested in all studied varieties.

440

The ablative morpheme *=sʌj may be traced back to PCB along with the use of
ljam in a determinative or descriptive compound and as a nominal directional morpheme.

(494) ʌ̃,   al=to=haŋ   i=le,   i=ljam
       yes   go=NMZ:ADV2=LOC1 PROX=DIS   PROX=ABL

       Madi   ghum=ʌ=ti   al=sa.
       Madi   go.about=LN=SEQ1   go=NMZ1

       'Yes, when going, it's this, one goes from here, diverting by Madi.'
       CH_CTW_BBC_PID_011520_7_Chepang_Raute

(495) *ama-ba=ko   kim=taŋ   dʑun   ljam=ljami*
       parents=GEN   house=ALL   whatever   path=ABL

       ŋa=ko   kim=taŋ   djahn=te=a.
       1SG=GEN   house=ALL   come.back=2=PST

       'You came back to my house from whatever path around the parents' house.'
       CH_CTW_BBC_POL_111720_1_Cing_Lan

(496) *tʌɾʌ   tɕini   kʰe=ja=taŋ,   kaʔm-ljam*
       but   sugar   COP=COND=ATT   downwards-ABL

       mʰeʔ   dʌdaw=o   bela=haŋ,
       fire   light.up=NMZ:REL   moment=LOC1

       dzʌm=ʌ=dʰʌj   al=tɕja   kʰe=to.
       curdle=LN=PROG   go=IRR   COP=REM.PST

       'But if there is sugar (in the honey), when lighting up fire from below, it would be
       that it would go curdling / it would start and continue to curdle.'
       CH_CTW_BBC_POL_102420_2_Tu'm

The morphosyntactic behaviors of native relator nouns are summarized in Table
122.

Table 122. Morphosyntactic behavior of native relator nouns

	ADV	ADV	NP	NP	POST	POST/REL. N.	REL. N.	REL. N.
locational (space)	ADV V	**RN**=LOC.CASE V	**RN** N	**RN**=ko N	NP **POST**	NP **POST**=LOC.CASE	NP=ko **RN**	NP=ko **RN**=LOC.CASE
kaʔmʌ~kaʔm 'downwards'	✓	✓	✓	✓	-	✓ <	✓	✓
tjawʌ~tjaw 'upwards'	✓	✓	✓	✓	-	✓ <	✓	✓
tajli 'under, below'	-	-	-	-	✓ <	✓ <	✓ <	✓
mʌlgʌ~eleŋ 'above'	-	-	-	-	-	-	-	✓
laŋkʰa 'above'	✓ <	✓ <	-	-	-	-	-	✓
tuʔŋ 'foot (of tree)'	-	-	-	-	-	-	-	✓
dʑuba~tɕjo 'top (of tree…)'	-	-	-	-	-	-	-	✓
lon 'down edge (of field)'	-	-	-	-	-	-	-	✓

3.4.6.3. Postpositions and relator nouns borrowed from Nepali

Nineteen postpositional elements, postpositions or relator nouns, are borrowed from Nepali and used in the same type of constructions as they occur in Nepali. They are briefly described in the following sub-sections where these are grouped according to their functions and the constructions within which they occur, as follows:

- relator noun, postposition, and adverb: intrinsic frame of reference (§ 3.4.6.3.1)
- relator noun, postposition, and adverb: relative and intrinsic frame of reference (§ 3.4.6.3.2)
- postpositions: *matʰi* 'above, up,' *muni* 'under, below' (§ 3.4.6.3.3)
- postpositions: *dekʰi* 'from, since,' *sʌmmʌ* 'up to, until' (§ 3.4.6.3.4)
- relator nouns: *madʑʰʌ* 'middle of, amongst,' *bitɕ~bitɕtɕʌ* 'middle of' (§ 3.4.6.3.5)
- relator noun to postposition: *mʌddʰe* 'amongst' (§ 3.4.6.3.6)
- relator nouns: *lagi* 'for,' *nimti* 'for, for the sake of' (§ 3.4.6.3.7)
- postpositions: *ʌnusaɾ* 'according to,' *bahek* 'except,' *bina* 'without' (§ 3.4.6.3.8)

The morphosyntactic behavior of relator nouns and postpositions is summarized in Table 123.

Table 123. Morphosyntactic behavior of relator nouns and postpositions

	ADV	ADV	NP	NP	POST	POST	POST	POST/REL. N.	REL. N.	REL. N.
locational (space / time)	ADV V	RN=LOC.CASE V	RN N	RN=ko N	NP=bʰnda POST	NP=sʌj POST	NP POST	NP POST=LOC.CASE	NP=ko RN	NP=ko RN=LOC.CASE
ʌgaɖi 'in the front (of), before'	✓	✓	-	✓	✓	-	✓	✓<	✓	✓
pʌtɕʰaɖi 'in the back (of), after'	✓	✓	-	✓	✓	-	✓	✓<	✓	✓
bʰitrʌ 'inside (of)'	✓	✓	-	✓	✓	-	✓	✓<	✓	✓
bahjrʌ 'outside (of)'	✓	✓	-	✓	✓	-	✓	✓<	✓	✓
wari 'deictic side'	✓	✓	-	✓	✓	-	✓	✓	✓	✓
pari 'deictic opposite side'	✓	✓	-	✓	✓	-	✓	✓	✓	✓
wari-pari~pari-wari 'around'	✓	✓	-	✓	✓	-	✓	✓	✓	✓
muni 'under, below'	✓	-	-	✓	✓	-	✓	-	✓	-
matʰi 'above'	✓	-	-	✓	✓	-	✓	-	✓	-
sʌmmʌ(n) 'up to, until'	-	-	-	✓	-	-	✓	-	-	-
dekʰi(n) 'from, since'	-	-	-	-	-	-	✓	-	-	-
madʰʌ 'middle (of), amongst'	✓	-	-	✓	-	-	-	-	-	✓
bitɕ~bitɕʌ 'middle (of)'	-	✓	-	✓	-	-	-	-	-	✓
other types	ADV V	RN=LOC.CASE V	RN N	NP RN=ko	NP=bʰnda POST	NP=sʌj POST	NP POST	NP POST=LOC.CASE	NP=ko RN	NP=ko RN=LOC.CASE
mʌddʰe 'amongst'	-	-	-	-	-	-	✓	✓<	-	✓
ʌnusar 'according to'	-	-	-	✓	-	✓	✓	✓<	-	-
bahek 'except'	-	-	-	-	-	-	✓	-	-	-
bina 'without'	-	-	-	-	-	-	✓	-	-	-
lagi 'for'	-	-	-	-	-	-	-	-	✓	✓
nimti 'for, for the sake of'	-	-	-	-	-	-	-	-	✓	✓

3.4.6.3.1.   Relator noun, postposition, and adverb: intrinsic frame of reference

The morphemes ʌgaɖi 'in the front of, before,' pʌtɕʰaɖi 'in the back of, after,' bʰitrʌ 'inside of' and bahjrʌ 'outside of' function as relator nouns, postpositions, and adverbs within an intrinsic frame of reference. They situate in space for the later or in space and time for the former, the location of a Figure against a Ground. Some examples are given in (497) to (505). These morphemes also function as adverbs, as in (506) and (507).

(497) *i   dwi   pusta   ʌgadi=ko   kura   hʌ.*
       PROX   two   generation   front=GEN   thing   COP

'This is something that goes back to two generations.'
CH_MKW_BBC_SIL_032820_1_Lanrang

(498) *ʌ=ŋ=sʌj,   o=kaj,   rama=i,   tin   tɕokta,*
       DIST=LOC1=ABL   DIST=DAT   sickle=INST three piece

       talaŋ, ʌni   i=haŋ,   bitɕ=ko   bʰak,
       head   and   PROX=LOC1   middle=GEN   part

443

ʌni	pʌtɕʰaɖi=ko	bʰak,	tɕokta,	tɕokta,	dʑahŋ=a.
and	behind_after=GEN	part	piece	piece	do_make=PST

'From there, we make pieces, one after another, we (cut) it (the bat) with the sickle in three pieces, the head, and here, the middle part, and the back part.'

CH_MKW_MRNDC_SIL_081818_3_Bats

(499)	ʌni	bop=i	o	gʰaŋ=ko,	gʰaŋ	bʰitrʌ=ko
	so	snail=ERG	DIST	hole=GEN,	hole	inside=GEN

	um=lʌm=kaj	kʌn=na=u.
	egg=SML=DAT	look=NPST=3O/DIR

'So, the snail looks at the eggs inside that hole.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(500)	juk=ma	naŋ=ko	ʌgadi=le	kʰe=to	ba.
	rhesus=ADD	2SG=GEN	front=DIS	COP=REM.PST	PART

'The rhesus monkey was definitely in front of you.'

CH_MKW_BMB_KW_BAN_103119_10_Conversation_Friends

(501)	i=tʌ	tahŋ=o	kim=ko	pʌtɕʰadi
	PROX=NMZ:ADV1	be.huge_be.like=NMZ:REL	house=GEN	behind_after

	pʌlʌŋ	nʰap=ti=taŋ	bʌj=ka=nʌ.
	bed	lay.down=SEQ1=ATT	give=2/3.PST=DIR/TR

'In the back of a house like this it gives space to lay down a sleeping place.'

CH_MKW_SC_SIL_122619_2_Dhobini_rani

(502)	ruiŋ=ko	bʰitrʌ	pok=o	bela=haŋ
	bamboo=GEN	inside	enter=NMZ:REL	moment=LOC1

'At the time of entering inside the bamboo,'

CH_MKW_SPMC_SIL_100921_1_Tiɾidu'm_Basʌdu'm

(503) o-mi   kim=bʰʌnda   pʌtɕʰadi   mu=na=i.
      DIST-PL.H   house=CMPR   behind_after   COP=NPST=PL

'They are in the back of the house.'
CH_MKW_1_21-22_CPR_BAN_100817_1_E

(504) kʌti   kuɾa=lʌm,   endʑijo   bʰitɾʌ
      how.much   thing=SML   NGO   inside

      sasʌn=ma   dʌb=ʌ=o   mʌ=na.
      governance=ADD   put.out.of.sight=LN=NMZ:REL   COP=NPST

'How much things, the governance also has been inaccessible inside NGOs.'
CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(505) ŋi=bʰʌnda   ʌgadi,   dwi   pusta   ʌgadi   waŋ=o.
      1PL=CMPR   front   two   generation   front   come=PERF

'Before us, (our ancestors) have arrived (here) before two generations.'
CH_CTW_SRP_GUN_102620_3_Gundi_settlement

(506) ʌni   bimaɾi=lʌm   waŋ=o   bela=kʰa,
      so   sick=PL   come=NMZ:REL   moment=LOC2

      badʑja,   naŋ=i,   kʰas,   ʌgadi   gʌ=tʌ
      elder.M   2SG=ERG   really   front   INT=NMZ:ADV1

      dʑahŋ=te=na=u ?
      do_make=2=NPST=3O/DIR

'So when sick people come, grandpa, really, how do you go about it first?'
CH_CTW_KRC_HAT_012120_Being_Shaman

(507) lʌw   djah,   doh   haj=sa   kja,
      well   now   what   do=NMZ1   PART

      o   jom=pa=taŋ   bahjɾʌ=taŋ   gljuŋ=a.
      DIST   bear=DIS=ATT   outside=ATT   get.out=PST

'Well now, what to do hey, that bear got out outside.'

3.4.6.3.2. Relator noun, postposition, and adverb: relative and intrinsic frame of reference

The morphemes *wari* 'deictic side,' *pari* 'deictic opposite side,' *wari-pari~pari-wari* 'around' situate a Figure within a relative frame of reference, the deictic center is the *origo* or speaker, or within an intrinsic frame of reference where the Figure is located against a Ground which becomes the deictic center. This is illustrated in (508) to (511). In these examples, they function as relator nouns and postpositions, but they can also occur as locational adverbs, with or without the presence of locational case markers, as shown in (512) and (513).

(508) ʌ,   i   *Gundi=ko=taŋ,*   *Gunai raʣa,*
       uh   PROX   Gundi=GEN=ATT   Gunay king

       ʌ,   *pari=ko,*   *Mime raʣa.*
       uh   opp.deic.side=GEN   Mime king

       'Uh, the one of this Gundi village, the king Gunay, uh, the one of the other side, the king Mime.'
       CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

(509) *Sarling*   *u=tang*   *pʌr=na,*
       Sarling   REM=ALL   have.to_fall=NPST

       niŋ-ʣi   *miskʰa*   *al=o=bʰʌnda*   *pari-pʌti.*
       2DU   just.now   go=NMZ:REL=CMPR   opp.deic-side

       'Sarling is over there, on the opposite side of where you two just went.'
       CH_CTW_BBC_POL_102520_1_Polkim

(510) *ʌntari-kantɕʰi*   *ane*   *i*   *pari*   *Gaibang=ko*
       eightth.F-younger.F   so   PROX   opp.deic.side   Gaibang=GEN

har-haɾi   *mu=na.*

line   COP=NPST

'My eighth daughter then, she lives on the line of the Gaibang village, on the other side of her.'

CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana_Archive

(511)   *ʌ*   *i*   *paɾi=ma*   *tɕiŋ-lan=ko*   *kim=le,*

uh   PROX   opp.deic.side=ADD   Cing-spirit=GEN   house=DIS

'Uh, on the other side of this too there are habitations of the Cing spirit.'

CH_CTW_BBC_POL_111720_3_Cing_Lan

(512)   *ʌ,*   *i*   *Syamrang=ko*   *Ganamani*   *radʑa*   *ɾʌ*

uh   PROX   Syamrang=GEN   Ganamani   king   and

paɾi,   *Sarling=ko=tɕʌhe,*   *Dagu*   *radʑa,*

opp.deic.side   Sarling=GEN=DIS   Dagu   king,

'Uh, this King Ganamani of Syamrang and on the other side, the king Dagu of Sarling,'

CH_CTW_BBC_POL_102520_1_Polkim

(513)   *o=tʌ*   *didi=kʰe=taŋ*   *mʰʌɾ=a*   *wari*   *ɾʌ*

DIST=NMZ:ADV1   elder.sister=DIS=ATT   think=PST   deic_side   and

paɾi=haŋ   *bibahʌ*   *pʌɾ=o,*

OPP.DEIC.SIDE=LOC1   wedding   have.to_fall=PERF

'Like that the elder sister thought, here and on the other side there is a wedding,'

CH_CTW_BBC_POL_111720_5_Two Sisters

3.4.6.3.3.   Postpositions: *matʰi* 'above, up' & *muni* 'under, below'

Although such locational semantics is natively carried by relator nouns, the Nepali postpositions *matʰi* 'above, up' and *muni* 'under, below' are also used in the

language, and in particular in the varieties that have been more greatly influenced by Nepali, such as MAN-4 of Handikhola. They are illustrated in (514) and (515).

(514) *wa*   *dʰuɾi*   *matʰi*   *mu=na.*
      bird   axle   above   COP=NPST

     'The bird is up the axle.'

     CH_MKW_1_21-22_CPR_BAN_100817_1_E

(515) *ŋa*   *siŋ*   *matʰi*   *hampʰal=ti*   *glʌj*   *juk*   *lekʰa*   *hampʰal=ti.*
      1SG   tree   above   jump=SEQ1   langur   rhesus   like   jump=SEQ1

     'I (would) jump up the tree like the langur and rhesus monkeys would do.'

     CH_MKW_SBC_BGR_101719_1_Life

3.4.6.3.4.   Postpositions: *dekʰi* 'from, since' & *sʌmmʌ* 'up to, until'

     The postpositions *dekʰi~dekʰin* 'from, since' and *sʌmmʌ~sʌmmʌn* 'up to, until' are used as postposition to express location in space and time, as shown in (516) to (518). These morphemes may as well function as genitive nominal modifiers, as in (519) and (520).

(516) *juin-ɾaj=ko*   *lagi,*   *ʌni*   *ʥuhɾ=sa=ko*   *lagi=paj,*
      story=GEN   for   and   tell.riddle=NMZ1=GEN   for=DIS

     himal   *dekʰin*   *Tarai,*   *Tarai*   *dekʰin*   *pahad,*
     mountain   from   Tarai   Tarai   from   hill

     dah=to   *ʨiʔ=na=ŋ*   *ŋa=i.*
     reach=NMZ:ADV2   know=NPST=1   1SG=ERG

     'For the stories, and to tell riddles, I have knowledge reaching the plain of Tarai from the mountains, and the hills from the plain of Tarai.'

     CH_MKW_SC_SIL_010220_3_Life

(517) s^j   rupja   le?=to=n=i   t^,   Bastipur
       hundred   rupee   take_buy=REM.PST=DIR/TR=PL   PART   Bastipur

s^mm^=kaj,
up.to=DAT

'The thing is that they used to take hundred rupees (to go up to) Bastipur.'

CH_MKW_SLP_MMRBP_AJI_102519_1_Conversation

(518) ^,   ŋa   ^tʰara   dekʰin unnais   b^rs^   s^mm^n,
       uh   1SG   eighteen from   nineteen   year   until

ŋa   i   lju?mpuk=haŋ,   t^r^   kim   mu=na,
1SG   PROX   cave=LOC1   but   house   COP=NPST

'Uh, from the age of eighteen to nineteen, I (lived) in this cave, but I have a house
(too, at that time).'

CH_MKW_LC_SIL_113019_1_Cave

(519) i   kur̥a   pr^manik,   p^hila dekʰi=ko   pur̥kʰa=l^m=i
       PROX   thing genuine   before from=GEN   ancestor=PL=ERG

lek=ti   waŋ=o   kur̥o,   i   kur̥o=tɕ^j,
tell_relate=SEQ1   come=NMZ:REL   thing   PROX   thing=DIS

ŋi   manta   b^j=sa   p^r̥=na.
1PL   person   give=NMZ1   have.to_fall=NPST

'This type of proof is the thing that we know having been related by the ancestors
of all times, and for this thing we have to provide people.'

CH_CTW_RLC_JIM_101920_Language and Culture

(520) ama-baba=ko   ŋa=i   tɕi?=tɕjuk   s^mm^n=ko   kur̥a
       mother-father=GEN   1SG=ERG   know=QTY   up.to=GEN   thing

'(This is) up to as much I know about the things of my parents.'

CH_MKW_SRNDC_SIL_081818_Life

3.4.6.3.5. Relator nouns: *madʑʰʌ* 'middle of, amongst' & *bitɕ~bitɕtɕʌ* 'middle of'

The morphemes *madʑʰʌ* 'middle of, amongst' and *bitɕ~bitɕtɕʌ* 'middle of' are relator nouns. They occur as nouns modified by a genitive construction, as in (521) and (522), or as modifiers marked with the genitive, as in (523) and (524). Adverbial uses of *madʑʰʌ* 'middle of, amongst' and *bitɕ~bitɕtɕʌ* 'middle of' are also attested, as in (525) and (526).

(521) *ŋa=i*   *git,*   *hʌdʑuɾ=lʌm=ko*   *madʑʰʌ=haŋ,*
     1SG=ERG   song   2PL.HON=PL=GEN   middle=LOC1

       Ram Krishna Dhakal   *me=o*   *git,*
       Ram Krishna Dhakal   song=NMZ:REL   song

       ŋa=ko   *Chepang=ljam=sʌj,*
       1SG=GEN   Chepang=DIR=ABL

       ʌ   *git*   *me=na=ŋ,*   *lʌw !*
       uh,   git   sing=NPST=1   PART

       'I('m going to sing) a song, amongst you all, a song (originally) sang by Ram Krishna Dhakal, in my Chepang language, uh, I'm going to sing a song, all right!'
       CH_MKW_SC_BGR_101619_4_Songs

(522) *o*   *ɾo=ko*   *bitɕ=haŋ*   *sjaʔn,*
     DIST   flower=GEN   middle=LOC1 insect

       'In the middle of this flower, an insect,'
       CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(523) *dada-kʰada*   *u=ko*   *madʑʰʌ=ko*   *mu=na*   *nitʌ,*
     hill-ridge   CAT=GEN   middle=GEN   COP=NPST   PART

       '(The school of Dhirang) is in the middle of that hill ridge, as you can see.'
       CH_MKW_STC_SIL_120619_2_E_4

(524) ʌ=ŋ=sʌj,   o=kaj,   rama=i,   tin   tɕokta,
DIST=LOC1=ABL   DIST=DAT   sickle=INST   three   piece

talaŋ,   ʌni   i=haŋ,   bitɕ=ko   bʰak,
head   and   PROX=LOC1   middle=GEN   part

ʌni   pʌtɕʰaɖi=ko   bʰak,   tɕokta,   tɕokta,   dʑahŋ=a.
and   behind_after=GEN   part   piece   piece   do_make=PST

'From there, we make pieces, one after another, we (cut) it (the bat) with the sickle in three pieces, the head, and here, the middle part, and the back part.'
CH_MKW_MRNDC_SIL_081818_3_Bats

(525) o=ŋ=sʌj   sat   samundrʌ   tʌr=o   bela=haŋ,
DIST=LOC1=ABL   seven ocean   cross=NMZ:REL   moment=LOC1

abʌ,   kwi   ʌgʰi,   beraw madʑʰ,   ʌ,   juʔ   pʌtɕʰaɖi   hʌ.
now   dog   front cat   middle   uh   mouse behind_after COP

'Then, at the time of crossing the seven oceans, so, the dog is in the front, the cat in the middle, and uh, the mouse in the back.'
CH_CTW_BBC_POL_111720_6_Jogi

(526) sali   bitɕtɕʌ=haŋ=ma   dah=taŋ=a.
wife's.younger.sister   middle=LOC1=ADD   reach=ATT=PST

'The wife's younger sister reached the middle.'
CH_CTW_BBC_GUN_102620_1_Cing_Lan

3.4.6.3.6. Relator noun to postposition: *mʌddʰe* 'amongst'

The morpheme *mʌddʰe* 'amongst' functions as a relator noun, as in (527), and as a postposition, as in (528) and (529). As a relator noun, it is followed by the locational case marker *=haŋ*, as in (527), which can also occur when used as a postposition, as in (528).

(527) Chepang=lʌm=ko   mʌddʰe=haŋ=ma,   ekdam   pʰʌrʌk
Chepang=PL=GEN   amongst=LOC1=ADD very   different

bʰasa	*mu=na.*
language	COP=NPST

'Amongst the Chepangs also, there are very different languages.'

CH_MKW_BBC_SIL_032820_2_Pambung

(528)
ʌm,	*tɕʰʌ=ota*	*wa*	*tɕah=ʌ=na*		*dahj=na=i,*	*pande=lʌm,*
yes	six=CL1	hen	need=LN=NPST		say=NPST=PL	shaman=PL

tɕʰʌ=ota	*wa*	*mʌddʰe=haŋ*	*tin=ota*	*bʰalja,*
six=CL1	hen	amongst=LOC1	three=CL1	male

tin=ota	*potʰi.*
three=CL1	female

'Yes, the shamans say that six hens are necessary, amongst the six hens, three male and three female.'

CH_MKW_BBC_SIL_032920_3_Origin_Christianity

(529)
Adibasi	*Janajati*	*Uthan*	*Rastriya*	*Pratisthan=ko*
Indegenous	Nationalities	Development	National	Foundation=GEN

unansatʰi	*u,*	*dzʌndzati*	*mʌddʰe*	*Chepang*	*jadzjo*	*hʌ.*
fifty-nine	CAT	group	amongst	Chepang	one	COP

'Amongst uh, the fifty-nine indigenous groups of the National Foundation for Development of Indigenous Nationalities, Chepang is one.'

CH_MKW_BBC_SIL_032920_1_Chepang people

3.4.6.3.7.   Postpositions: *ʌnusar* 'according to,' *bahek* 'except,' & *bina* 'without'

The postpositions *ʌnusar* 'according to,' *bahek* 'except' and *bina* 'without' are borrowed from Nepali. They are illustrated in (530) to (536). They all primarily behave like postpositions, following nouns. By contrast with *bahek* 'except' and *bina* 'without,' the postposition *ʌnusar* 'according to,' can be followed by the locational case marker =*haŋ*, as in (531), and be marked with the genitive to modify the noun that precedes it, as

in (532). Finally, the posposition *bahek* may as well follow a noun marked with the ablative case marker *=sʌj* which functions as a compartive, as in (534).

(530)   *dʰerʌj*   *ɾadʑa*   *sjaw=ka=i*   *ma=ba,*
          a.lot   king   become=2/3.PST=PL   PART=PART

          tol   *ʌnusaɾ*   *ɾadʑa*   *sjaw=ka=i ?*
          division   according.to   king   become=2/3.PST=PL

          'There were a lot of kings it sounds like, were there kings according to the different areas?'
          CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

(531)   *o*   *ʌnusaɾ=haŋ*   *deŋe*   *biswas*   *dʑahŋ=na=ŋ=su.*
          DIST   according.to=LOC1   now   Christianity   do_make=NPST=1=1PL.EXCL

          'According to that, we now believe in Christianity.'
          CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(532)   *pʌhile*   *ŋi=ko*   *mʌwlik*   *pʌɾʌm-pʌɾa*   *ʌnusaɾ=ko*
          before   1PL=GEN   original   traditions   according.to=GEN

          ŋi=ko   *biʃwaʃ*   *mʌ=na.*
          1PL=GEN   belief   COP=NPST

          'Before, we had beliefs according to our original traditions.'
          CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(533)   *ŋa*   *bahek*   *sʌppej*   *al=na=i.*
          1SG   except   all   go=NPST=PL

          'Except me, they all go.'
          CH_MKW_PC_SIL_E

(534)   *ʌni*   *ŋi=ko*   *bajbʌl=ko*   *kʌtʰa=haŋ=ma*
          so   1PL=GEN   Bible=GEN   story=LOC1=ADD

to=o	*mʌ=na*	*mʌ=sa=kaj=paj,*	
tell_say=PERF	COP=NPST	COP=NMZ1=DAT=DIS	

adam	*rʌ*	*habba=sʌj*	*bahek*	*na=lʌ.*
Adam	and	Eve=ABL	apart	COP=NEG

'So, in the story of our Bible also, as for what is said, apart from Adam and Eve, there was nothing.'

CH_CTW_RC_KCR_101920_2_Myth_Origin

(535)
lʌw,	*sjaw=na,*	*ŋi-tɕi*	*naŋ*	*bina,*	*ŋi-tɕi*
well	become=NPST	1DU	2SG	without	2DU

pahj=ŋʌ=tɕʌ=lʌ,	
leave=1=1/3DU=NEG	

'Well, that works, we two, without you, we two won't leave.'

CH_CTW_BBC_POL_111720_6_Jogi

(536)
sabun	*bina*	*dʰaw=gar=ŋʌ=lʌ.*
soap	without	wash=DES=1=NEG

'I don't want to wash without soap.'

CH_MKW_SC_SIL_120619_4_E_2

3.4.6.3.8. Relator nouns: *lagi* 'for' & *nimti* 'for, for the sake of'

The relator nouns *lagi* 'for' and *nimti* 'for, for the sake of,' are borrowed fom Nepali. They occur as a nominal head modified by a genitive construction and may or may not be followed by the locational case marker *=haŋ*, as illustrated in (537) to (539).

(537)
Ross	*Caughley*	*sajep,*	*pʌtɕis*	*sal=haŋ,*	*ʌni*
Ross	Caughley	sir	2025	year=LOC1	so

bʰasa=ko	*lagi*	*waŋ=a.*
language=GEN	for	come=PST

b^hasa=ko	lagi=haŋ	waŋ=a	$p\Lambda t\varepsilon^hi$, (…)
language=GEN	for=LOC1	come=NMZ2	after

'Sir Ross Caughley, in the year 2025 VS (1968 CE), so he came for the language.

After he came for the language, (…)'

CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(538) | *Chepang=ko* | $\Lambda d^hika\Gamma$=ko | nimti=tɕ\Lambdahe, |
|---|---|---|
| Chepang=GEN | right=GEN | for=DIS |

'For the sake of the Chepangs' rights, (…)'

CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(539) | *tɕʰjo,* | *mu=\Lambda,* | *naŋ* | *l\Lambdaj=ko* |
|---|---|---|---|
| enough | COP=2SG.IMP.INTR | 2SG | SLF.INTS=GEN |

eɾija=ko	*nimti=haŋ,*
area=GEN	for=LOC1

'Enough, you stay, for the sake of your area,'

CH_CTW_BBC_POL_102420_3_Chepang_Kings

CHAPTER IV

PRONOUNS AND DEMONSTRATIVES, DEIXIS AND DISCOURSE ENDOPHORA

This chapter describes pronouns, along with their deictic and discourse uses. Pronouns share with nouns most of their morphosyntactic devices. They constitute a specific part of speech category since some of their morphosyntactic behaviors are distinct from that of nouns: pronouns can substitute for a noun-phrase or higher syntactic unit, like a text; they have discourse anaphoric and cataphoric referential functions; and they cannot be modified by an adjective or relative clause.

Personal independent pronouns are presented in § 4.1, the nominal exclusive morpheme *=jal~jal* 'only' is described in 4.2, other types of demonstrative pronouns in § 4.3, possessive pronouns in § 4.4, and interrogative pronouns in § 4.5.

4.1.   Personal independent pronouns

The personal independent pronouns, or personal pronouns, are presented in their absolutive form in Table 124. All inflect in number (singular, dual, plural).

Clusivity in personal pronouns is not attested with dual and plural, as can be the case in other TH languages (§ (DeLancey 2019). I describe the uses of the pronouns in the following sub-sections. Details about 1st person forms are given in § 4.1.1, and about 2nd person in § 4.1.2. The 3rd person pronouns *i* and *o* distinguish human vs. non-human in their plural and dual forms. Deixis plays a major role in the choice between the proximal and distal forms *i* and *o,* in addition to discourse referential and contrastive functions. A certain degree of politeness is associated with 3rd person pronouns, mainly linked with their deictic functions. The form *u* is specifically used to express remote or absent referents. Two additional 3rd person pronouns borrowed from Nepali are *u* and *uh.* The former has a discourse cataphoric function, and the latter a deictic locational function that can be used to refer to anything with a mirative or attentional effect triggered on the addressee. 3rd person pronouns are described in § 4.1.3.

These personal independent pronouns are widely used in all varieties. The allomorphs attested for 1st and 3rd person are not restricted to specific varieties.

In addition to these pronouns, two innovations are attested for 2nd and 3rd person (distal). They are presented in Table 125.

Table 124. Personal independent pronouns

	SG	DU		PL	
1	*ŋa*	*ŋi~ni-tɕi*		*ŋi~ni*	
2	*naŋ*	*niŋ-dʑi*		*niŋ*	
3		**HUM**	**N.HUM**	**HUM**	**N.HUM**
PROX	*i*	*i=nis=tɕak* *i=nis*	*i*	*i-mi~mʌj=lʌm* *i-mi~mʌj* *i=lʌm*	*i*
DIST	*o*	*o=nis=tɕak* *o=nis*	*o*	*o-mi~mʌj =lʌm* *o-mi~mʌj* *o=lʌm*	*o*
REM	*u*	*u=nis=tɕak* *u=nis*	*u*	*u-mi~mʌj=lʌm* *u=lʌm*	*u*
1/2HON, IMPS	*lʌj*				
2HON	*hʌdʑuɾ* (<N.) *uha* (<N.)			*hʌdʑuɾ=lʌm* (<N.)	

457

Table 125. Innovations attested for independent pronouns

		SG	DU		PL	
RAP-13/11	**2HON**	*tʌpaj* (<N.)	*tʌpaj=nis* (<N.)		*tʌpaj=lʌm* (<N.)	
			HUM	**N.HUM**	**HUM**	**N.HUM**
MAN-4	**3**	*ow*	*ow=nis=tɕak*	*ow*	*ow-mi~mʌj=lʌm*	*ow*
RAP-13	**DIST**		*ow=nis*		*ow-mi~mʌj*	
					ow-lʌm	

4.1.1. 1ˢᵗ person pronoun

The 1ˢᵗ person singular pronoun in Chepang is *ŋa*, a widespread form in TH languages. Velar initial pronouns are reconstructed back to PTH for 1ˢᵗ person singular: *#ŋa* by Benedict (1972), Thurgood (1985), Matisoff (2003), and DeLancey (2019).

For the 1ˢᵗ person dual and plural forms, two allomorphs are used in free variation in all varieties: *ŋi~ni-tɕi* 1DU and *ŋi~ni* 1PL. Both forms *ŋi* and *ni* can be used interchangeably by a single individual. However, I have observed that in MAN-4, the form *ni* occurs more predominantly.

The form *ŋi* is likely the oldest, result of the combination of a velar nasal /ŋ/, like the initial velar nasal that forms the 1ˢᵗ person singular pronoun, with the vowel /i/ which may originally be a plural marker, like the plural morpheme *=i* found in verbal inflectional morphology. It is possible that the velar nasal /ŋ/ palatalized in front of /i/ and developed the allomorph *ni*.

The 1ˢᵗ person dual *#ŋa-tsi* is reconstructed by DeLancey (2019). The 1ˢᵗ person plural *#i* is reconstructed by Bauman (1975), van Driem (1993b) and DeLancey (2019). Another 1ˢᵗ person plural form *#ka* functioning as an exclusive marker is reconstructed by Benedict (1972), Bauman (1975), and van Driem (1993b).

4.1.2. 2nd person and innovative formal pronoun

The 2nd person singular pronoun is *naŋ* is another widespread form in TH languages. The 2nd person dual is *niŋ-dʑi*, and 2nd person plural *niŋ*.

The 2nd person dual form -*dʑi* could be analyzed as cognate with the form -*tɕi* found with 1st person dual, result of voicing assimilation. But this development is unlikely since this distinction is also present in the verbal morphology, which is unlikely to be innovative. Benedict (1972) reconstructs *#naŋ* for 2nd person singular and *#niŋ* for 2nd person plural. DeLancey (2019) reconstructs *#naŋ* for 2nd person singular *#naŋ-tsi* for 2nd person dual and *#ni* for 2nd person plural.

The form *naŋ* is used to address children, people with whom the speaker feels close or comfortable with, but also elderly people. The use of *naŋ* with elderly people does not show disrespect, rather some kind of tenderness or affection, similar to that expressed to children.

The use of *naŋ* is avoided with people that are not personally known by the speaker, or that the speaker just met. In this case, either the speaker completely avoids using any term of address, relying solely on argument indexation present on the verb or the use of nominalized constructions, or the innovative honorific pronoun *lʌj* or borrowed honorific pronouns from Nepali are used instead, such as *hʌdʑuɾ* and *uha*. In addition to these strategies, in the varieties spoken in RAP-11 and RAP-13, speakers have developed the use of *tʌpaj* as a formal 2nd person pronoun, borrowed from the Nepali 2nd person honorific pronoun तपाइ <tapāi>. In Nepali, तपाइ <tapāi> combines also with number. In Chepang too, the formal 2nd person *tʌpaj* is attached with the native dual morpheme =*nis* and plural morpheme =*lʌm*. In the case of *tʌpaj*, the 2nd person dual marker -*dʑi* found with *niŋ* is not used.

4.1.3. 3rd person pronoun and demonstrative

The referent of the 3rd person pronouns or demonstratives *i* 'proximal' and *o* 'distal' can be a person, an object, an animal, an event or fact, or a place. They are illustrated in (540) and (541) with a person, in (542) and (543) with a natural object, and

in (544) and (545) with an animal. These pronouns also function as demonstrative determiners, as illustrated in (546) to (547).

There is no morphological difference between 3rd person pronouns and demonstrative determiners, nor is there a difference between 3rd person pronouns and demonstrative pronouns. The 3rd person singular and dual pronouns likely developed from demonstrative pronouns. This directionality is typologically widespread (Diessel 1999; Klausenburger 2000; Heine & Song 2011). The plural forms have developed differently in a construction involving the demonstrative determiners rather than the demonstrative pronouns.

(540) *i*   *ŋa=ko*   *majli-tɕoʔ=le.*
       PROX   1SG=GEN   second.eldest.F-child=DIS

       'She/This is my second daughter.'
       CH_MKW_SC_SIL_120619_3_E

(541) *o=i*   *ŋa=kaj*   *mʌdʌt*   *bʌj=o*   *mʌ=na*   *hʌj !*
       DIST=ERG   1SG=DAT   help   give=PERF   COP=NPST   PART

       'S/He has helped me a lot, for sure!'
       CH_MKW_SC_SIL_120619_2_E_1

(542) *i*   *gal=o*   *badʌl*   *ma.*
       PROX   be.black=NMZ:REL   cloud   COP

       'It/This is a black cloud.'
       CH_MKW_1_40-47_CPR_BAN_101217_1_E

(543) *ɾuiŋ*   *tat=a,*   *tat=bʌtiko,*   *o=kaj*   *tɕiɾ=ʌ=ti,*
       bamboo   cut=PST   cut=SEQ2   DIST=DAT   slit=LN=SEQ1

       'The bamboo is cut, after it's cut, one/people slits it,'
       CH_MKW_MRNDC_SIL_081818_6_Dikalak

(544) *ma,*   *i*   *mʌj=o*   *wa*   *ma,*
       yes   PROX   be.small=NMZ:REL   hen   COP

       mʌj=o   *kʰal=ko*   *wa.*
       be.small=NMZ:REL   type_kind=GEN   hen

       'Yes, it/this is a small hen, a hen of a small kind.'

       CH_MKW_1_40-47_CPR_BAN_101217_1_E

(545) *tɕum=sa*   *kʰaj=lʌ*   *o=kaj,*   *tɕum=sa*   *kʰaj=lʌ.*
       catch=NMZ1   be.able=NEG   DIST=DAT   catch=NMZ1   be.able=NEG

       'One/people can't catch it (porcupine), one/people can't catch (it).'

       CH_MKW_PMRC_LAM_081618_4_Porcupine

(546) *pʌhila,*   *i*   *kim=haŋ*   *mu=o*   *kʰe=to,*   *ŋa.*
       before   PROX   house=LOC1   COP=PERF   COP=REM.PST   1SG

       'Before, I had lived in this house.'

       CH_MKW_KKBP_LPK_101917_1_Life

(547) *ga=haŋ*   *al=a,*   *o*   *ɾadza ?*
       INT=LOC1   go=PST   DIST   king

       'Where did he go that king?'

       CH_CTW_SBC_BBC_GUN_012120_Chepang Kings

     The pronouns *i* and *o* inflect for number: singular, dual, and plural. Only human referents show dual and plural marking. With non-human referents (animal and objects), unless anthropomorphologized, only the pronouns *i* and *o* occur, with no distinction of dual or plural number, as shown in (548) with hens and in (549) with children.

(548) *i=kaj*   *kʌn=u*   *da !*
       PROX=DAT   look=2SG.IMP.TR   PART

       'Keep an eye on them, okay!'

       CH_CTW_SPC_POL_E

(549) i=lʌm=kaj   kʌn=u   da !
       PROX=PL=DAT   look=2SG.IMP.TR   PART

'Keep an eye on them, okay!'
CH_CTW_SPC_POL_E

The dual function can be marked with two morphemes, =nis=tɕak, or simply =nis, which is a morpheme also found attached to nouns. The form =nis means 'two' and shows four allomorphs in free variation when followed by the morpheme =tɕak, such as: =nis~nih~nitVd~ni. The final voiceless fricative consonant of =nis can thus entail glottalization (=nih), fortition (=nitVd), or deletion (=ni). The same allomorphy is attested when =nis is followed by the form =dʑjo~dʑjaŋ with which they together form a determiner to refer to objects, or by the form =dʑana borrowed from Nepali and used to determine nouns that refer to people. I did not include this allomorphy in Table 124 since their surface realization is somewhat unpredictable, mainly tight to speech rate. The morpheme =tɕak is likely cognate with the nominal suffix -tɕa that derives pairs of individuals of the same kind (§ 3.3.4.2). The absence of =tɕak seems to affect politeness, showing a lower degree of address.

The morpheme used with human 3rd person plural referent has two allomorphs, also attested in free variation: -mi~mʌj. This variation occurs in all varieties and can sometimes vary in the speech of a single individual. The morpheme -mi~mʌj can combine with the plural form =lʌm which occurs with pronouns referring to humans. According to some speakers, the presence of -mi~mʌj conveys more respect than its absence. According to others, there is no difference. The question of degrees of politeness to refer to 3rd person is described in § 4.1.3.2.

The morphemes i and o were likely used as demonstrative determiners, such as *i/o mi 'this/that person' or *i/o mi=lʌm 'these/those persons' before -mi~mʌj encliticized to the demonstrative pronouns i and o, a noun-phrase construction then reanalyzed as a 3rd person plural pronoun.

Finally, in the language varieties spoken in MAN-4, the distal 3rd person morpheme o underwent labialization to became ow. This form is also attested in RAP-13 in free variation with o.

462

The morpheme *i* can be described as marking a so-called proximal reference and the morpheme *o*, a distal reference. However, this proximal vs. distal distinction is mainly present deictically or in discourse when two objects (or else) or persons are to be expressed comparatively, as shown in (550) and (551).

(550) *naŋ=ko*   *ga=o=kʰe*   *mobajl ?*   *i*   *ja,*   *o*   *ja ?*
       2SG=GEN   INT=NMZ:REL=DIS   phone   PROX or   DIST or

       'Which one is your mobile? this one or that one?'
       CH_CTW_SP_E

(551) *ane,*   *i*   *Silinge*   *dada*   *rʌ*   *o=taŋ*
       so   PROX Silinge   hill   and DIST=ALL

       Sidam pʌti   *ja=dʑo=le*   *noʔ=sa=paj.*
       Sidam side   one=CL1=DIS speak=NMZ1=DIS

       'So, this Silinge hill, and over there, on the side of Sidam, speaking is similar.'
       CH_MKW_RC_JMC_SIL_120119_Conversation

We will see that the functions of *i* and *o* cannot be reduced to a difference in spatial position. In fact, the distinction between *i* and *o* is primarily deictic, but also contrastive (§ 4.1.3.1). Degrees of politeness regarding 3rd person are linked with deixis (§4.1.3.2). In addition to *i* and *o*, the remote pronoun *u* has deictic and discourse cataphoric functions. Two other forms that should not be confused with the remote pronoun *u* are borrowed from Nepali: *u* and *uh*. They serve pronominal functions as well. The form *u* is used as a cataphoric device to hold the floor of discourse while remembering what was to be said. The form *uh* is a pronoun that can refer to both a person and a location with an additional mirative or attentional grasp effect sought to be triggered in the addressee. The three morphemes *u*, *u*, and *uh* are described in § 4.1.3.3, § 4.1.3.4, and § 4.1.3.5, respectively.

4.1.3.1. Deictic and discourse anaphoric uses of 3ʳᵈ person *i* and *o*

The 3ʳᵈ person proximal pronoun *i* is illustrated in (552) with an object referent and in (553) with a person referent, both through the expression of a deictic reference and a discourse anaphoric reference. In other words, the referent of *i* in both examples is deictically accessible to the speaker and the addressee. While its first mention is clearly a deictic reference, the second mention of the referent does not necessarily constitute an anaphoric reference even though the referent was just mentioned in discourse, since it is still present deictically, i.e., the deictic reference is still active.

In (552), the speaker is describing the various products that the Indian butter tree offers, and in this excerpt, he is talking about the oil on display in front of him. The first occurrence of *i* introduces the oil to the addressee while deictically pointing at it. The second occurrence of *i* refers anaphorically to the oil previously introduced with *i*.

In (553), the speaker explains to the addressee why the person next to them does not remember what her native village is like.

In both examples, all the instances of the 3ʳᵈ person proximal pronoun *i* express referents accessible deictically.

(552) *i=kʰe,*   *jʌ-sati.*
       PROX=DIS   Indian.butter.tree-oil

       i   *kjan=haŋ*   *kaɾ=ti*   *dʑe=sa*   *njum=na.*
       PROX   dish=LOC1   fry=SEQ1   eat=NMZ1   be.tasty=NPST

       'As for this, it is Indian butter tree oil. Eating this having fried it in a dish is tasty!'
       CH_MKW_MRNDC_SIL_081818_7_Indian Butter Tree

(553) *dʌgi=paj*   *i*   *mʰea=ka=n !*
       nowadays=DIS   PROX   forget=2/3.PST=DIR/TR

       i=paj   *i=tɕuk*   *mi=to*   *kʰe=to !*
       PROX=DIS   PROX=QTY   small=NMZ:ADV2   COP=REM.PST

       'Nowadays, she forgot! She was this little!'

464

The 3rd person distal pronoun *o* is illustrated in (555), in a conversation between a daughter and her parents who try to convince her to marry a boy she doesn't like. This is the beginning of the story. The boy is introduced in (554) with the distal demonstrative determiner *o*. A couple of sentences later, after the girl told her parents that she did not like the boy, using the distal pronoun *o*, the parents insist, telling her she should like him and marry him anyway (555). The fact that the boy is not present deictically in this conversation clearly triggers the use of *o*, by contrast with the above instances of *i* in (552) and in (553).

(554) ʌ,   *baba*   *rʌ*   *ama=i*   ʌ,   *keti=kaj,*
       uh   father and   mother=ERG   uh   girl=DAT

       o   *dʑʰak=ma=l=o*   *buda=kaj*   *bʌj=ka=n=i.*
       DIST   like=NEG=COP=NMZ:REL   husband=DAT   give=2/3.PST=DIR/TR=PL

       'Uh, the parents gave the girl to that husband she didn't like.'
       CH_MKW_SC_SIL_122619_2_Dhobini rani

(555) *"ʌhʌ,*   *naŋ*   *pahj=sa=le*   *pʌr=na*
       no   2SG   leave_go.home=NMZ1=DIS   have.to_fall=NPST

       o=kaj=le,   *naŋ*   *dʑʰak=sa*   *pʌr=na,*
       DIST=DAT=DIS   2SG   like=NMZ1   have.to_fall=NPST

       pahj=ʌ   *o=kaj=le."*
       leave_go.home=2SG.IMP.INTR   DIST=DAT=DIS

       'No, you have to marry him, you have to like him, marry him!'
       CH_MKW_SC_SIL_122619_2_Dhobini rani

Regardless of the presence of an earlier mention of the referent, deixis triggers the choice of using the proximal pronoun *i* or distal pronoun *o*. That is, deixis prevails over discourse, and both *i* and *o* can be used as anaphoric pronouns. This is confirmed in (556), from a story where a girl likes a boy, but the girl's parents do not want her to

marry him. The narrator uses the pronoun *o* to refer to the parents of the girl, to introduce the boy the girl likes, and to refer to the girl when the boy comes visit in a conversation where the parents, the girl and the boy are all present. The parents tell the boy that they do not want to give him the girl to marry, referring to the girl with the proximal pronoun *i*, before the narrator continues to describe the story using the demonstrative determiner *o*.

(556)   *o=ko*   *tɕoʔ-djaŋ*   *jadʑo*   *mʌtɾej=taŋ*   *mu=to*   *na.*
         DIST=GEN   daughter   one   only=ATT   COP=REM.PST   COP

      ʌni   *o=ko*   *tɕoʔ-djaŋ,*   *ekdʌm*   *sundʌɾi,*
      so   DIST=GEN   daughter   very   beautiful

      pe=to   *kja !*
      be.nice=NMZ:ADV2   PART

      ekdʌm,   *maha*   *pe=to=taŋ*   *kʰe=to=taŋ*   *na.*
      very   very   be.nice=NMZ:ADV2=ATT   COP=REM.PST=ATT   COP

      ʌni   *o*   *keta=taŋ*   *ekdʌm*   *dʑʰak=a,*
      so   DIST   boy=ATT   very   like=PST

      o   *keta=ma=taŋ*   *kʰandan=ko*   *tɕoʔ*   *le !*
      DIST   boy=ADD=ATT   rich.descent=GEN   child   COP

      ʌni   *jat*   *din*   *waŋ=ja*   *pʌtɕʰi,*   *o*   *keti=kaj*
      so   one   day   come=NMZ2   after   DIST   girl=DAT

      kʰandan=ko   *tɕoʔ,*   *"i=paj*   *aʔl=sa*   *sjaw=lʌ,*
      rich.descent=GEN   child   PROX=DIS   take.away=NMZ1   become=NEG

      bʌj=sa   *sjaw=lʌ,*   *i=kaj=paj"*   *dahj=ti,*
      give=NMZ1   beome=NEG   PROX=DAT=DIS   say=SEQ1

      o   *keti=ko*   *ama-ba=i,*   *tɕoʔ-djaŋ=ko*   *ama-ba=i,*
      DIST   girl=GEN   mother-father=ERG   daughter=GEN   mother-father=ERG

bʌj=ma=ja		*pʌtɕʰi,*	*ʌni*	*o*	*keta=kaj*
give=NEG=NMZ2		after	so	DIST	boy=DAT

upʌja=taŋ	*waŋ=a.*
idea=ATT	come=PST

'They had only one daughter. And their daughter was so beautiful! So nice hey! She was extremely beautiful. And she very much liked that boy. That boy was even from a rich family! So when he came one day, to the girl, that rich family boy: "You can't marry her, we can't give her to you!" said that girl's parents, the daughter's parents; since they would not give her (to him), then that boy got an idea.'

CH_CTW_RC_KCR_101920_2_Sja'n

But the deictic distinction between *i* and *o* is not necessarily that straightforward. While the morpheme *i* is used to refer to proximal referent deictically, it also underlines some contrast between the entity that it refers to and any potential others.

In (557), the speaker reports what he was told by someone giving him the Bible for the first time, when the first Chepangs started to be converted to Christianity. In this person's discourse, the proximal 3rd person pronoun *i* refers to the Bible that is deictically present in their hand. Then, the speaker explains to the addressee that after he (speaker) read it, he realized it was the Bible, that the book talked about Jesus. He uses the 3rd person pronoun *o* to refer to the Bible in the clause *o pʌrʌti* 'having read that,' anaphorically referring to the Bible he just talked about. Indeed, the Bible is not present at the time of this conversation. Then, in the main clause that follows, the speaker uses the 3rd person proximal pronoun *i* to refer to the Bible in a non-verbal predicative construction. He is here clearly putting emphasis on the Bible, that he realized at the time was this particular book, by contrast with any other kind of book. In (558), the speaker reports more about what he was told at the time. The first two instances of *i* also refer to the Bible deictically. But in the last clauses, the 3rd person proximal *i* refers to Jesus. The person who tries to convert the speaker uses *i* and not *o* to refer to Jesus whom he tells the speaker to worship, by contrast with worshiping anything else. He then insists that if

he worships Jesus, nothing will happen to him, that he will not need to see any Shaman as people use to do, that he will not need anything else.

(557) *"naŋ=i   bʌlʌ   bʌlʌ   pʌɾ=ʌ=sa   tɕiʔ=te=na=u,*
     2SG=ERG   little   little   study=LN=NMZ1   know=2=NPST=3O/DIR

i   pʌɾ=ʌ=ʌ,   naŋ."
PROX   study=LN=2SG.IMP.INTR   2SG

o   pʌɾ=ʌ=ti,   ane   Yesu=ko   baɾe=ko
DIST   study=LN=SEQ1   so   Jesus=GEN   subject=GEN

kuɾa   to=o   i.
thing   tell_say=NMZ:REL   PROX

'"You know how to read a little, read this." Having read that, then (I realized that) this is what is told about Jesus.'
CH_MKW_PSC_MAI_012620_1_Becoming Christian

(558) *"i   pʌɾ=ʌ=ʌ,   i   pʌɾ=ʌ=ti   naŋ=i*
     PROX   study=LN=2SG.IMP.INTR   PROX   study=LN=SEQ1   2SG=ERG

i=kaj   man=ʌ=sa   kʰaj=te=na=u.
PROX=DAT   worship=LN=NMZ1   be.able=2=NPST=3O/DIR

i=kaj   man=ʌ=a   pʌtɕʰi
PROX=DAT   worship=LN=NMZ2   after

naŋ   doh=ma   dʌj=te=lʌ,   naŋ=kaj
2SG   what=ADD   happen_occur=2=NEG   2SG=DAT

pande=ma   pʌɾ=ʌ=lʌ,   doh=ma   pʌɾ=ʌ=lʌ,"
shaman=ADD   have.to_fall=LN=NEG   what=ADD   have.to_fall=LN=NEG

'Read this, and having read this, you will be able to worship him (Jesus). After having worshiped him (Jesus), nothing will happen to you, you will not need any shaman, you will not need anything,'
CH_MKW_PSC_MAI_012620_1_Becoming Christian

In (559), the speaker talks about the stories she heard from her father. She tells the addressee what she understood about the way these stories are told, and how powerful they are. She tells the addressee what she realized when listening to these stories, reporting her own thoughts. She uses the 3rd person proximal pronoun *i* to differentiate each story while describing the particular thing she learned from each of them. This differentiation prevails over the deictic character of *i*. Although she reports the thoughts she was having while listening to the stories, in the last clause, she uses the distal demonstrative determiner *o* to refer to the last type of story.

This shows that using *i* has also the function of contrasting referents with one another, similarly to when two referents deictically present are distinguished.

(559) *ŋa=ko*   *baba=i*   *dʑjasta*   *kʌtʰa*   *ta=u=to,*
     1SG=GEN   father=ERG   excessive   story   tell.story=3O/DIR=REM.PST

     dʰeɾe *kʌtʰa* *tɕiʔ=u=to*   *ŋa=ko*   *baba=i.*
     a.lot   story   know=3O/DIR=REM.PST   1SG=GEN   father=ERG

     ŋa=ko   *baba=i*   *ta=o*   *kʌtʰa,*
     1SG=GEN   father=ERG   tell.story=NMZ:REL   story

     jat   *ŋa=kaj*   *sʌbej*   *mu=na.*
     one   1SG=DAT   all   COP=NPST

     i=kaj   *i=tʌ*   *to=sa*   *pʌɾ=na*
     PROX=DAT   prox=NMZ:ADV1   tell_say=NMZ1   have.to_fall=NPST

     ɾejsʌ,   *i=kaj*   *i=tʌ*   *dahj=sa*
     COP.MIR   PROX=DAT   prox=NMZ:ADV1   say=NMZ1

     pʌɾ=na   *ɾʌjsa,*   *i=kaj*   *i=tʌ*
     have.to_fall=NPST   COP.MIR   PROX=DAT   PROX=NMZ:ADV1

     dahj=ti,   *i*   *tʰaw=haŋ*   *dak=laŋ*   *al=sa*
     say=SEQ1   PROX   place=LOC1   escort_lead=PUR   go=NMZ1

     pʌɾ=na   *ɾejsʌ.*
     have.to_fall=NPST   COP.MIR

o	*kʌtʰa*	*i*	*tʰaw=haŋ*	*al=na*	*ɾejsʌ.*
DIST	story	PROX	place=LOC1	go=NPST	COP.MIR

'My father used to tell so many stories, he knew a lot of stories, my father. The stories that my father told, I have one, I have them all. "For this one, it turns out, we must tell it this way! For this one, it turns out, we have to say it like this! for this one, saying it like this, it turns out that you must go lead (people) to this place! That story, it turns out, goes to this place.'
CH_MKW_SC_SIL_010220_2_Life

Another last example in (560) shows that indeed, the 3rd person proximal pronoun *i* has a contrastive effect that can apply whether the referent is present deictically or not. The narrator here reports the thoughts of a boy who became a caterpillar to marry the girl he likes. Having become a caterpillar, he now makes money from the seeds of the Bauhinia Vahlii creeper. The boy is deictically referring to the money in front of him with the proximal 3rd person demonstrative determiner *i* and then uses the 3rd person pronoun *i* to talk about the father of the girl he wants to marry. The father is not present deictically at the time the caterpillar is thinking, but the use of *i* here emphasizes that his plan is to marry the daughter of this one father specifically, by contrast with any other father.

(560)

ʌni	*"lʌw,*		*i*	*pʌjsa*	*bʌj=ti*	
so	well		PROX	money	give=SEQ1	

ŋa=i	*i=ko*	*tɕoʔ-djaŋ*	*aʔl=na=ŋ"*		*dahj=ti*
1SG=ERG	prox=GEN	daughter	take.away=NPST=1		say=SEQ1

o=i=kʰe	*lʌj=ko*	*mʌn=haŋ*	*sotɕ=ʌ=ti,*
DIST=ERG=DIS	SLF.INTS=GEN	heart=LOC1	think=LN=SEQ1

'So, "Well, giving this money, I will take away his daughter," he said, he was thinking in his own heart,'
CH_CTW_RC_KCR_101920_2_Sjaʔn

470

Finally, the contrastive function of the 3rd person proximal pronoun *i* is also present when *i* occurs as a demonstrative determiner, as shown in (561). Here, doing the work of shamans is clearly contrasted with doing the work of Jesus.

(561) *niŋ=i*　　　 *i*　　 *pande=lʌm=ko*　　　 *kam*　 *dzahŋ=ljam.*
　　　2PL=ERG　　PROX　shaman=PL=GEN　　work　do_make=IMP.NEG

　　　niŋ=i　　　 *pɾʌbʰu=ko*　 *kam*　 *dzahŋ=nu.*
　　　2PL=ERG　　lord=GEN　　work　do_make=2PL.IMP.TR

　　　'You all, don't do the work of these shamans, you, do the work of Jesus.'
　　　CH_MKW_PSC_MAI_012620_2_Becoming Christian

4.1.3.2.　　　Degrees of politeness for 3rd person *i* and *o*

According to some speakers, different degrees of politeness apply to the use of the 3rd person pronouns *i* and *o*. According to others, there is no difference. In fact, Chepang mainly relies on verbal morphology for politeness strategies. I will briefly describe the points of views speakers expressed regarding politeness in reference to 3rd person.

Some speakers say that when talking about a single individual with whom a person feels close to or is familiar with, the 3rd person proximal pronoun *i* is favored. This applies as well for dual and plural forms. In addition, they note that using the full forms *i~nis=tɕak* PROX=DU=CL2 and *i-mi~mʌj=lʌm* PROX-PL.H=PL conveys more respect than the shorter forms *i=nis* PROX=DU and *i=lʌm* PROX=PL. They explain that these latter are usually used to talk about children. The plural shorter form *i-mi~mʌj* PROX-PL.H is nevertheless considered to be a form that conveys as much respect as its full form *i-mi~mʌj=lʌm* PROX-PL.H=PL.

The 3rd person distal pronoun *o* is said to be preferred to talk about someone the speaker does not know personally or people that holds a prominent position in society, politics, etc. Similarly to what conveys the use of the full dual and plural forms with the 3rd person proximal pronoun *i*, the use of *o=nis=tɕak* DIST=DU=CL2 and *o-mi~mʌj=lʌm* DIST-PL.H=PL show more respect than their shorter counterparts *o=nis* DIST=DU and

o=lʌm DIST=PL. The short form *o-mi~mʌj* DIST-PL.H conveys a little more respect than the two other shorter dual and plural forms.

Other speakers say that there are no such politeness degrees associated with 3rd person pronoun forms. They note that an impolite way to refer to 3rd person when deictically present at the time of speech would be to use the 3rd person distal pronoun *o* rather than the 3rd person proximal *i*. According to those who do not see politeness differences associated with 3rd person, only an accurate use of the deictic functions of *i* and *o* matters.

Indeed, the examples given about children being referred to as i may merely be triggered by the fact that the children are often present deictically while talking about them.

However, it is nevertheless possible that the use of full forms when it comes to marking dual and plural is seen as more polite for certain Chepang communities. The speakers who described the different degrees of politeness are from rap-13 and those who did not see a difference are from rak-6. These contrastive views could be further explored since it would be interesting to see how these pronominal politeness strategies are used, if they exist in certain varieties, and how they are perceived by speakers of other varieties. In fact, the varieties spoken in RAP-13 are the ones which also innovated the use of the Nepali 2nd person honorific form *tʌpaj* to address more politely 2nd person. In fact, it is likely that politeness distinction was not a function native to Chepang and that contact with Nepali speaking communities clearly influenced its emergence. The only other form that conveys politeness for 1st and 2nd person is the use of the self intensifier morpheme *lʌj*, a calque construction that developed as well under the influence of Nepali.

4.1.3.3. Deictic and discourse anaphoric uses of 3rd person *u*

The 3rd person remote pronoun *u* is used to refer deictically to objects, places, or persons that are situated far away from the location where the conversation takes place. The referent can be visible or not, or merely indicated through a direction. This is illustrated in (562).

(562) *u*   *Bhimbuŋ*   *dada*   *kʰe=na.*
       REM   Bhimbung   hill   COP=NPST

       'This is the hill of Bhimbung.'
       CH_MKW_PC_SIL_E

In discourse, the 3rd person remote pronoun *u* is used to express a remote, non-visible referent, often used to talk about people, places, times, or facts, that do not exist anymore in the world, as in (563), where the speaker talks about local Chepang kings fighting with one another with bows and arrows. The same function is attested when *u* is used as a remote demonstrative determiner, as illustrated in (564), referring to the topic of the story the speaker just narrated, i.e., spirits that now belonged to the story that is finished. In (565), the two instances of the 3rd person remote pronoun and demonstrative determiner *u* refer to a practice which is no longer used in the Chepang community, specifically slash and burn.

(563) *u=lʌm=i*   *ap=ti*   *lʰok=o*   *bela=haŋ*
       REM=PL=ERG   shoot=SEQ1   send=NMZ:REL   moment=LOC1

       ljumpuk=haŋ *pok=bʌt*   *ljuhŋ=ti*   *mu=to.*
       cave=LOC1   enter=SEQ2   waiting=SEQ1   cop=REM.PST

       'At the time when they were shooting and sending (arrows), they entered in a cave and were waiting.'
       CH_CTW_BBC_POL_102520_1_Polkim

(564) *u*   *tɕiŋ-lan=ko*   *tɕoʔ=ko*   *kuɾa*   *sjaw=a.*
       REM   Cing-spirit=GEN   child=GEN   thing.matter   become=PST

       'The topic of the child of the Cing spirit is done.'
       CH_CTW_BBC_POL_111720_1_Cing Lan

(565) *mʰeʔ   law=ʌ=sa   pʰe=ka=n=i,*
       fire   apply_wear=LN=NMZ1   leave_abandon=2/3.PST=DIR/TR=PL

       ane   u   mʰeʔ   law=ʌ=sa   pʰe=ti, (...)
       so   REM   fire   apply_wear=LN=NMZ1   leave_abandon=SEQ1

       u   pahw=lʌ   kja !
       REM   be.allowed.to=NEG   PART

       'They stopped setting up fire, so, having stopped this setting up fire thing, (…)
       this (practice) is not allowed, that's the thing.'
       CH_MKW_RC_JMC_SIL_120119_Conversation

4.1.3.4.   Discourse cataphoric use of the borrowed pronoun *u*

In addition to the 3rd person remote pronoun *u*, another morpheme *u* has a discourse cataphoric function. This morpheme was borrowed from Nepali which has a 3rd person pronoun ऊ <u> associated with a low degree of politeness. This pronoun developed this cataphoric function.

Like in Nepali, the morpheme *u* is used when the speaker is looking for the referent they intended to express, a way to hold the floor of the conversation while sustaining the attention of the addressee. This is illustrated in (566) with *u* used as a substitute for the noun *tɕʌlʌn* 'habit, custom.'

(566) *laʔ=ma   tɕʌlaw=sa=ko,   u   kʰe=to,*
       arrow=ADD   use=NMZ1=GEN   CAT   COP=REM.PST

       tɕʌlʌn   kʰe=to,   hojnʌ ?
       habit_custom COP=REM.PST no

       'There was that, the habit of using arrows, right?'
       CH_CTW_SBC_BBC_GUN_012120_Chepang Kings

The cataphoric pronoun *u* can function as a substitute for, not only a noun, but also a verb, a noun-phrase, or even a clause, as in (567) with a purposive subordinate

clause. Like in Nepali, the Chepang cataphoric pronoun *u* can carry nominal plural inflection, similarly as well to the native Chepang personal pronouns *i*, *o*, and *u*. This is illustrated in (568). But in by contrast with Nepali, it can inflect with a locative case marker, like *=haŋ* in (569) and (570).

(567) *o*   *Nak*   *pan*   *din*   *sʌmʌn*   *al=a,*
       DIST   Nag   five   day   until   go=PST

       u,   *ŋaʔ*   *kweʔ=laŋ !*
       CAT   fish   fish=PUR

       'That Nag went for five days, that… to fish!'
       CH_MKW_SC_SIL_122619_1_Nākko_Co'

(568) *ʌni*   *ŋa=bʰʌnda=ma*   *kuɾa,*   *u=taŋ=ko*   *u=lʌm,*
       then   1SG=CMPR=ADD   thing   REM=ALL=GEN   CAT=SML

       ʃristi=lʌm=kʰe,   *ŋa=kaj*   *tʰaha*   *na=lʌ*   *ane,*   *ʌm…*
       origin=SML=DIS   1SG=DAT   know   COP=NEG   PART   yes

       'So, these things are of that time before me, that, about the origin, I don't know then, yes…'
       CH_MKW_SRNDC_SIL_081818_Life

(569) *ʌm,*   *ʌ…*   *u=haŋ,*   *Narayanghat=haŋ,*   *pãs*   *bʌɾsʌ.*
       yes   uh   CAT=LOC1   Narayanghat=LOC1   five   year

       'Yes, uh… there, in Narayangath, five years.'
       CH_MKW_PMRC_SIL_081818_1_Life

(570) ʌni   o   dʌgi, u=haŋ   pʌr=ʌ=ti   mu=o
      then   DIST   now   CAT=LOC1   study=LN=SEQ1   COP=PERF

      kʰe=to…   u,   nãw=ma   waŋ=ŋʌ=lʌ   kja,   ʥʰʌttʌj…
      COP=REM.PST CAT   name=ADD   come=1=NEG   PART   instantly

      u   kja,   sistʌr=haŋ !
      CAT   part   sister=LOC1

'The now she, she had studied there… that… even the name doesn't come to me instantly, that (frustrating), at the sister's (place to learn about Christianity)!'
CH_MKW_PMRC_SIL_081818_1_Life

What is interesting is that, in addition to the presence of nominal inflectional morphology, Chepang developed attaching as well verbal morphology to the cataphoric pronoun *u* when substituting for a verb or nominalized verb. This is illustrated in (571) with a nominalized verbal argument, in (572) with a nominalized relative clause, in (573) with a main verb inflected for 2nd person and in (574) for 3rd person.

(571) hʌtpʌt   i=kaj   tjaw=taŋ   u=sa,
      rapidly   PROX=DAT   up_above=ALL   CAT=NMZ1

      kum=ti   lʌhn=sa   gahro
      weaving.in.circle=SEQ1   climb_reach.top=NMZ1   difficult

      sjaw=na.
      become=NPST

'To do that, rapidly making it reach the top weaving in circle becomes difficult.'
CH_MKW_MRNDC_SIL_081818_6_Dikalak

(572) u=o,   sʌʔj=o   baburi=lʌm,   ro=lʌm,
      cat=NMZ:REL   thread=NMZ:REL   wild.basil=SML   flower=SML

'That, threaded wild basils, flowers,'
CH_MKW_GBC_CYO_120119_Conversation

(573) *"ane   u=te=a,   i   tʰãw=kaj   naŋ   ʣʰak=te=a !"*
so   CAT=2=PST   PROX   place=DAT   2SG   like=2=PST

"ʣʰak=alaŋ   baba,   ʣʰak=alaŋ"   dajh=ti=taŋ   ʨoʔ-bʰaw
like=1.PST   father   like=1.PST   say=SEQ1=ATT   son-in-law

dahj=a.
say=PST

'"Well, you did that, you liked this place!"; "I liked it father, I liked it." said the son-in-law.'

CH_MKW_SC_SIL_122619_1_Nakko_Co'

(574) *kwej   u=na=i,   ʣʰja=na=i,*
some   CAT=NPST=PL   shaman.practice=NPST=PL

'Some do that, they practice shamanic ceremonies.'

CH_MKW_STC_SIL_120619_2_E_4

4.1.3.5.   Deictic locational use of the borrowed pronoun *uh*

Another related 3rd person pronoun is *uh*, borrowed from Nepali. It functions as a locational pronoun that means 'here' or 'over there' but is specifically used to deictically refer to people, animal, things, or places in the following context, often accompanied either with pointing or head gesture: pointing at or showing the direction of anything located nearby or farther away from the location of the speaker and the addressee, while triggering a mirative and attentional effect on the addressee. If a location is to be emphasized as extremely far, the form *uh* will not be used but a long and high pitch vowel /uː/. The locational pronoun *uh* is illustrated in (575) with a location, in (576) with goats, and in (577) with a person.

(575) *ama=kaj*   *uh,*   *Charkila=sʌj*   *waʔn=o*   *Kucur*
mother=DAT   PRO:LOC   Charkila=ABL   bring=PERF   Kucur

       to=o   *tʰãw=sʌj.*
       tell_say=NMZ:REL   place=ABL

       'My mother, from over there, (my father) brought her from Charkila, a place called Kucur.'

       CH_MKW_PMRC_SIL_081818_1_Life

(576) *uh,*   *metɕʰja*   *rʰam=ti*   *wah=sa,*   *tʰikkʌ !*
       PRO:LOC   goat   graze=SEQ1   walk=NMZ1   good

       'Over there look, the goats are hanging out, grazing, just well!'

       CH_MKW_RC_JMC_SIL_120119_Conversation

(577) *ŋa=ko*   *moʔm*   *uh !*
       1SG=GEN   granddaughter   PRO:LOC

       'My granddaughter, here!'

       CH_MKW_RC_JMC_SIL_120119_Conversation

The 3rd person remote pronoun and determiner *u*, the cataphoric pronoun *u*, and the location mirative or attentional morpheme *uh* are used together in (578).

(578) *deŋi*   *u*   *saj,*   *u...*   *saj*   *saj=a,*   *uh !*
       now   REM   fruit   CAT   fruit   grow=PST   PRO:LOC

       'Now, that fruit, what's that… the fruit grew, over there look!'

       CH_MKW_RC_JMC_SIL_120119_Conversation

Table 126 summarizes the uses of the 3rd person pronouns and determiners *i*, *o* and *u*, and of the two pronominal forms *u* and *uh* borrowed from Nepali.

Table 126. 3rd person pronouns and determiners functions

pro/det	deictic				discourse		
	present anaphoric	present contrast	absent / present remote	location to be noticed by the addressee	absent anaphoric	absent / anaphoric contrast	cataphoric holding the floor
i PROX	✓					✓	
o DIST					✓		
i … o		✓				✓	
u REM			✓				
u CAT							✓
uh LOC				✓			

4.1.3.6.   Discourse cataphoric use of the morpheme *waj*

The morpheme *waj* is also used in Chepang as a cataphoric pronoun to talk about people whose referent is forgotten at the time of speech. Like with the morpheme *u*, the morpheme *waj* is used to hold the floor the time of remembering what was to be said. This morpheme is used in particular in RAP-13. This is illustrated in (579).

(579) ʌ,   *waj=lʌm=i,*   *ʤʌnta=lʌm=i*   *dinka-din*
       uh   CAT=PL=ERG   people=PL=ERG   every.single.day

       gwej   *djaw=laŋ*   *al=dʰʌj,*
       yam   dig.up=PUR   go=PROG

       'Uh, they, the people are going to dig up yams every single day,'
       CH_CTW_BBC_POL_102420_3_Chepang_Kings

4.2. Nominal exclusive enclitic morpheme *=jal~al* 'only'

The encliticization of a morpheme *=jal~=al* to any personal pronoun or noun can occur to mean 'only.' It is illustrated in (580) with 1st person *ŋa*, in (581) with 1st person plural *ŋi*, in (582) with 3rd person *o*, and in (583) with the noun *kʌlʌm* 'pen.' This construction is largely being replaced by the adverb मत्र <matra> borrowed from Nepali, as shown in (584).

(580)   *ten,*   *ŋa=jal*   *waŋ=o.*
         today   1SG=EXCL   come=PERF

         'Today, only I came.'
         CH_CTW_SPC_POL_071521_E

(581)   *o=haŋ,*   *nʌja*   *manta,*   *ŋi=jal*   *hʌ,*   *ŋi=jal.*
         DIST=LOC1   new   person   1PL=EXCL   COP   1PL=EXCL

         'There, new people, (no,) there is only us, only us.'
         CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(582)   *(...) su=ko=ma*   *bʰaggja*   *na=lʌ*   *kʰe=ja=kaj*
         who=GEN=ADD   luck   COP=NEG   COP=COND=DAT

         ʌ,   *bʌstu*   *sjas=ti=ma*   *sjaw=lʌ,*
         uh   cattle   raise=SEQ1=ADD   become=NEG

         rʌ   *o=al=ma*   *kʰe=lʌ,*
         and   DIST=EXCL=ADD   COP=NEG

         'If someone is not fortunate, raising a cattle doesn't work either, and that is not the only thing,'
         CH_CTW_BBC_POL_111720_6_Jogi

(583) *ŋa=kusi*　　*ja=ʥo*　　*kʌlʌm=jal*　　*mu=na.*
　　　　1SG=COM　　one=CL1　　pen=EXCL　　COP=NPST

'I only have one pen.'

CH_CTW_SPC_POL_071521_E

(584) *"ŋa*　　*matɾʌ*　*mʌ=na=ŋ*　　*ɾaʥa,"*　　　*ɾja=ti,*
　　　　1SG　　only　　COP=NPST=1　king　　　　　say=SEQ1

Dagu　*ɾaʥa*　*ɾja=o*　　　　*kʰe=to.*
Dagu　　king　　say=PERF　　　COP=REM.PST

'"Only I remain king," he said, the king Dagu had said.'

CH_CTW_BBC_POL_102420_3_Chepang_Kings

4.3.　Other demonstrative pronouns

In § 4.1.3, I described the uses of the 3rd person pronouns *i*, *o* and *u*, and showed that they were not different from demonstrative pronouns and that they also functioned as determiners.

In Table 127, I present other types of demonstrative pronouns formed with the pronouns *i*, *o* and *u* to refer to: locational, quantity and size, manner and type, and time. When variation in pronunciation is attested, it is specified.

Table 127. Other types of demonstrative pronouns

	PROX	DIST	REM
location	*i=haŋ* 'here'	*o=haŋ* 'there'	*u=haŋ* 'over there'
	[ihãŋ], [ehẽŋ], [ẽŋ],[jʰãŋ]	[ohãŋ], [ohõŋ], [õŋ]	
	i=kʰa 'here'	*o=kʰa* 'there'	*u=kʰa* 'over there'
time	*i bela=haŋ* 'at this time'	*o bela=haŋ* 'at that time'	*u bela=haŋ* 'at that time'
			ui bela=haŋ
	i bela=kʰa	*o bela=kʰa*	*u bela=kʰa*
			ui bela=kʰa
quantity	*i=tɕuk~tɕjuk* 'this much'	*o=tɕuk~tɕjuk* 'that much'	*u=tɕuk~tɕjuk* 'that much'
	i=tɕik		
quantity	*i-tɕikajto* 'this little'		
manner, type	*i=tʌ* 'like this, this type'	*o=tʌ* 'like that, that type'	*o=tʌ* 'like that, that type'

4.4. Possessive pronouns and determiners

Possessive pronouns and determiners are similar, as shown in (585) and (586). They are formed with independent pronouns encliticized with the genitive morpheme =*ko* (§ 3.4.4.1). No specific morphophonological change is attested at morpheme boundary.

Possessive pronouns and determiners are presented in Table 128, and the possessive forms attested with the innovated pronouns in Table 129.

(585)   *ʌni*   *baʥja,*   *naŋ=ko*   *nam=kʰe*   *doh*   *ane ?*
        then   elder.M   2SG=GEN   name=DIS   what   PART

        'So, grandpa, what is your name then?'

        CH_MKW_KRC_HAT_012120_Being_Shaman

(586)   *e…*   *naŋ=ko=paj*   *sʌppej*   *naŋ=ko=le ?*   *naŋ=ko*   *eɾija=le ?*
        EXPR   2SG=GEN=DIS   all   2SG=GEN=DIS   2SG=GEN   area=DIS

        'I see… as for yours, all is yours? your territory?'

        CH_MKW_SC_SIL_122619_1_Nakko_Co'_Archive

Table 128. Possessive pronouns and determiners

	SG	DU		PL	
1	*ŋa=ko*	*ŋi~ni-tɕi=ko*		*ŋi~ni=ko*	
2	*naŋ=ko*	*niŋ-ʥi=ko*		*niŋ=ko*	
3		**HUM**	**N.HUM**	**HUM**	**N.HUM**
PROX	*i=ko*	*i=nis=tɕak=ko* *i=nis=ko*	*i=ko*	*i-mi~mʌj=lʌm=ko* *i-mi~mʌj=ko* *i=lʌm=ko*	*i=ko*
DIST	*o=ko*	*o=nis=tɕak=ko* *o=nis=ko*	*o=ko*	*o-mi~mʌj =lʌm=ko* *o-mi~mʌj=ko* *o=lʌm=ko*	*o=ko*
REM	*u=ko*	*u=nis=tɕak=ko* *u=nis=ko*	*u=ko*	*u-mi~mʌj=lʌm=ko* *u=lʌm=ko*	*u=ko*
1/2HON, IMPS	*lʌj=ko*				
2HON	*hʌʥuɾ=ko* (<N.) *uha=ko* (<N.)			*hʌʥuɾ=lʌm=ko* (<N.)	

Table 129. Possessive pronouns and determiners with innovated forms

		SG	DU		PL	
RAP-13/11	**2HON**	*tʌpaj=ko* (<N.)	*tʌpaj=nis=ko* (<N.)		*tʌpaj=lʌm=ko* (<N.)	
			HUM	**N.HUM**	**HUM**	**N.HUM**
MAN-4	**3**	*ow=ko*	*ow=nis=tɕak=ko*	*ow=ko*	*ow-mi~mʌj=lʌm=ko*	*ow=ko*
RAP-13	**DIST**		*ow=nis=ko*		*ow-mi~mʌj=ko*	
					ow=lʌm=ko	

4.5. Interrogative pronouns

Interrogative pronouns are presented in Table 130. The number of the illustrative examples showing their use are noted in the table.

Table 130. Interrogative pronouns

FORM	MEANING	e.g.
do~doh	'what'	(587), (588)
su~su-manta	'who'	(589), (590)
ga=o	'which'	(591)
gʌlʌ~gʌlʌ=kʰa	'when'	(592), (593)
ga=haŋ	'where'	(594)
ga=taŋ	'to where'	(595)
gʌ=tʌ	'how'	(596)
gʌ=tʌ haj=ti	'how to'	(597)
gʌ=tʌ tahŋ=to	'how is it'	(598)
gʌ~gu=tɕuk~tɕjuk	'how much'	(599)
dʌjti	'why'	(600)

(587)   *djah*   *o*   *lan=kaj*   *doh*   *ɖzahŋ=sa ?*
        now   DIST   spirit=DAT   what   do_make=NMZ1

        'Now, what to do for that spirit?'

        CH_CTW_JMC_PYK_101920_Cing_Lan

(588)   *ʌni*   *naŋ=ko*   *budʰa=i*   *doh*   *kam*   *ɖzahŋ=na=u ?*
        so   2SG=GEN   husband=ERG   what   work   do_make=NPST=3O/DIR

        'So, what work does your husband do?'

        CH_MKW_SC_DAM_112819_Conversation_with_Bipana

(589)   *oho*   *ten*   *su*   *waŋa ?*
        EXPR   today   who   come=PST

        'Oho, who came today?'

        CH_MKW_CPR_BAN_102817_1_Mit_Co'

(590)   *su*   *manta*   *waŋ=a ?*
        who   person   come=PST

        'Who came?'

        CH_MKW_1_73-79_CPR_BAN_102417_2_Verb

(591)   *ʌni*   *ʌɾu*   *ʌɾu*   *ga=o*   *ga=o*   *sjaʔn ?*
        so   other   other   INT=NMZ:REL   INT=NMZ:REL   insect

        'So, which other insects (are there)?'

        CH_MKW_BMB_KW_BAN_110619_Conversation_About Minuscule_Worms in love

(592)   *i=kaj*   *gʌlʌ*   *meme-lan=i*   *ɖze=na=tʰʌ=i ?*
        PROX=DAT   when   Meme-spirit=ERG   eat=NSPT=INV=3>3SG

        'When is the Meme spirit going to eat her?'

        CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

(593)　ʌhʌ　　ŋa=ko　　tokɾak　kwa　　　gʌlʌ=kʰa　　dah=na　　ane ?
　　　　EXPR　1SG=GEN　　frog　　bound.friend　when=LOC2　reach=NPST　PART

'When is my bound friend frog gonna arrive then?'

CH_CTW_SP_POL_111420_Ream_Tokrak

(594)　metɕʰja　　　ga=haŋ　　　dʑjal=a　　da ?
　　　　goat　　　　　INT=LOC1　　leave=PST　PART

'Where did the goat leave, hey?'

CH_MKW_RC_JMC_SIL_120119_Conversation

(595)　"aj　　　　　　　ga=taŋ　　　al=a ?"=tʌ　　to=na=ŋ,
　　　　mother.in.law　　INT=ALL　　go=PST=REP　tell_say=NPST=1

　　　　naŋ　　tjaŋ=ʌ.
　　　　2SG　　reply=2SG.IMP.INTR

'I ask you "Where did your mother-in-law go to?," You reply.'

CH_MKW_RC_JMC_SIL_120119_Conversation

(596)　lʌw　gʌ=tʌ　　　　　gʌ=tʌ　　　　　　dʑahŋ=sa
　　　　well　INT=NMZ:ADV1　INT=NMZ:ADV1　do_make=NMZ1

　　　　pʌɾ=o　　　　　　　　ba ?
　　　　have.to_fall=NMZ:REL　　PART

'Well, how do we have to do, what do you think?'

CH_CTW_KRC_HAT_012120_Being_Shaman

(597)　ŋa　　gʌ=tʌ　　　　　haj=ti　　wah=tɕja=ŋ ?
　　　　1SG　INT=NMZ:ADV1　do=SEQ1　walk=IRR=1

'How would I do to leave?'

CH_CTW_JMC_PYK_101920_Cing_Lan

(598)　ʌni　ten=ko　　　kjan　gʌ=tʌ　　　　　tahŋ=to
　　　　so　　today=GEN　dish　INT=NMZ:ADV1　be.huge_be.like=NMZ:ADV2

rʌjsʌ ?
COP.MIR

'So, how does today's dish turn out be?'

CH_MKW_SLP_MMRBP_AJI_102519_1_Conversation

(599) *gu=tɕuk*   *din ?*   *dʌs*   *din*   *mu=ljam,*
INT=QTY   day   ten   day   cop=2SG.IMP.NEG

dwi   *hʌpta*   *mu=ja=ma*   *sjaw=na.*

two   week   COP=COND=ADD   happen=NPST

'How many days? don't stay ten days, even if you stay two weeks, that works.'

CH_MKW_BMB_KW_BAN_103119_10_Conversation_Friends

(600) *naŋ*   *dʌjti*   *ŋa=kusi*   *noʔ=te=lʌ*   *ane ?*
2SG   why   1SG=COM   speak=2=NEG   PART

'Why don't you speak with me then?'

CH_MKW_SC_SIL_122619_2_Dhobini_rani

4.6.   Interrogative-based indefinite pronouns and determiners

Interrogative-based indefinite pronouns and determiners form constructions that result from complex morphosyntactic processes, such as nominalization or the use of an additive morpheme *=ma.* They are presented in Table 131 along with the number of the associated example.

Table 131. Interrogative-based indefinite pronouns and determiners

	FORM	MEANING	e.g.		FORM	MEANING	e.g.
positive	*do~doh, do do*	'whatever'	(601)	negative	*do~doh=ma~mʌ*	'nothing'	(610)
	do=laŋ	'something'	(602)				
	su	'whoever'	(603)		*su=ma~mʌ*	'noone'	(611)
	su=laŋ	'someone'	(604)				
	ga=o	'whichever'	(605)				
	gʌlʌ=ma~mʌ	'whenever'	(606)		*gʌlʌ=ma~mʌ*	'never'	(612)
	ga=haŋ	'wherever'	(607)		*ga=haŋ=ma~mʌ*	'nowhere'	(613)
	do=haj=ti	'however'	(608)				
	gʌ~gutɕuk	'however much'	(609)				

(601) *do   mʌ=na,   o   ŋa=kaj   bʌj=tɕi.*
        what   COP=NPST   DIST   1SG=DAT   give=2SG>1SG.IMP

'Whatever it is, give that to me.'

CH_MKW_PC_SIL_062421_E

(602) *do=laŋ   sjaw=a.*
        something   happen=PST

'Something happened.'

CH_MKW_BC_SIL_011920_E

(603) *su=i   ŋa=ko   ɾaŋ   jaw=na=u,*
        who=ERG   1SG=GEN   field   plough=NPST=3O/DIR

o=ko   ɾaŋ   jaw=na=ŋ.

DIST=GEN   field   plough=NPST=1

'Whoever ploughs my field, I'll plough theirs.'

CH_MKW_PC_SIL_062421_E

(604) *su=laŋ*   *waŋ=a.*
        someone   come=PST

        'Someone came.'

        CH_MKW_BC_SIL_011920_E

(605) *ga=o*   *saj*   *mu=na,*   *o*   *waʔn=ʌ.*
        INT=NMZ:REL fruit COP=NPST   DIST bring=2SG.IMP.INTR

        'Whichever fruit it is, bring that.'

        CH_MKW_PC_SIL_062421_E

(606) *gʌlʌ=ma*   *waŋ=ŋʌ=to,*   *mu=jak=te=lʌ.*
        when=ADD   come=1=REM.PST   COP=REM.PST=2=NEG

        'Whenever I came, you were not (there).'

        CH_MKW_PC_SIL_062421_E

(607) *ga=haŋ*   *al=te=na,*   *ŋa=ma*   *al=na=ŋ.*
        INT=LOC1   go=2=NPST   1SG=ADD   go=NPST=1

        'Wherever you'll go, I'll go too.'

        CH_MKW_PC_SIL_062421_E

(608) *do=haj=ti*   *sjan=o,*   *o=tʌ*   *haj=ti*
        what=do=SEQ1 teach_learn=PERF DIST=NMZ:ADV1 do=SEQ1

        dzahŋ=nu.
        do_make=2PL.IMP.TR

        'However you've learnt it, do it doing it that way.'

        CH_MKW_PC_SIL_062421_E

(609) *gʌ=tɕuk*   *kam*   *ʥahŋ=na=u,*
INT=QTY   work   do_make=NPST=3O/DIR

     o=tɕuk   *pʌjsa*   *waŋ=na.*
     DIST=QTY   money   come=NPST

     'However much work one does, one gets that much money.'
     CH_MKW_PC_SIL_062421_E

(610) *doh=ma*   *sjaw=lʌ,*   *bistaɾi-bistaɾi*   *ʥahŋ=dʰʌj=le,*
     what=ADD   become=NEG carefully   do_make=PROG=DIS

     'Nothing will happen, we'll do that carefully,'
     CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(611) *ʌni*   *ʌɾu,*   *su=ma*   *noʔ=waj=lʌ.*
     so   other who=ADD   speak=CERT=NEG

     'So, as for someone else, no one speaks at all.'
     CH_MKW_MRC_LPK_080918_3_Chepang_Language

(612) *o*   *gʌlʌ=ma*   *bʰet* *sjaw=ma=l=o*
     DIST   when=ADD   meet become=NEG=COP=NMZ:REL

     manta   *ma=ba.*
     person   PART=PART

     'He is someone one never meets, don't you think?'
     CH_MKW_BRC_CYO_120119_Yukdhung

(613) *joh=ma*   *kim=kʰa=le*   *mu=o,*
     yesterday=ADD house=LOC2=DIS   COP=PERF

     ga=haŋ=ma   *al=ŋʌ=lʌ.*
     INT=LOC1=ADD   go=1=NEG

     'Yesterday too, I stayed at home, I didn't go anywhere.'
     CH_MKW_STC_SIL_120619_1_E

CHAPTER V

VERBS AND VERBAL MORPHOLOGY

In this chapter, I describe Chepang verbs and associated derivational and inflectional morphology.

I start by defining what a verb is by contrast with nouns (§ 5.1). I show that at the root level, verbs have preserved morphological traces of old productive derivational suffixes that have cognates in other TH languages (§ 5.2). I present two non-productive morphemes *kʌ* and *sʌ* that are found with some verbs and whose function remains unclear while likely being as well traces of old derivational morphology (§ 0). I describe the morphosyntactic differences attested between stative intransitive and intransitive verbs and show that stative intransitive verbs are used as adjectivals in nominalized constructions (§ 5.4). I briefly describe the morphological means by which verbs are borrowed from Nepali (§ 5.5). I provide a verbal template that shows all the morphemes attested in a verbal stem in their paradigmatic and syntagmatic distributions (§ 5.7). I describe the productive derivational morphology attested with verbal roots (§ 5.6), the causative morpheme *=tak* (§ 5.6.1) and the emotional anti-experiencer morpheme *=se~si~sje~sja* (§ 5.6.2).

I then describe Chepang verbal inflectional morphology (§ 5.8). I start by giving an introduction (§ 5.8.2) which summarizes previous attested accounts of Chepang verbal argument indexation (§ 5.8.2.1) and shows that while it is possible to consider that Chepang verbal indexation forms a direct-inverse system, this latter, as in other TH languages, is highly non-canonical by comparison to other direct-inverse systems attested typologically, and questions the functional validity a hierarchical system beyond such system (§ 5.8.3.3). I then provide an analysis of Chepang verbal argument indexation attested with each type of verbal construction along with their full paradigms, annotated for variation and compared with forms attested in earlier literature (§ 5.8.3). I finally give an overview of morphological tense, aspect and modality (§ 5.8.4).

I use the following terms and abbreviations: intransitive verbs (vi.) have a single argument (S) and a possible (T) argument; transitive (vt.) verbs two arguments, i.e., a

most Agent-like argument (A) and a most Patient-like argument (P); ditransitive (ditr.) have three arguments, a most Agent-like argument (A), a most Patient-like argument (P), and a Recipient (R).

5.1. Noun and verb distinction

As briefly laid out in § 3.2, Chepang nouns and verbs are distinguished by the type of inflectional and derivational morphology they can take and their access to syntactic functions, i.e., that of argument or predicate.

A noun needs the presence of a copula to serve the function of predicate in a clause, as with *pande* 'shaman' in (614), while verbs do not; verbs require nominalizing morphology to carry the function of an argument in a clause, as with *tɕiʔ=ma=l=o* 'people we don't know' in (615), while nouns do not.

(614) ʌ̃   *dʑʰja=te=a=ma,*
       yes   shaman.pratice=2=PST=ADD

       naŋ   pande   kʰe=na   ɾa ?
       2SG   shaman   COP=NPST   PART

       'Yeah right, you practiced a shamanic ceremony, you think you're a shaman?'
       CH_MKW_RP_SP_CHI_102519_2_Conversation

(615) *tɕiʔ=ma=l=o=kaj=paj*   *ŋi   to=waj=lʌ.*
       know=NEG=COP=NMZ:REL=DAT=DIS   1PL   tell_say=CERT=NEG

       'We don't tell (it) at all to people we don't know.'
       lit. 'We don't tell (it) at all to those who are not known.'
       CH_CTW_KRC_HAT_012120_Being_Shaman

Some roots may occur as noun or verb in absence of such morphosyntactic devices, showing both nominal and verbal properties. While these roots likely come from the same historical origin or etymological source morpheme, they may either have preserved a similar meaning, as in (616), or shifted semantically, as in (617). A historical

492

analysis of verbal and nominal morphology allows us to better understand the semantic origin and evolution of such forms.

Chepang features root final and sonorant pre-final consonants that are traces of ancient derivational morphology. Such endings form pairs, triplets or more members of sets constituted of verbs and/or verbs and nouns that can be reconstructed back to the same origin. While the meanings of the pair in (617) seem to be semantically unrelated and their form the result of arbitrariness, the existence of the verb in (618) helps understanding its historical origin and semantic and morphosyntactic development in addition to that of the other members of the group: the verb *laʔ- (ti)* 'swim' comes from the meaning of 'grab, seize water (make the water be grabbed, seized),' and *laʔ* from the meaning of 'the one that grabs, seizes (make something be grabbed, seized),' as a result of the morphological presence of the glottal stop /ʔ/ on the transitive verb *la-* 'grab, seize.' The glottal stop is a historical causative derivational morpheme *-ʔ*.

Such derivational traces along with examples of pairs, triplets, or more members, are described in § 5.2.

(616)   *kueʔ~keʔ*
        'fishhook'

        kueʔ~keʔ- (ŋaʔ)
        'fish'

        lit. 'hook (fish)'

(617)   *laʔ*
        arrow

        laʔ- (ti)
        'swim'

        lit. 'make be grabbed (water)'

(618)   *la-*
        'seize, grab'

5.2. Ancient derivational morphological traces at the root level

Amongst the 1,200 verbal roots collected, around 140 verbs belong to a pair, triplet or to a group consisting of up to five members that may include verbs or verb(s) and noun(s) whose semantics is related. The various final consonants of such sets correspond to old traces of derivational morphology, such as: *-h*, *-s*, *-k*, *-ʔ*, *-t*, and *-p*. In this section, I will describe each of them, in a synchronic and diachronic approach, with a selection of eleven pairs and seven triplets.

To better understand the patterns and observations presented here, our team members and I compared all the possible argument structures of 280 verbs, which included determining what triggers the presence or absence of overt arguments, the type of case marking they may take, the type of argument indexation occurring on the verb, and the nature of the context in which they occur to account for possible pragmatic effects.

In the following sub-sections, I discuss the distribution of the derivational consonants over attested pairs, triplets, or groups. They are reported in Table 132.

The distribution in this table is organized by type of grouping, i.e., pair, triplet, or larger group, and reports the number of sets (in the column headed by #) attested for a particular co-occurrence pattern of derivational consonants (noted ✓); for instance, four pairs are attested where each consists of a verbal root ending in a vowel and a verbal root featuring the derivational consonant *-k* (first line of the table). In addition, when the final consonant of a root underwent a phonological change at a morpheme boundary due to the presence of a laryngeal derivational consonant (observed when in combination with nasals), I note * the reconstructed final consonant of the root and < its current phonological shape.

Table 132. Distribution of derivational consonants over pairs, triplets, and groups

		#VF	#VF~VCF[son] *-h						*-s	*-k	#VF~VCF[son] *-ʔ						*-t	*ʔ	#VCF[son]				
	#	V	h	hl	hɾ	hm	hn	hŋ	s	k	ʔ	ʔl	ʔɾ	ʔm	ʔn	ʔŋ	t	p	l	ɾ	m	n	ŋ
pairs	4	✓								✓													
	4	✓									✓												
	2	✓	✓																				
	1	✓							✓														
	1	✓																✓						
	1							✓															<	*
	3		✓							✓													
	3								✓	✓													
	1		✓								✓													
	1						<	*																✓
	1															<	*							✓
	1							✓									✓							
	2				✓																✓			
	1					✓																✓		
	1								✓									✓						
	2														✓							✓		
	2																✓							✓
	1										✓							✓						
	1											✓									✓			
	1															✓		✓						
	1																	✓					✓	
	1		~	✓														✓						
triplets	2	✓	✓								✓													
	2	✓										✓						✓						
	2	✓							✓	✓													
	2		✓						✓	✓													
	1				✓									✓							✓			
	1					✓									✓							✓		
	1								✓	✓								✓						
	1								✓		✓							✓						
4	1	✓	✓							✓								✓					
5	1	✓														✓		✓	✓				✓	
	1	✓							✓	✓	✓							✓						

495

Given the observed distribution in Table 132, different preliminary observations can be made along with the raise of a few questions:

(1) The derivational morphological traces attested are the following: *-h*, *-s*, *-k*, *-ʔ*, *-t*, and a somewhat unexpected consonant -*p* which has not been described in historical TH literature as having derivational properties.

(2) The type of root final endings to which such derivational morphology is observed are final vowels and the final sonorant consonants /m/, /n/, /ŋ/, /l/, and /ɾ/. While the combination of a laryngeal derivational consonant -*h* or -*ʔ* and a verbal root whose final is an approximant /w/ or /j/ is attested for many verbs, no significant pair[48], triplet or group was observed; therefore, I did not include such combination in this selection.

Note that synchronically, the glottal fricative /h/ and stop /ʔ/ are phonologically transcribed as preceding root final sonorant consonants, which corresponds to their pronunciation when the root is followed by a consonant initial syllable. When followed by a vowel initial syllable, the glottal fricative /h/ may shift from its position to that of initial position of the following syllable, as in: *dʑahŋ-sa* [dʑãhŋ.sa] 'do, make' vs. *dʑahŋ-alaŋ* [dʑãŋ.ha.laŋ] 'I did/made.' This shift, which is not attested in all cases, seems to be triggered as well by prosody or intonation, where articulatorily, the higher the pitch is, the less likely the shift is to happen. This type of shift has not yet been observed with the glottal stop /ʔ/.

The synchronic position of the derivational consonant within the root, when preceding final sonorants does not undermine the fact that such consonants may be reconstructed as suffixing or encliticizing morphology – rather it is likely that such laryngeal features have developed the property to shift around sonorant segments, since it is still observed synchronically with -*h*.

(3) As for the verbal roots' final endings, a consonant /n/, while expected to occur in a pair, triplet, or group where a laryngeal consonant *-h* or *-ʔ* would occur, is attested in a pair where the other member is a derivational consonant *-t*. This makes the consonant /n/ a new member of the derivational consonant series by default. It is possible

---

[48] Only one pair is attested with the root final approximant /j/; it is described in § 5.2.3.4.

that in such a case, the form /n/ may be the result of a sound change involving for instance the final glottal *-ʔ and a verbal root final consonant, or that a member of an original triplet formed with a laryngeal and a verbal root final /n/ disappeared; another pair is found where the derivational consonants *-t and *-ʔ are both present, and where the glottal *-ʔ combines with a verbal root final /n/ but where an expected verb ending in /n/ is absent.

(4) One can observe the presence of a higher number of pairs and triplets constituted by *-h, *-s, *-k, and *-ʔ, with the following distribution:

- final vowel, *-k
- final vowel, *-ʔ
- final vowel, *-k, *-ʔ
- final vowel, *-h, *-k
- *-h, *-k
- *-s, *-k
- *-h, *-k, *-ʔ

(5) Any type of complementary distribution, or in other words, the presence of a consonant corresponding to the absence of another throughout the pairs, triplets, or groups, could explain potential predictable sound change in a given environment or the loss of a member. The co-occurrence of a consonant with another inside pairs, triplets, and groups is represented in Table 133.

One can see that: (a) *-k and *-ʔ co-occur in any pair, triplet, or group with any type of other derivational consonant; (b) *-t co-occurs in any pair, triplet, or group with any type of other derivational consonant but *-h; (c) *-s and *-h do not co-occur in any pair, triplet, or group; (d) -p co-occurs in any pair, triplet or group with any derivational consonant but *-s; (e) as for the laryngeal *-h and *-ʔ, which may attach to a vowel final root or combine with a sonorant final root, two roots involving the presence of the same glottal feature in a pair, triplet, or group is not attested.

(6) Finally, the distribution of such derivational consonants over specific verbal roots that feature a final vowel or a sonorant, entails an asymmetry with regard to other types of verbal roots' final endings.

In Chepang, as described in § 2.3.2. The final consonant of a morpheme may natively either be a voiceless stop (/p/, /t/, /k/, /ʔ/) or fricative (/h/, /s/) or a non-breathy sonorant (/m/, /n/, /ŋ/, /ɾ/, /l/, /j/, /w/). This raises a few questions regarding the nature and origin of such derivational morphology.

First, not all verbs belong to a set of semantically related verbs showing different types of valency or argument structures; that is, although the series of derivational consonants are similar to the possible native final consonants attested in Chepang, it does not mean that these latter are all traces of such derivational morphology, since no consistent patterns are observed between each of them and their associated valency or argument structures, although some do, even when they do not belong to a set.

Second, if all final consonants similar to the derivational consonants are morphosyntactically unrelated, what type of derivational morphology then occurred at the same proto stage where they were used productively? Would it be possible to consider that part of the verbs that do not entail a correspondence between their final consonants and their argument structure would come from another stock of vocabulary, or would it be rather the case that all verbs were formed through the presence of these derivational morphemes before their majority shifted semantically? Is it possible that today's distribution of the derivational consonants may be explained through the distribution of synchronically productive derivational morphology, such as the causative morpheme =*tak* or the emotional anti-experiencer morpheme =*se~si~sje~sja*. Such questions are important to better understand what can be reconstructed at the level of Proto-Chepang (PC) and beyond, since cognate derivational morphological traces are attested in other TH languages; they will nevertheless be left aside for now since such questions may only find answers through the comparison of all possible verbal argument structures with regard to their morphology and their possible modification through today's derivational processes.

Table 133. Co-occurrence of derivational consonants within pairs, triplets, and groups

	*-h	*-s	*-k	*-ʔ	*-t	*-p	*-n
*-h		-	✓	✓	-	✓	-
*-s	-		✓	✓	✓	-	-
*-k	✓	✓		✓	✓	✓	-
*-ʔ	✓	✓	✓		✓	✓	✓
*-t	-	✓	✓	✓		✓	✓
*-p	✓	-	✓	✓	✓		✓
*-n	-	-	-	✓	✓	✓	

To analyze the different derivational functions of the consonants *-h, *-s, *-k, *-ʔ, *-t, I compared the semantics, valency, and argument structures of pairs, triplets, and groups. In addition to comparing the derived verbs within their pair, triplet, or group, I compared these latter with one another to confirm the possible argument structure(s) derived by each consonant. I only focused on *-h, *-s, *-k, *-ʔ, and *-t, since no significant explanation could be found yet for the presence of -p, which could be the result of a morphophonological change at morpheme boundary between the root and derivational consonant.

Since it is old unproductive morphology, it is not always possible to completely recover the original semantics, valency and argument structure(s) of the verb to which they attach or to understand that of intermediate stages of the derived form whose semantics, valency or argument structure(s) may have changed. Indeed, it is likely that such derivational morphology was productive at least before the split from a common ancestor between Chepang-Bhujel and Magar and probably beyond since cognate verb final roots are found across TH languages.

I present the main observations that result from historical and synchronic analyses made for each derivational consonant, in addition to examples and possible reconstructed forms, as follows:

- Derivational consonant *-h
- Derivational consonant *-s
- Derivational consonant *-ʔ
- Derivational consonant *-k
- Derivational consonant *-t

5.2.1. Derivational consonant *-h

The derivational consonant *-h is a valency decreasing device that was used to derive a transitive verb into an intransitive middle causative.

Structurally, only an S argument occurs, and the verb is intransitive. The S argument holds the semantic role of both the causer and causee. The derived verb is a middle causative because the S argument is both the A and P argument of a semantic transitive process where A acts on itself (noted between [ ]).

This intransitive middle causative derivation may be translated as follows:

'x makes self be VERBed'

This intransitive middle causative derivation can be schematized as follows, where x and y indicate the identity of particular participants' referents:

{Ay Px Vtr} > {S [Ay Py] Vintr}

This is illustrated through the following three pairs and four triplets:

root	type	meaning	reconstruction
gljun-	tr	'take out'	*gljun- 'y take out x'
gljuhŋ-	intr	'go out'	*gljun-**h** 'y make self be taken out'
djan-	tr	'return, send sth back'	*djan- 'y return x'
djahn-	intr	'come back, return'	*djan-**h** 'x make self be returned'
gul-	tr	'drive away'	*gul- 'y drive away x'
guhl-	intr	'follow, chase'	*gul-**h** 'x make self be driven away'

500

ja-	tr	'filter liquid'	*ja-* 'y settle x'
	tr	'block, settle non-liquid parts'	*ja-* 'y settle x'
jah-	intr	'be night, set (sun)'	*ja-**h** 'x make self be settled'
jak-	tr	'catch, stop, prevent from falling'	*ja-k* 'y cause x to be settled'
	intr	'settle, stop in location'	*ja-k* 'x caused to be settled (by y)'
	intr	'be stuck (food in throat)'	*ja-k* 'x caused to be settled (by y)'
ˡˡʰu-	tr	'pour out (fertilizer, dirt, sand)'	*ˡˡʰu-* 'y pour out x'
lʰuh-	intr	'be poured out'	*ˡˡʰu-**h** 'x make self be poured out'
lʰuk-	tr	'powder sb'	*ˡˡʰu-k* 'y cause x to be poured out on'
sɾa-	tr	'sow, scatter (grains, seeds)'	*sɾa-* 'y scatter x'
sɾah-	intr	'be mixed together (grains, seeds)'	*sɾa-**h** 'x make self be scattered'
		'be colorful (grains, seeds)'	
sɾak-	tr	'spray, splash, sprinkle on'	*sɾa-k* 'y cause x to be scattered on'
juɾ-	tr	'squeeze boil, furuncle'	*juɾ-* 'y squeeze x'
juhɾ-	intr	'dive, be buried into'	*juɾ-**h** 'x make self be squeezed'
juʔɾ-	tr	'extract (fruit juice)'	*juɾ-ʔ* 'y cause x to be squeezed'
	intr	'be extracted (fruit juice)'	*juɾ-ʔ* 'x caused to be squeezed (by y)'

5.2.2. Derivational consonant *-s*

The derivational consonant *-s* is a valency increasing device that was used to derive a stative intransitive verb into a transitive causative. It also has the function of an applicative when the original verb is transitive. This causative applicative derivation turns the transitive verb into a ditransitive verb where the oblique locational argument is promoted to the function of R argument.

Two processes may happen, given the original valence of the verb, such as:

(1) The S of the stative intransitive verb becomes the P argument of the derived transitive causative verb; it holds the semantic role of causee. A new A argument is expressed and holds the semantic role of causer.

This transitive causative derivation may be translated as follows:

'y causes x to be VERBed'

This transitive causative derivation can be schematized as follows:

{Sx Vst.intr} > {Ay Px Vtr}

(2) The A argument of the original verb remains the A argument of the derived verb and holds the semantic role of causer. The P argument of the transitive verb becomes the T argument of the derived ditransitive causative, while the oblique locational argument becomes the R argument, which assumes the semantic role of causee.

This causative applicative derivation may be translated as follows:

'y causes z to be VERBed x'

This causative applicative derivation can be schematized as follows:

{Ay Px Vtr (LOCz} > {Ay Tx Rz Vditr}

This is illustrated through the following triplet and pair:

root	type	meaning	reconstruction
nu-	st.intr	'be hidden'	**nu-* 'y be hidden'
nus-	tr	'hide sth'	**nu-s* 'y cause x to be hidden'
nuk-~nukʌ	intr	'hide self'	**nu-k* 'x caused to be hidden (by y)'
ka-	tr	'put sth (in sth)'	**ka-* 'y put x (inside z)'
kas-	ditr	'feed sb sth'	**ka-s* 'y cause z to be put in x'

5.2.3. Derivational consonant *-ʔ

The derivational consonant *-ʔ is a valency increasing device that was used to derive intransitive and transitive verbs. The valence is increasing through the addition of an argument that holds the semantic role of causer; the causee physically performs the process expressed by the caused event, not the causer.

The derived construction may be syntactically expressed intransitively or transitively, which means that such a causative derivation entails labiality, and in particular P lability.

When attached to an intransitive verb, the derivation results in a transitive causative (1) or an intransitive anticausative verb (2), such as:

(1) the S argument of the intransitive verb becomes the P argument of the derived transitive causative verb, while a new A argument is added, and the verb becomes transitive.

This transitive causative derivation may be translated as follows:

'y causes x to VERB'

This transitive causative derivation can be schematized as follows:

{Sx Vintr} > {Ay Px Vtr}

(2) the S argument of the intransitive verb becomes the S$_P$ argument of the derived intransitive anticausative verb. In this case, no additional A argument may be expressed, and the verb remains intransitive.

This intransitive anticausative derivation may be translated as follows:

'x is caused to be VERBed'

This intransitive anticausative derivation can be schematized as follows:

{Sx Vintr} > {S$_P$$x$ Vintr}

This is illustrated in the following four pairs and one triplet:

root	type	meaning	reconstruction
al-	intr	'go'	**al-* 'x go'
a?l-	tr	'take away'	**al-?* 'y cause x to go'
	intr	'be taken away'	> **al-?* 'x be taken away/caused to go (by y)'
waŋ-	intr	'come'	**waŋ-* 'x come'
wa?n-	tr	'bring'	**waŋ-?* 'y cause x to come'
	intr	'be brought'	> **waŋ-?* 'y be brought/caused to come (by y)'
na-	intr	'be, be many'	**na-* 'x is'
na?-	tr	'give birth'	**na-?* 'y cause x to be'
	intr	'be born'	**na-?* 'x be born/caused to be (by y)'
no	n	'ear'	**no* 'ear'
*no-	*st.intr	'be heard'	**no-* 'x be heard'
no?-	*tr	*'speak'	**no-?* 'y cause x to be heard'
	intr	'speak'	**no-?* 'x caused to be heard (by y)'
saj-	(st)intr	'be grown, grow'	**saj-* 'x be grown/grow'
sa?j-	tr	'hear, listen'	**saj-?* 'y cause x to grow'
	intr	'be heard'	**saj-?* 'x caused to grow (by y)'

When attached to a transitive verb, the derivation does not result in a ditransitive verb but in a transitive indirect causative verb (1) and an intransitive anticausative verb (2), such as:

(1) The A argument of the transitive verb is expressed and remains the syntactic A argument of the derived transitive indirect causative verb; it holds the semantic role of causer of the caused event but is not the direct causer. A new semantic causer participant is introduced semantically; it performs the action of the caused event, but it is not expressed or present in the argument structure of the derived verb. The P argument of the

504

transitive verb remains the P argument of the derived transitive indirect causative verb. The P argument holds the semantic role of causee.

This transitive causative derivation may be translated as follows:

'*y* causes *x* to be VERBed (by *z*)'

This transitive causative derivation can be schematized as follows:

{A*y* P*x* Vtr} > {A*y* P*x* Vtr (by *z*)}

(2) The P argument of the transitive verb becomes the S argument of the derived intransitive anticausative verb and holds the underlying function of a S$_P$ argument. The A argument, performer causer of the caused event, is not expressed and the verb becomes intransitive.

This intransitive anticausative derivation may be translated as follows:

'*x* is caused to be VERBed (by *z*)'

This intransitive anticausative derivation can be schematized as follows:

{A*y* P*x* Vtr} > {S$_P$*x* Vintr (by *z*)}

This is illustrated through the following triplet and pair:

root	type	meaning	reconstruction
juɾ-	tr	'squeeze boil, furuncle'	**juɾ-* '*y* squeeze, extract, express *x*'
juhɾ-	intr	'dive, be buried into'	**juɾ-h* '*x* make self be squeezed'
juʔɾ-	tr	'extract (fruit juice)'	**juɾ-ʔ* '*y* cause *x* to be expressed/extracted'
	intr	'be extracted (fruit juice)'	**juɾ-ʔ* '*x* caused to be expressed/extracted (by *z*)'
om-	tr	'cover (cloud)'	**om-* '*y* cover *x*'
oʔm-	tr	'brood (egg)'	**om-ʔ* '*y* cause *x* to be covered (by *z*)'

When attached to a transitive verb, the derivation may result in another set of derived verbs: a transitive causative verb (1) or an intransitive anticausative verb (2).

(1) The A argument of the transitive verb remains the A argument of the derived transitive causative verb; it holds the semantic role of causer. The P argument of the transitive verb remains the P argument of the derived transitive verb; it holds the semantic role of causee. However, A and P of the predicate of cause are inversely P and A of the predicate of effect (noted between [ ]).

This transitive causative derivation may be translated as follows:

'y causes x for (y) be VERBed'

This transitive causative derivation can be schematized as follows:

{Ay Px vtr} > {Ay Px [Ax Py] vtr} or {Ay[Py] Px[Ax] vtr}

(2) The derived causative transitive verb in (1) can be used in an intransitive anticausative construction where only the A argument of the predicate of cause or P argument of the predicate of effects remains.

This intransitive anticausative derivation may be translated as follows:

'y be VERBed'

This intransitive anticausative derivation can be schematized as follows:

{Ay Px vtr} > {SPy vintr}

This is illustrated through the following two triplets:

root	type	meaning	reconstruction
la-	tr	'seize, grab sth'	*la- 'y grab x'
***la?-** (ti)*	*tr	*?	*la-$?$ 'y cause x for y to be grabbed' (x=water)
	intr	'swim'	*la-$?$ 'y be grabbed (by x)'
lat-	tr	'carry on arms'	*la-t 'x cause y to grab x' (< *x=child)
go- (siŋ)*	tr	'look for, collect (wood)'	*go- 'y look for x'
***go?-**	tr	'call sb's attention'	*go-$?$ 'y cause x for y to be looked for'
	intr	'chirp' (bird)	*go-$?$ 'y be looked for (by x)'
got-	tr	'call sb to come over'	*go-t 'x cause y to look for x'

This latter set of verbal derivations entails a complex construction which may find a historical explanation that would go back to the common ancestor of Chepang-Bhujel and Magar. In fact, such constructions may have developed off an intransitive middle causative derived verb. Cognate constructions of the verb roots *go-* 'look for, collect,' and *la-* 'seize, grab' are found in Western Magar (Pons, data), as follows:

Western Magar cognates

morpheme	type	meaning
la-	tr	'take, keep, receive sth'
lah-	intr	'stick self, closely remain'
lak-	tr	'bring sth'
		'clay the house, paint'
goh-	tr	'look for, collect'
goh-ak-	ditr	'make sb look for, collect sth'

Given these Magar cognate roots, two intransitive middle causative derivations that are not attested in Chepang can clearly be reconstructed back to the ancestor of Chepang-Bhujel and Magar, since their derivational process is the one that has been described for the Chepang reconstructed derivations (§ 5.2.1), as follows:

(1) **la-h* 'x make self be grabbed'
> CH: *not attested*
> MAG: 'stick self'

(2) **go-h* 'x make self be looked for/collected'
> CH: *not attested*
> MAG: 'look for, collect'

It is possible that the causative derivation of such intransitive middle causative derivations at the proto level was structurally preserved in Magar, and that in Chepang,

the combination of the derivational consonant *-h with the derivational morpheme *-k (which is still attested in Magar as a morpheme -k and -ak) may have entailed a morphophonological change at morpheme boundary where the combination of *-h and *-k turned into *-ʔ, merging with the derivational causative consonant *-ʔ found with intransitive roots.

Such constructions may originally have been used in as a generic or impersonal statement, such as 'x can be verbed' or 'one can verb x,' since the referent of x is and was probably too originally an inanimate object.

At the proto-level, the following derivations may then be reconstructed with a meaning similar to the Chepang reconstructed forms *la-ʔ- 'x causes y for x to be grabbed/make self be grabbed' and *go-ʔ- 'x cause y for x to be looked for/make self be looked for,' which in Magar developed as 'x cause y to look for z,' as follows:

*la-h-k 'x causes y for x to be grabbed/make self be grabbed'
> CH: *la-h-k   *la-ʔ 'x causes y for x to be grabbed/make self be grabbed'
                *la-ʔ 'x be grabbed'
> MAG: *not attested*

*go-h-k 'x causes y for x to be looked for/make self be looked for'
> CH: *go-h-k   *go-ʔ 'x causes y for x to be looked for/make self be looked for'
                *go-ʔ 'x be looked for'
> MAG: *go-h-ak 'y cause x to look for z'

In the following sub-sections, I provide a detailed analysis of a few verb pairs formed with an intransitive verb and a transitive causative verb derived with the consonant *-ʔ, along with illustrative examples from natural discourse, as follows:
- *waŋ- – waʔn-*
- *al- – aʔl-*
- *na- – naʔ-*
- *no – noʔ- (bʰasa) & saj- – saʔj-*

5.2.3.1. *waŋ- – waʔn-*

The verb *waŋ-* 'come' is intransitive, its argument structure is: {S*x* Vintr}, as in (619). The verb *waʔn-* is intransitive anticausative 'be brought' or transitive causative 'bring;' its corresponding argument structures may be: {SP*x* Vintr}, as in (620), or {A*y* P*x* Vtr}, as in (621). The original meaning of the verb **waŋ-ʔ* is hypothesized to have been '*x* be caused to come (by *y*); *y* cause *x* to come.'

While the final consonant of the verb *waʔn-* 'be brought, bring' is not a velar nasal as in the verb *waŋ-* 'come,' such a phonological change is observed in other pairs involving the combination of a laryngeal derivational consonant and root final nasals.

(619) *nʌw   bʌɾsʌ=haŋ   ŋa=ko=tɕʌhe   guru   waŋ=a.*
       nine   year=LOC1   1SG=GEN=DIS   spiritual.teacher   come=PST

'At the age of nine, my spiritual teacher arrived.'
CH_CTW_BBC_PID_011520_1_Being_Shaman

(620) *ʌni   ja-dʑo   ŋa=ko   nata-natini*
      so   one-CL   1SG=GEN   grandchildren

      biha   dʑahŋ=ti   waʔn=a,   nek.
      wedding   do_make=SEQ1   be.brought|bring=PST   last.year

'And one granddaughter of mine married and got brought (to her husband's), last year.'
CH_MKW_PMRC_SIL_081818_1_Life

(621) *i   tɕoʔ-djaŋ=i,   ane   su=ko   tɕoʔ*
      PROX   daughter=ERG so   who=GEN   child

      waʔn=aka=n ?
      be.brought|bring=PST.2/3=DIR/TR

'This daughter, so whose son did she bring?'
CH_MKW_DBC_MAI_1_020220_Local_History

5.2.3.2. **al- – aʔl-**

The verb *al-* 'go' is intransitive; its argument structure is: {Sx Vintr}, as in (622). The verb *aʔl-* is intransitive anticausative 'be taken away' and transitive causative 'take away;' its corresponding argument structures may be: {SPx Vintr}, as in (623), or {Ay Px Vtr}, as in (624) and (625). The original meaning of the verb *al-ʔ* is hypothesized to have been '*x* be caused to go (by *y*); *y* cause *x* to go.'

(622) *naŋ=ko*   *ʨoʔ-djaŋ*   *ga=haŋ*   *al=a ?*
      2SG=GEN   daughter   INT=LOC1   go=PST

      'Where did your daughter go?'
      CH_MKW_SC_SIL_122619_1_Nakko_Co'

(623) *Bharatpur=ma*   *aʔl=a,*   *kjansʌr-ʌspʌtal=ma*
      Bharatpur=ADD   take.away=PST   cancer-hospital=ADD

      aʔl=a.
      be.taken.away|take.away=PST

      'He was taken to Bharatpur too, he was taken to the cancer hospital.'
      CH_MKW_SPMC_LC_SIL_100921_3_Conversation

(624) *Kathmandu*   *talim,*   *ʌ,*   *Chepang*   *bʰasa*   *bʌj=laŋ*
      Kathmandu   training   uh   Chepang   language   give=PUR

      aʔl=aka=n.
      be.taken.away|take.away=PST.2/3=DIR/TR

      'They took him to a training in Kathmandu, uh, to give the Chepang language.'
      CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(625)   *o*   *bela=haŋ=ma,*   *i*   *badʑa*   *Bhabikan=kaj*
        DIST   moment=LOC1=ADD   PROX   elder.M   Bhabikan=DAT

        i   *bʰasa=ko*   *lagi*
        PROX   language=GEN   for

        aʔl=a=tʰʌ=i   *Kathmandu.*
        be.taken.away|take.away=PST=INV=3>3SG   Kathmandu

        'At that time, this elder Bhabikan was taken to Kathmandu for this language.'
        CH_MKW_PSC_MAI_012620_1_Becoming_Christian

5.2.3.3.   *na-* – *naʔ-*

The verb *na-* 'be, be many' is intransitive; its argument structure is: {S*x* Vintr}, as in (626). The verb *naʔ-* is intransitive anticausative 'be born, be given birth' and transitive causative 'give birth;' its corresponding argument structures may be: {SP*x* Vintr}, as in (627) and (628), or {A*y* P*x* Vtr}, as in (629) and (630). The original meaning of the verb *na-ʔ is hypothesized to have been '*x* be caused to be (by *y*); *y* cause *x* to be.'

(626)   *bʌn=kʰa=le*   *na=na,*   *o=ma.*
        jungle=LOC2=DIS   be=NPST   DIST=ADD

        '(The porcupines) are in the jungle, they too are.'
        CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

(627)   *dʑetʰi-tɕoʔ-djaŋ*   *naʔ=a*   *dʌj !*
        elder.F-daughter   be.born|give.birth=PST   PART

        'My elder daughter was born, hey!'
        CH_MKW_BMB_BAN_100617_5_Birth

(628)   *ʌni*   *kwej=ko*   *gʰoda*   *tekʌ=ti*   *naʔ=na=i,*
        so   some=GEN   knee   stand.on=SEQ1   be.born|give.birth=NPST=PL

        kwej=ko   *namlo=haŋ*   *dʑʰundjaw=ti*   *naʔ=na=i,*
        some=GEN   rope=LOC1   hang.on=SEQ1   be.born|give.birth=NPST=PL

511

'So some people's children are given birth to while standing on the knees, some while hanging on a rope.'

CH_MKW_BMB_BAN_100617_5_Birth

(629) ʌni   ŋa=i=kʰe ?   ŋa=i=kʰe   dʑetʰi-tɕʰoɾi
      so   1SG=ERG=DIS   1SG=ERG=DIS   elder.F-daughter

      naʔ=alaŋ   dʌj !
      be.born|give.birth=1.PST   PART

'And me? Me, I gave birth to my elder daughter.'

CH_MKW_BMB_BAN_100617_5_Birth

(630) o=haŋ=sʌj=ko   ŋa=kaj=kʰe   ama=i
      DIST=LOC1=ABL=GEN   1SG=DAT=DIS   mother=ERG

      naʔ=ak=ta=ŋ.
      be.born|give.birth=PST=INV=1

'After that, my mother gave birth to me.'

CH_MKW_BMB_BAN_100617_5_Birth

5.2.3.4.   *no – noʔ- (bʰasa)* & *saj- – saʔj-*

The analyses of the two pairs respectively formed by the noun *no* 'ear' and the verb *noʔ-* 'speak,' and the verb *saj-* 'grow (seed)' (which is also a noun *saj* that means 'seed, fruit' and the verb *saʔj-* 'hear, listen,' go together since they developed through the same derivational process described in the above sub-sections and share together part of the story of their evolution.

The original verb root **no-* may be reconstructed as a stative intransitive verb carrying the meaning of 'be heard, listened to' later replaced by the verb *saʔj-* 'hear, listen.' The reconstructed intransitive root **no-* had the argument structure of a stative intransitive verb, such as {Sx Vst.intr}.

The verb *noʔ-* 'speak' is intransitive and its argument structure is {S Vintr}. It is illustrated in (632). It is an intransitive verb that can occur with a noun, such as *bʰasa*

'language,' result of a construction that historically had the following anticausative intransitive argument structure: {SP*x* Vst.intr}.

Given the properties of the derivational anticausative/causative morpheme *-ʔ*, the absence of transitive argument structure, and the reconstructed form *no-* 'x be heard,' the reconstructible transitive causative argument structure of *no-ʔ*, whose construction is not used synchronically, was {A*y* P*x* Vtr} with the meaning of '*y* cause *x* to be heard,' and that of the intransitive anticausative argument structure of *no-ʔ* was {SP*x* Vst.intr} with the meaning of '*x* be caused to be heard.' In such intransitive anticausative construction this is the thing heard or speech that is the referent of SP or *x* and not the speaker. This explains also why the verb *noʔ-* 'speak' allows the presence of a noun in its synchronic argument structure while the verb remains intransitive. The verb 'speak' thus developed from a form meaning '*x* (language, speech, voice) caused to be heard.'

Further evidence for the argument structure attested with the verb *noʔ-* is shown in (633), where the speaker gives his agreement to share the recordings; he uses the verb *noʔ-* 'speak' derived with the productive causative morpheme *=tak* with 'recordings' as a referent of the T argument of the derived ditransitive causative form. The meaning of *noʔ=tak-* is not 'make the recordings speak,' but rather 'make the recordings caused to be heard, listened to.'

(631) *o*   *dʰanteʌro*   *waŋ=ti*   *saj=na,*   *saj=na,*
       DIST   corn.tassel   come=SEQ1   grow=NPST   grow=NPST

'That corn tassel comes out and grows, grows,'
CH_MKW_BMB_BAN_100717_4_Corn

(632) *ʌm,*   *o=haŋ=sʌj*   *ama*   *noʔ=a,*
       yes   DIST=LOC1=ABL   mother   speak=PST

Chepang bʰasa,   *ba*   *noʔ=a,*   *Thakuri,*
Chepang language   father speak=PST   Thakuri/Nepali

'Yes, from there, my mother spoke, the Chepang language, and my father spoke, Thakuri/Nepali.'
CH_CTW_BBC_POL_102520_1_Polkim

(633) ʌ,   noʔ=tak=ja=ma   sjaw=na,
       uh,   speak=CAUS=COND=ADD   become=NPST

       doh=ma   dʌj=lʌ !
       what=ADD   happen=NEG

       'Uh, even if you make (the recordings) caused to be heard it's fine, that's
       nothing!'
       CH_CTW_JBC_BHR_111720_3_Agreement

The verb *saʔj*- 'hear, listen' comes from the derivation of the intransitive verb *saj*- 'be grown, grow (seed),' such as, **saj-ʔ* whose meaning was '*x* be caused to grow, *y* cause *x* to grow.' The verb **saj-ʔ* was and *saʔj*- is still an anticausative and transitive causative verb since their argument structures are {SP*x* Vintr} and {A*y* P*x* Vtr}. Like with the original argument structure of *noʔ*- 'speak,' the referent of SP or *x* was not the speaker but the speech. The difference between **saj-ʔ* and *saʔj*- is now semantic since *saʔj*- metaphorically developed the meaning of 'hear, listen,' from the idea of causing something said to grow like a seed in somebody's mind, i.e., '*x* (speech) caused to be heard/grown (in sb's mind); *y* cause *x* (speech) to be heard/grown (in sb's mind).'

The verb *saj*- 'grow' is illustrated in (631) and *saʔj*- 'hear, listen,' in (632) to (636).

(634) mel=ko   sahu   si=a   bʰʌnne   dahj=ti
       mill=GEN   owner   die=PST   say.NMZ   say=SEQ1

       kuɾa   saʔj=o   mu=na,
       thing   be.heard|hear=PERF   COP=NPST

       o   kaɾʌni   dah=ma=tʌ   mu=ti,
       dist   because   reach=NEG=NMZ:ADV1   COP=SEQ1

       'The mill's owner died, the thing is heard, that's why I didn't go and stayed,'
       CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

(635)   - e,   ʌni   no   saʔj=na=u   pe=to=le ?
        EXPR   SO   ear   hear=NPST=3O/DIR   be.nice=NMZ:ADV2=DIS

        - no=paj   saʔj=u=lu,   saʔj=u=lu !
        ear=DIS   hear=3O/DIR=NEG   hear=3O/DIR=NEG

'- I see (after the interlocutor said she didn't see well), then the ear hears well?

- The ear either doesn't hear, doesn't hear!'

CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana

(636)   ʌŋ,   didi-bʌhini,   dadʑu-bʰaj=lʌm=i,
        yes   elder.sister-younger.sister   elder.brother-younger.brother=PL=ERG

        sʌbʌj=le,   saʔj=pa=n=i,
        all=DIS   hear=OPT=DIR/TR=PL

'Yes, sisters and brothers, may they all listen (to the story),'

CH_CTW_BBC_POL_111720_1_Cing_Lan

5.2.4.   Derivational consonant *-*k*

The derivational consonant *-*k* is a valency increasing device that was used to derive intransitive and transitive verbs. While the valence is semantically increasing with the addition of a new participant, the causer, this latter is not expressed syntactically.

When attached to an intransitive verb, the derivation results in an intransitive anticausative verb, such as:

(1) The S argument of the intransitive verb becomes the P argument of the derived transitive causative verb while the verb remains intransitive, and no A argument is expressed. This type of derivation is underlyingly transitive causative and P-labile since the resulting construction is an anticausative one. This causative derivation is indirect since the causer is not expressed syntactically.

This intransitive anticausative derivation may be translated as follows:

'*x* is caused to VERB (by *y*)'

This intransitive anticausative derivation can be schematized as follows:

$\{Sx \text{ Vintr}\} > \{SPx \text{ Vintr (by } y)\}$

This is illustrated in the following pair and two triplets:

root	type	meaning	reconstruction
ga-	intr	'open mouth'	**ga-* 'x open mouth'
gak-	intr	'laugh out loud'	*ga-k 'x caused to open mouth (by y)'
na-	intr	'be, be many'	**na-* 'x is'
naʔ-	tr	'give birth'	*na-ʔ 'y cause x to be'
	intr	'be born'	*na-ʔ 'x be born/caused to be (by y)'
nak-	*intr>tr	'hold a baby in arms'	*na-k 'x caused to be present (by y)'
**nu-*	st.intr	'be hidden'	**nu-* 'y be hidden'
nus-	tr	'hide sth'	*nu-s 'y cause x to be hidden'
nuk-~nukʌ	intr	'hide self'	*nu-k 'x caused to be hidden (by y)'

When attached to a transitive verb, the derivation results in either a transitive causative verb (1) or in an intransitive anticausative construction (2), such as:

(1) The P argument of the transitive verb remains the P argument of the derived transitive causative verb, while a new A argument is expressed.

This transitive causative derivation may be translated as follows:

'y causes x to be VERBed'

This transitive causative derivation can be schematized as follows:

$\{Ay \text{ P}x \text{ Vtr}\} > \{Ay \text{ P}x \text{ Vtr}\}$

(2) The P argument of the transitive verb becomes the S argument of the derived transitive causative; the underlying A argument is not expressed.

This intransitive anticausative derivation may be translated as follows:

'x caused to be VERBed'

This intransitive anticausative derivation can be schematized as follows:

$$\{Ay\ Px\ vtr\} > \{SPx\ vintr\}$$

This is illustrated in the following three triplets:

root	type	meaning	reconstruction
ja-	tr	'filter liquid'	**ja-* 'y settle x'
	tr	'block, settle non-liquid parts'	**ja-* 'y settle x'
jah-	intr	'be night, set (sun)'	**ja-h* 'y make self be settled'
jak-	tr	'catch, stop, prevent from falling'	**ja-**k*** 'y caused x to be settled'
	intr	'settle, stop in location'	**ja-**k*** 'x caused to be settled (by y)'
		'be stuck (food in throat)'	**ja-**k*** 'x caused to be settled (by y)'
*lʰu-	tr	'pour out (fertilizer, dirt, sand)'	**lʰu-* 'y pour out x'
lʰuh-	intr	'be poured out'	**lʰu-h* 'x make self be poured out'
lʰuk-	tr	'powder sb'	**lʰu-**k*** 'y caused x to be poured out on'
sɾa-	tr	'sow, scatter (grains, seeds)'	**sɾa-* 'y scatter x'
sɾah-	intr	'be mixed together (grains, seeds)'	**sɾa-h* 'y make self be scattered'
		'be colorful (grains, seeds)'	
sɾak-	tr	'spray, splash, sprinkle on'	**sɾa-**k*** 'y cause x to be scattered on'

5.2.5. Derivational consonant *-t*

The derivational consonant *-t* is a valency increasing device that was used to derive transitive verbs. The derivation results in a transitive causative verb, such as:

(1) The P argument of the transitive verb becomes the A (A1) argument of the derived transitive causative verb. The A of the transitive verb becomes the P (P1) of the derived transitive causative verb which underlyingly functions as an A (A2) argument whose P (P2) argument is the A (A1) argument. The underlying P (P2) argument is not overtly expressed.

517

However, A and P of the predicate of cause are inversely P and A of the predicate of effect (noted between [ ]).

This transitive causative derivation may be translated as follows:

'x caused y to VERB (x)'

This transitive causative derivation can be schematized as follows:

$\{Ay\ Px\ vtr\} > \{Ay\ PAx\ [Py]\ vtr\}$

This is illustrated in the following two triplets:

root	type	meaning	reconstruction
la-	tr	'seize, grab sth'	**la-* 'y grab x'
laʔ- (ti)	intr	'swim'	**la-ʔ* 'y causes x for (y) to be grabbed'
			> **la-ʔ* 'y be grabbed'
lat-	tr	'carry on arms'	**la-t* 'x causes y to grab x' (< **x*=child)
go- (siŋ)	tr	'look for, collect (wood)'	**go-* 'y look for x'
goʔ-	(in)tr	'call sb's attention'	**go-ʔ* 'y causes x for (y) to be looked for'
		'chirp' (bird)	> **go-ʔ* 'y be looked for'
got-	tr	'call sb to come over'	**go-t* 'x causes y to look for x'

5.2.6. Summary

The historical analysis of Chepang old derivational traces is important at three levels. First, it allows us to better understand the synchronic argument structures of such verbs and the morphosyntactical behavior of the arguments and their historical development along with the semantic changes that may have occurred.

Second, this sheds light on the fact that P labiality or anticausative structures are present in the language. Such constructions, which are historically transitive but express an SP argument that holds the function of a P argument are synchronically difficult to analyze without the knowledge of the existence and origin of such P labile constructions; indeed, since Chepang does not require the presence of the A argument when

contextually recoverable, such constructions may be analyzed as transitive whose argument indexation is not overtly expressed. If such an argument is analyzed as a P rather than an SP, it entails the unsubstantiated hypothesis that Chepang is losing argument indexation. At the same time this shows that such a setting may indeed meet all the required morphosyntactic conditions for such system to actually lose argument indexation, as it happened in other TH languages.

Finally, this analysis allows to provide reconstructions of the proto-forms at the lower level of PCB, which then can be compared to the proto-forms of higher-level clades.

This analysis is far from complete and needs to be developed in comparison to the verbal roots attested in other closely related TH languages and beyond.

In Table 134, I summarize the type of verbal constructions derived by the reconstructed series of derivational consonants *-h, *-s, *-ʔ, *-k, *-t.

Table 134. Reconstructed series of derivational consonants *-h, *-s, *-ʔ, *-k, *-t

der.	verb	argument structure	derived verbal construction		argument structure
*-h	tr	{Ay Px Vtr}	intransitive	middle causative	{S [Ay Py] Vintr}
*-s	intr	{Sx Vst.intr}	transitive	causative	{Ay Px Vtr}
	tr	{Ay Px Vtr LOCz}	ditransitive	caus. applicative	{Ay Tx Rz Vditr}
*-ʔ	intr	{Sx Vintr}	transitive	causative	{Ay Px Vtr}
	intr	{Sx Vintr}	intransitive	anticausative	{SPx Vintr}
	tr	{Ay Px Vtr}	transitive	causative	{Ay Px Vtr (by z)}
	tr	{Ay Px Vtr}	intransitive	anticausative	{SPx Vintr (by z)}
	< *mid	{Ay Px Vtr}	transitive	causative	{Ay Px [Ax Py] Vtr}
	< *mid	{Ay Px Vtr}	intransitive	anticausative	{SPy Vintr}
*-k	intr	{Sx Vintr}	intransitive	anticausative	{SPx Vintr (by y)}
	tr	{Ay Px Vtr}	transitive	causative	{Ay Px Vtr}
	tr	{Ay Px Vtr}	intransitive	anticausative	{SPx Vintr}
*-t	tr	{Ay Px Vtr}	transitive	causative	{Ay PAx [Py] Vtr}

5.3. Verbs ending in *-kʌ and *-sʌ

A few verb roots are found with irregular dissyllabic structure whose second syllable is /kʌ/ or /sʌ/. These forms are likely allomorphs of the old derivational consonants *-k and *-s (§ 5.2). The attested examples are scarce but allow some significant observations that lead towards an explanation for their origin.

The form *-kʌ is primarily found in intransitive constructions where the referent of the S argument is affected by the process. This suggests that such verbs may historically be derived intransitive causative verbs from a transitive verbs and that *-kʌ and *-k are allomorphs (§ 5.2.4).

Some verb roots that feature a derivational consonant *-k show a synchronic allomorph ending in -kʌ, as with nukʌ-~nuk- 'hide self' and rʌkʌ-~rʌk- 'prise off (from under shell),' or else grekʌ-~grjak-~grek- 'press on, pound, crush, smash (condiments or cereals having a peel like rice, millet, black pepper, sesame, cardamon)' (RAP-13).

In addition, *-kʌ is not only found following a vowel segment, but also consonants, as with ʌskʌ- 'be so irritated that one can't stand it anymore (noise, people's behavior),' tɕʰiʔnkʌ-(RAP-13, RAK-6)~tɕʰeŋkʌ-(RAK-6) 'separate the rubbish from flour, by tilting downwards and chaking the winnowing tray,' hʌŋkʌ- 'sting, be inflamed, sore (wound),' and dzʰʌskʌ- 'be started, distressed.' This suggests that the allomorphy may be the result of a morphophonological change at morpheme boundary, from *-k to *-kʌ, when following a consonant segment rather than a vowel.

If this is the case, an explanation is still needed for the presence of *-kʌ after vowel segments, as with drʌkʌ- 'trickle (water),' hikʌ- 'grind (teeth),' jokʌ- 'hang sth on, hang self on,' rukʌ-(RAP-13, RAK-6)~lukʌ-(RAP-13) 'rinse one's mouth,' sokʌ- 'caress/slightly massage where sb's got hurt,' takʌ- 'prevent, obstruct, stop, delineate, to stick (nail)' (RAP-13).

Only one example of verb featuring *-sʌ is found and this is kʌjsʌ- 'be pissed off, take things personally, badly.' This form could be derived from the verb kaj- 'quarrel, argue with, get worked up,' given their similar semantics.

The verbs formed with *-kʌ and *-sʌ are reported in Table 135.

Table 135. Verbal roots ending in *-kʌ and *-sʌ

root	meaning
drʌkʌ-	'trickle (water)'
dʑʰʌskʌ-	'be startled, disturbed'
gɾjak-~gɾek-~gɾekʌ- (RAP-13)	'press on, pound, crush, smash (condiments or cereals having a peel like rice, millet, black pepper, sesame, cardamon)'
hikʌ-	'grind (teeth)'
hʌŋkʌ-	'sting, be inflamed, sore (wound)'
jokʌ-	'hang sth on, to hang self on'
nukʌ-~nuk-	'hide self'
ɾukʌ- (RAK-6, RAP-13) ~lukʌ- (RAP-13)	'rinse one's mouth'
ɾʌkʌ-~ɾʌk-	'prise off (from under shell)'
sokʌ-	'caress/slightly massage where sb's got hurt'
takʌ- (RAP-13)	'prevent, obstruct, stop, delineate, to stick (a nail)'
tɕʰiʔnkʌ- (RAP-13) tɕʰeŋkʌ- (RAK-6)	'separate the rubbish from flour, by tilting downwards and chaking the winnowing tray'
ʌskʌ-	'be so irritated that one can't stand it anymore (noise, people's behavior)'
kʌjsʌ-	'be pissed off, take things personally, badly.'

5.4. Stative intransitive vs. dynamic intransitive verbs

Stative intransitive and dynamic intransitive verbs are distinguished through their access to predication. A stative intransitive verb may occur as adjectival predicate in combination with the presence of a copula. In such a construction, the stative intransitive verb is encliticized with the nominalizer morpheme =to. This is illustrated in (637) and (638). As shown in (638), adjectivals may also modify another adjectival, which underlies their ability to also function as an adverbial. Stative intransitives may also modify a noun through the presence of the nominalizer =o, as shown in (639).

Dynamic intransitive verbs can function as predicate without the presence of a copula, as shown in (640) and (641).

Some stative intransitive verbs can however also function as predicate when their process is expressed as dynamic, that a change occurs from one state to another. This is illustrated in (642).

(637) *baŋ=ma*   *du=to*   *du=to*
stone=ADD   be.red=NMZ:ADV2   be.red=NMZ:ADV2

wiʔ=l=o   *tahŋ=to*   *sjaw=na*   *ni !*
blood=COP=NMZ:REL be.huge_be.like=NMZ:ADV2 become=NPST PART

'The stone becomes red, so red, like blood!'
CH_MKW_JMC_SIL_120619_1_Barbalyak

(638) *Khora*   *badʑe,*   *tuhuɾa-ɾadʑa=ko*   *kʌtʰa*
Khora   elder.M   orphan-king=GEN   story

pe=ma=to   *lok=to*   *mu=na*   *i=paj !*
be.good=NEG=NMZ:ADV2 be.far=NMZ:ADV2   COP=NPST   PROX=DIS

'(It was told by) the elder Khora, the story of the orphan king, it is badly long this one!'
CH_MKW_SC_SIL_122619_4_Song

(639) *ʌ,*   *jat-dʑo*   *bʰam=o*   *kʌtu* *kʰe=to,*
uh   one-CL   be.white=NMZ:REL   short COP=REM.PST

ʌni   *du=o,*   *u*   *kʰe=to,*   *gʌ=t=taŋ,*
and   be.red=NMZ:REL   CAT   COP=REM.PST   INT=MAN=ATT

gʌndʑi   *kʰe=to.*
sweater   COP=REM.PST

'Uh, there was one white short, and a red, what's that, there was a red sweater.'
CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(640) ʌ,　　ʌbʌ　　ʌm=mʌj　　　ʌm=na=i,　　　　ʌm=ma=mʌj
　　　uh　　now　　heal=NMZ:H　　heal=NPST=PL　　heal=NEG=NMZ:H

　　　ʌm=i=li,　　　o=le　　　hʌ.
　　　heal=PL=NEG　DIST=DIS　　COP

　　　'Uh, now, those who heal heal, those who don't heal don't heal, that's what it is.'
　　　CH_CTW_SMC_JIM_101920_1_Being_Shaman

(641) sjahŋ　　　　ŋa　　dinbʰʌɾi　　eʔn=na=ŋ !
　　　tomorrow　　1SG　all.day　　sleep=NPST=1

　　　'Tomorrow, I'll sleep all day!'
　　　CH_MKW_STC_SIL_120619_1_E

(642) "kliʔ=le　　　dadz=ʌ=ti　　　kʌn=dzu,
　　　shit=DIS　　compare=LN=SEQ1　show=2DU.IMP.TR

　　　ʌ　　kliʔ　　dadz=ʌ=to=haŋ　　　　　　　su=ko　　　　bʰam=na
　　　uh　　shit　　compare=LN=NMZ:ADV2=LOC1　who=GEN　　be.white=NPST

　　　su=ko　　　jaɾ=na,"　　　　ɾja=ti,
　　　who=GEN　　be.yellow=NPST　say=SEQ1

　　　'"You two show your shits having compared them, uh, when comparing them,
　　　(let's see) whose whitens, whose becomes yellow," he said,'
　　　CH_CTW_BBC_POL_102420_3_Chepang_Kings

(643) duŋ=a　　　　pʌtɕʰi,　　ʌni　　aw=kaj　　god=ʌ=ti,
　　　plant=NMZ2　　after　　then　　DIST=DAT　scrape.around=LN=SEQ1

　　　mʌl　　　　ka=ti,　　　ʌni　　bɾawna,
　　　fertilizer　pour=SEQ1　then　　be.big=NPST

　　　'After having been planted, then having scraped at the foot, poured fertilizer, then
　　　it becomes big,'
　　　CH_MKW_BMB_BAN_100717_4_Corn

5.5. Native and borrowed verbs

The main source of borrowing for the renewal of Chepang verbal roots is Nepali. When a verb is borrowed from Nepali, the enclitic morpheme =ʌ attaches to the borrowed verb form. This is a productive morphological process. This is illustrated in (644) to (646).

(644) *tɕʰak-ʥaʔ=paj*   *o=paj*   *mi=to*
dead.tiger.spirit=DIS   DIST=DIS   be.small=NMZ:ADV2

      sjaw=ti   *nisk=ʌ=na,*
      become=SEQ1   come.out=LN=NPST

      'The tiger spirit of the dead, it comes out having become small,'
      CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'

(645) *bit=a*   *pʌtɕʰi*   *abʌ*   *ŋa=pa*
die=NMZ2   after   now   1SG=DIS

      tuhuɾi=le   *sjaw=alaŋ.*
      orphan.F=DIS   become=1.PST

      'After (my mother) passed, let's say, I became an orphan.'
      CH_MKW_KKBP_LPK_101917_3_Chepang_Language

(646) *Praja Mobile House,* *Lothar bʌʥaɾ=haŋ,* *kʰol=ʌ=alaŋ,*   *ŋa=i*
Praja Mobile House   Lothar town=LOC1   open=LN=1.PST   1SG=ERG

      'I opened the Praja Mobile House in the town of Lothar.'
      CH_MKW_SKP_DAM_112819_Conversation_with_Bipana

When a verb is transitive causative, which in Nepali is characterized by the presence of a final form <āu>, the morpheme =ʌ does not occur and the Nepali ending <āu> is preserved, as in (647).

(647) *tɕoʔ=kaj*   *pʌdaw=sa*   *lag=ʌ=alaŋ,*
child=DAT   study=NMZ1   apply=LN=1.pst

'I made my children study,'

CH_MKW_SC_SIL_010220_1_Life

5.6. Productive derivational morphology

Two productive derivational morphemes are attested in Chepang: the causative morpheme *=tak* (§ 5.6.1) and the emotional anti-experiencer morpheme *=se~si~sje~sja* (§ 5.6.2). While both morphemes are productive, the emotional anti-experiencer marker occurs less frequently than the causative. In our corpus, around 200 examples of verbs derived with *=tak* are found, against 10 examples for *=se~si~sje~sja*.

5.6.1. Causative *=tak*

The causative morpheme *=tak* is a valency increasing device. The resulting derivation can be an intransitive anticausative, or a transitive or ditransitive causative construction.

When the causative morpheme *=tak* attaches to an intransitive verb, the S argument of the intransitive verb becomes the P argument of the transitive causative verb while an additional A argument occurs. In the derived construction, the A holds the semantic role of causer and the P that of causee. This is illustrated in (648). The causative morpheme *=tak* can also occur in imperative construction, as shown in (649), expressing a request of permission.

This transitive causative derivation may be translated as follows:

'*y* causes *x* to VERB'

This transitive causative derivation can be schematized as follows:

{S*x* Vintr} > {A*y* P*x* Vtr}

(648) *(...) biɾko=kʰe*   *saʔ=taŋ*   *klaʔ=tak=ka=nʌ.*
     cap=DIS   ground=ALL   fall=CAUS=2/3.PST=DIR/TR

'(The caterpillar) made the cap fall on the ground.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(649) *ama*   *tuk=haŋ*   *dah=tak=tɕi !*
     mother   belly=LOC1   reach=CAUS=2SG>1SG

e,   *baba*   *tuk=haŋ*   *dah=tak=tɕi !*

EXPR   father   belly=LOC1   reach=CAUS=2SG>1SG

'Let me reach my mother's wombs, let me reach my father's wombs!'

CH_CTW_JBC_BHR_102420_1_Cing_Lan

When the morpheme =*tak* attaches to a transitive verb, the derived verb becomes ditransitive causative, adding a new participant holding the function of A argument of the caused event and semantic role of causer. The P argument of the transitive verb becomes the T argument of the ditransitive causative, and the P argument of the ditransitive construction underlyingly holds the function of an A (A2) argument. This is illustrated in (650).

This ditransitive causative derivation may be translated as follows:

'*y* causes *z* to VERB *x*'

This ditransitive causative derivation can be schematized as follows:

{A*y* P*z* Vtr} > {A*y* PA*z* T*x* Vtr}

(650) *hela*   *dʑahŋ=na=u,*   *ʌ*   *purʌj*   *gʰas*   *siŋ*
     reprimand   do_make=NPST=3O/DIR   uh   completely   grass   wood

tat=tak=sa,   *bʌn=kʰa*   *garʰo*   *garʰo*

cut=CAUS=NMZ1   jungle=LOC2   hard   hard

kam   *dʑahŋ=tak=sa,*

work   do_make=CAUS=NMZ1

'(The mother-in-law) reprimands her, with no limit, making her cut grass and wood, making her work hard hard in the jungle,'
CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

The derivation of a transitive verb with the morpheme =*tak* may as well entail another type of construction, where the verb does not become ditransitive but transitive anticausative. This is illustrated in (651) in (652).

In such transitive anticausative derivation, while a new A argument is added, the A argument of the transitive verb is not expressed in the derived transitive anticausative verb. The P of the transitive verb remains the P of the derived transitive anticausative verb but still holds the underlying function of SP argument of the transitive verb in its P labile construction. Here, the derived verb is anticausative with regard to the A of the original transitive verb, since it is not expressed.

This transitive anticausative derivation may be translated as follows:

'*y* causes *x* to be VERBed (by *z*)'

This transitive anticausative derivation can be schematized as follows:

{A*z* P*x* vtr} > {A*y* (A*z*) Psx vtr}

(651) *ʤjaɾ*   *djaw=tak=a=ta=ŋ,*   *ʤjaɾ*   *djaw=ti*
       yam.*sp*   dig=CAUS=PST=INV=1   yam   dig=SEQ1

       ɾekʌt   *ʤahŋ=tak=a=ta=ŋ*   *ŋa=kaj,*
       record   do_make=CAUS=PST=INV=1   1SG=DAT

       glih   *djaw=ti*   *ɾekʌt* *ʤahŋ=tak=a=ta=ŋ,*
       yam.*sp*   DIG=SEQ1   record do_make=CAUS=PST=INV=1

'They made me dig up yam (*Dioscorea*), while digging them up, I got recorded, while digging up yam (*Dioscorea kamoonensis kunth*), I got recorded.'
CH_MKW_SC_SIL_010220_1_Life

(652) didi=ma   bʌlʌ   no   saʔj=u=lu=le   hʌ,
     elder.sister=ADD little   ear   hear=3O/DIR=NEG=DIS   COP

     didi=kaj=te   dʑʌjk=tak=te=tɕja=n=i.
     elder.sister=DAT=2   bite=CAUS=2=IRR=DIR/TR=PL

     'The thing is also, elder sister doesn't hear well, (they) would have made you

     being bitten, elder sister (by rhinoceroses).'
     CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

The derivational causative marker can also occur with borrowed verb, whether it
is an intransitive verb, as in (653).

(653) (...) dahj=ti,   al=ti,   Bharatpur   dʑʌnm=ʌ=tak=a,
     say=SEQ1   go=SEQ1   Bharatpur   be.born=LN=CAUS=PST

     '(…) they said and went, and he got given birth in Bharatpur.'
     CH_MKW_SC_SIL_010220_1_Life

It is likely that the morpheme =tak historically comes an auxiliary construction
formed with the verb tak- 'permit, let happen' before grammaticalizing into a causative
derivational morpheme in a serial verb construction. The verb tak- 'permit, let happen' is
illustrated in (654) and (655).

(654) gʌj=da=ko=paj   kuɾa=le   dʑahŋ=jak=n=i=li,
     rhino=GEN=DIS   thing=DIS   do_make=REM.PST=DIR/TR=PL=NEG

     kuɾa   dʑahŋ=sa   tak=jak=lʌ.
     thing   do_make=NMZ1   permit=REM.PST=NEG

     'We never made anything from rhinoceroses, doing this type of thing was not

     permitted.'
     CH_MKW_SCBKC_SIL_081918_5_Hunting

(655) *deŋe   goli   ap=sa   tak=ka=n=i,*
       now   bullet shoot=NMZ1   permit=2/3.PST=DIR/TR=PL

       bʌnduk   ap=sa   tak=ka=n=i,
       gun   shoot=NMZ1   permit=2/3.PST=DIR/TR=PL

'Now they let people shoot with bullets, they let people shoot with guns.'
CH_MKW_SCBKC_SIL_081918_5_Hunting

5.6.2.   Emotional anti-experiencer =se~si~sje~sja

The emotional anti-experiencer derivational construction formed with the morpheme *=se~si~sje~sja* is only found in an adjectival predication in combination with a copula and the encliticization of the nominalizer *=to*, as illustrated in (656) to (658). Its use seems restricted to verbs that convey emotional expressions; it can attach to transitive verbs like *ɾap-* 'like,' *ʣe=gaɾ-* 'desire eating' or experiencer verbs like *ɾʌj-* 'fear, be scared'. The P or Stimulus argument of the original verb, both marked with the dative case, becomes an S argument, i.e., an S argument that holds the semantic role of a Stimulus; the A or S experiencer argument of the original verb is not expressed, identifiable or specific. The process expressed is not realized in the sense that it merely expresses the idea of it.

(656) *tʰo=ti=ja,   manta=i   sat=na=n=i,*
       beat=SEQ1=COND   person=ERG   kill=NPST=DIR/TR=PL

       ɾʌj=sje=to   sjaw=na   nitʌ !
       fear=ANTIEXP=NMZ:ADV2   become=NPST   PART

'What if (they) beat (the shaman), people kill (other people), it's definitely scary!'
CH_CTW_BBC_PID_011520_1_Being_Shaman

(657) *ŋa   ɾap=se=to   mu=na=ŋ.*
       1SG   love=ANTIEXP=NMZ:ADV2   COP=NPST=1

'I'm lovable.'
CH_CTW _KC_GUN _071220_E_30_2

(658) i   tak-saj   ʥe=gaɾ=sje=to   mu=na.
       PROX   mango-fruit   eat=DES=ANTIEXP=NMZ:ADV2COP=NPST

'This mango makes you feel like you want to eat it.'

alt. 'This mango is eating-desirable.'

CH_MKW _PC_SIL_ E_32_8

It is possible that this emotional anti-experiencer construction may have developed from a serial verb or nominalized construction. Indeed, the verb *se-* 'feel' is still used synchronically and is likely the source for the development of the morpheme *=se~si~sje~sja*, as shown in (659). It also could have evolved in a nominalized construction, which is also attested, as illustrated in (660), before the construction be used without the nominalizer.

(659) ʨoʔ=lʌm   dukʰʌ   se~sje=na=i.
       child=PL   pain   feel=NPST=PL

'Children feel pain.'

CH_CTW _KC_GUN _071220_E_29_5

(660) - o=tʌ   jah-diʔŋ   ʨiŋ=tak=ti   gʌm=ja=paj
         DIST=NMZ:ADV1   night   stand.up=CAUS=SEQ1 keep=COND=DIS

       ɾʌj=a=se=to   ma=ba   didi !
       fear=NMZ2=feel/ANTIEX=NMZ:ADV2 PART=PART   elder.sister

       - hahah   ma=ba=ane !
       hahah   PART=PART=PART

'- If we stay like that up all night, it's definitely scary, don't you think, elder sister?
- (laugh) yeah, for sure!'

CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

5.7. Verbal template

The verbal template presented in Table 136 summarizes syntagmatic and paradigmatic distribution of verbal morphology that occurs around the root. The root is noted by the symbol $\sum$.

Fourteen verbal derivation and inflectional enclitics are attested after the verbal root (SLOT 0), while only two morphemes are found preceding it. Both the enclitic 2nd person morpheme =te (SLOT -1) and the enclitic negation morpheme =ma can precede (SLOT -2) the root and follow it (SLOT 6 and 8 respectively) or do both at once, occurring twice.

In the following sub-sections, I give a brief overview of the kind of morphemes found in each slot along with examples.

Table 136. Chepang verbal template and morphology

-2	-1	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
{=NEG}	{=2ND}	$\sum$	AUX		EPIS	ANT	EPIS	{NEG}	RECIP	{2ND}	TENSE	INVERSE	1ST	NUM	REFL	NEG
			MOD			REM					IRR		1INCL	S/A/P		REM PAST
			ASP			PAST					OPT		INV_3>1DU			IMPER NEG
			-----	-------									2>1SG			
			MOD	DERIV												

5.7.1. SLOT -1 / SLOT 8: 2nd person, and 1>2 person

SLOTS -1 and 8 are dedicated to indexing 2nd person by the 2nd person morpheme =te. In both slots, it occurs as an enclitic, that is, even when preceding the verbal root, it attaches to the first morpheme just before it, as shown in (661) for SLOT -1 and (662) for SLOT 8. When the 2nd person morpheme =te occurs in SLOT -1 it frequently also occurs in SLOT 8, as in (663). When the 2nd person morpheme =te occurs in SLOT -1, nothing occurs between it and the verbal root. As described in (§ 5.7.2) the enclitic negation morpheme =ma also attaches to the first morpheme immediately preceding it. As shown in (664), the position of the 2nd person marker =te may fall between the negation marker =ma and the verbal root. Like with the negation morpheme =ma, the 2nd person morpheme =te was historically procliticizing to the root.

SLOT 8 also hosts another morpheme dedicated to the marking of 1st person acting on 2nd person (1>2), the morpheme =ʨe~ne, as shown in (665).

(661) *naŋ*   *rʌjsʌ,*   *ŋa=kaj*   *ŋa=ko*
      2SG   COP.MIR   1SG=DAT   1SG=GEN

       ahm=te   *waj=a=ka=n !*
       rice=2   throw=PST=2/3.PST=DIR/TR

       'You're like that, you threw my rice at my face!'
       CH_MKW_DBC_MAI_1_020320_The two brothers

(662) *ŋa=ko*   *ʨoʔ* *mak=te=ka=n,*
      1SG=GEN   child devour=2=2/3.PST=DIR/TR

       'You devoured my child,'
       CH_CTW_JMC_PYK_101920_Cing_Lan

(663) *niŋ=paj*   *bani=te*   *sjaw=te=ka=i,*
      2PL=DIS   habit=2   become=2=2/3.PST=PL

       ŋa=paj   *rʌj=a=na=ŋ.*
       1SG=DIS   be.scared=EPIS=NPST=1

       'You all got used to that, me, I'm definetly scared!'
       CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(664) *naŋ=i=paj* *ten=paj*   *doh=ma=te* *ʥahŋ=u=lu*
      2SG=ERG=DIS today=DIS   what=NEG=2 do_make=3O/DIR=NEG

       da !
       PART

       'You didn't do nothing today!'
       CH_MKW_STC_SIL_120619_2_E_2

(665) *lʌ*   *bʰandʑa,*   *ŋa=i*   *ahm*   *kas=tɕe=na=ŋ* !
       well   nephew   1SG=ERG   rice   feed=1>2=NPST=1

'Okay nephew, I'm going to feed you rice!'

CH_CTW_JBC_BHR_102420_1_Cing_Lan

5.7.2. SLOT -2 / SLOT 6: Negation

SLOT -2 and SLOT 6 are dedicated to the position of the negation morpheme =*ma*, which in all cases is an enclitic.

This is historically the trace of a verbal proclitic which has cognates in other TH languages. It can be reconstructed as **ma=* at least back to the ancestor of Chepang-Bhujel and Magar, which features a cognate prefix negation morpheme *ma-* (Grunow-Hårsta 2008: 189).

The morpheme can occur once in either slot -2 or slot 6, as in (666) and (667) respectively, or occur twice, as in (668). The negation morpheme =*ma* also developed an additive function meaning 'not even' with negation and 'also, too' in positive constructions, as in (668) and (669).

(666) *ʌɾka=kʌ*   *pahj=sa=ma*   *aʔl=waj=ta=ŋ=lʌ,*
       other=NMZ:LOC   leave=NMZ1=NEG   take.away=CERT=INV=1=NEG

'I won't (even) be taken away to leave to another's.'

alt. 'I won't get married with another man.'

CH_MKW_SC_SIL_010220_1_Life

(667) *o*   *karʌn=sʌj*   *ŋa=i=kʰe,*   *pʌtɕʰi* *dekʰi*
       DIST   reason=ABL   1SG=ERG=DIS   after   from

       pʌjsa   *lʰok=ma=ŋʌ=lʌ.*
       money   send=NEG=1=NEG

'That is why, I, after that I didn't send money.'

CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

(668) *dʑjan=ma*   *lʌj=ma*   *kʰaj=ma=ŋʌ=lʌ,*
       body=ADD   SLF.INTS=NEG   be.able=NEG=1=NEG

'My body also can't by itself,'

alt. 'My body also is not able to do anything.'

CH_MKW_STC_SIL_120619_2_E_1

(669) *o=haŋ=le*   *ŋi=kaj=ma*   *aʔl=nu.*
       DIST=LOC1=DIS   1PL=DAT=ADD   take.away=2PL.IMP.TR

'You all take me away there too.'

CH_MKW_BMB_BAN_103119_8_Jungle

5.7.3. SLOT 1: Auxiliary (modality, aspect)

SLOT 1 is the position of a few auxiliaries that grammaticalized into aspectual and modality markers, as shown in (670) to (672).

(670) ʌ   *Tantula,*   *doh*   *sjaw=a,*   *tɕiʔ=lʌ,*
       uh   Tantula   what   become=PST   know=NEG

       ʌdʑʰʌ   *sʌmmʌ*   *dah=dʰaŋ=lʌ* !
       now   up.to   reach=IMPF.NEG=NEG

'Uh, as for Tantula, what happened, I don't know, she hasn't reached home yet!'

CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

(671) *naŋ=i*   *ŋa=ko*   *tɕoʔ-djaŋ=kaj*   *gʰanti*   *sat=kʰʌj=o.*
       2SG=ERG   1SG=GEN   daughter=DAT   beat=SEQ1   kill=INC=PERF

'You've beaten and started to kill my daughter.'

CH_MKW_BMB_BAN_090118_7_Chepang_marriage

(672) ʌni   o   paɾi=taŋ   kʰola   tʌɾ=ʌ=ti,
      so   DIST   DEIC.OPP=ALL   river   cross=LN=SEQ1

      al=gaɾ=te=na   ja,   al=gaɾ=te=lʌ   ja ?
      go=DES=2=NPST   or   go=DES=2=NEG   or

      'So, crossing the river, do you want to go to the other side or not?'
      CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

5.7.4. SLOT 2: Derivational

Slot 2 is dedicated to the position of two derivational markers, the causative morpheme *=tak* and emotional anti-experiencer morpheme *=se~si~sje~sja*, described in § 5.6. Both morphemes may follow the modality marker *=gaɾ* which occurs in Slot 1, as illustrated in (673) and (674). These two derivational morphemes cannot co-occur, nor be followed by any other aspectual or modality marker of Slot 1.

(673) o=kaj   ʥe=gaɾ=tak=alaŋ.
      DIST=DAT   eat=DES=CAUS=1.PST

      'I made her/him feel like eating.'
      CH_CTW_SPC_POL_ E

(674) lʌj   kʰaŋ=o=paj   ʌʥʌm=gaɾ=se=to
      SLF.INTS   cook=NMZ:REL=DIS   be.disgusting=DES=ANTIEXP=NMZ:ADV2

      kja !
      PART

      '(The lezard monitor) you cook by yourself seems so disgusting!'
      CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

5.7.5. SLOT 3: Epistemic

SLOT 3 hosts the morpheme *=a* which has an epistemic meaning; it expresses the speaker's belief that the process expressed by the verb will happen, as in (675). This

morpheme possibly developed from its original function of nominalizer before becoming a past tense marker.

(675) *"ŋa=paj* *gu=tɕjuk* *sjak=waj=na=ŋ* *ɾʌ,*
    1SG=DIS   INT=QTY   survive=CERT=NPST=1 PART

    djah=paj *ŋa=paj* *dʑe=a=na=ŋ*
    now=DIS   1SG=DIS eat=EPIS=NPST=1

    tɕoʔ=lʌm," *dahj=a,* *ama.*
    child=PL   say=PST mother

'For how much longer am I going to survive? Tell me, now, my dear children, I will definitely eat (the poison),' she said, my mother.'
CH_MKW_SC_SIL_010220_3_Life

5.7.6. SLOT 4: Remote past, anteriority

The morpheme *=jak* occurs in slot 4. It is used as a remote past marker in the negative remote past construction, as shown in (676) to (678). It also carries the meaning of 'before' when used in an imperative construction, as in (679). In either case it marks the anteriority of the process against a temporal deictic center that is the time of speech, with or without the presence of deictic adverbials.

(676) *o* *biswas* *dʑahŋ=sa*   *kʰaj=jak=n=i=li.*
    DIST Christinity do_make=NMZ1 be.able=REM.PST=DIR/TR=PL=NEG

'Before, they would not be able to practice Christianity.'
CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(677) *o-mi*   *lekʰʌ=sa=kʰe* *o* *bela=pa*
    DIST-PL.H write=NMZ1=DIS DIST moment=DIS

    tɕiʔ=ma=jak=n=i=li   *dʌjʰa.*
    know=NEG=REM.PST=DIR/TR=PL=NEG maybe

'They maybe did not how to write at that time.'

CH_MKW_SCBKC_SIL_081918_2_Chepang_king

(678) - *kaka=i=ma*   *ɾap=ta=ŋ=to.*
father's.younger.brother=ERG=ADD love=INV=1=REM.PST

  - *ʌ* *po=jak=te=lʌ*   *ja* *ane,*

  uh beat=REM.PST=2=NEG   or PART

'- My paternal uncle too used to love me,

- Uh, he wouldn't beat you or would he then?'

CH_MKW_SPMC_LC_SIL_100921_3_Conversation

(679) *lʌw,* *sat-bʰaɾi* *siŋ* *go=jak=u,*
   well, seven-load wood collect=ANT=2SG.IMP.TR

  o=ŋ=sʌj   *ŋa=kaj* *ʌ,* *kʰaŋ=ti* *dʑe=ljam,*

  DIST=LOC1=ABL 1SG=DAT uh cook=SEQ1 eat=NEG.IMP

'Well, first collect seven loads of wood, and then having cooked me food, don't eat.'

CH_CTW_KMC_TAP_102520_2_The bat and the crab

5.7.7. SLOT 5: Epistemic

The epistemic morpheme =*dʑe*, which possibly comes from the verb *dʑe*- 'eat,' serves the expression of different emotions that the speaker feels regarding a process portrayed as imperative in the case of an imperative statement, or as non-reversible for which nothing can be done for it to be changed, or else nostalgia.

It asserts more strongly the imperative statement of the speaker, as in (680). It conveys the feeling of helplessness and sadness regarding unfortunate or terrible events, as in (681). It can also convey a feeling of nostalgia regarding past events, as in (682).

(680) *talaŋ   hʌllaw=dʑe=ljam,   hahaha,*
head   shake=EPIS=NEG   (laugh)

'Don't shake your head, hahah (while I tell you this story),'

CH_MKW_SPMC_SIL_100921_1_Tiɾidu'm_Basʌdu'm

(681) *tɕoʔ-tɕoʔ-djaŋ   ɾja=dʑe=na=i   basi   ahm   te=dʑe=na=i,*
son-daughter   cry=EPIS=NPST=PL   stale   rice   beg=EPIS=NPST=PL

'Our children cry, they beg for stale rice,'

CH_CTW_YMC_TAP_102420_2_Song_Maru_sarasara

(682) *naŋ=i   ŋa=kaj   pʌhjla   pe=ma=tʌ*
2SG=ERG   1SG=DAT   before   be.nice=NEG=NMZ:ADV1

git   ke=dʑe=to   ane,
song   sing=EPIS=NMZ:ADV2   PART

'Before, you would sing songs to me so nicely!'

CH_MKW_STC_SIL_120619_2_E_1

5.7.8. SLOT 7: Reciprocal

The reciprocal marker *=kaj* occurs in SLOT 7, as illustrated in (683).

(683) *o=sʌj   palo   palo   ɾap=kaj=na=tɕʌ,*
DIST=ABL   turn   turn   love=RECIP=NPST=1/3DU

'From there, one after another, they make love to each other.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

5.7.9. SLOT 9: Tense, irrealis, optative

Tense markers occur in SLOT 9, as shown in (684) with non-past and in (685) with past. This position is shared with the irrealis, as in (686), and the optative marker, as in (687).

(684) *Prabhu=i*   *naŋ=kaj*   *jo=te=na,*
lord=ERG   2SG=DAT   see=2=NPST

Prabhu=i   *naŋ=kaj*   *bʰʌlo*   *dʑahŋ=te=na.*
lord=ERG   2SG=DAT   good_honest   do_make=2=NPST

'The lord sees you, the lord makes you good.'

CH_MKW_BMB_BAN_103119_1_Church

(685) *tʰik*   *dʑahŋ=te=ka=tɕi*   *dʰʌnjebad,*
good   do_make=2=2/3.PST=2SG>1SG   thank.you

'You did good to me, thank you.'

CH_CTW_KRC_HAT_012120_Being_Shaman

(686) *gʰas*   *tat=laŋ*   *al=o*   *bela=haŋ*
grass   cut=PUR   go=NMZ:REL   moment=LOC1

dʑaʔ=i=te   *dʑʌjk=te=tɕja=i,*
tiger=ERG=2   bite=2=IRR=PL

'When going to cut grass, a tiger would bite you,'

CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(687) *i=ko*   *loʔ*   *gʌ=tɕjuk*   *gaɾ=na=to*
PROX=GEN   leaf   INT=QTY   want=NPST=NMZ:ADV2

waʔn=pa=u !
bring=OPT=3O/DIR

'May she bring as much of her leaves as she wants, (I don't care!)'

CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan

5.7.10. SLOT 10: Inverse

The inverse marker =*ta* ~ *tʌ* ~ *tʰʌ* occurs in slot 10, as shown in (688) and (689) with 3rd person acting on 1st person, and 3rd person acting on 3rd person, respectively.

(688) *"ajʌ   djah   ŋa=kaj   dze=waj=ka=ta=ŋ,"*
      EXPR now   1SG=DAT   eat=CERT=2/3.PST=INV=1

      mʰʌɾ=ti   ɾʌj=a.
      think=SEQ1   be.scared=PST

      'Thinking "(The bear) is gonna definitely eat me now," she was scared.'
      CH_CTW_BBC_POL_111720_5_Two Sisters

(689) *Bhabikan   pʰeɾi, uh   Butwal=haŋ=ma   aʔl=ti*
      Bhabikan   again   CAT   Butwal=LOC1=ADD   take.away=SEQ1

      ka=a=tʰʌ=i   dzel=haŋ=paj.
      put=PST=INV=3>3SG   jail=LOC1=DIS

      'And Bhabikan, there, take him away to Butwal, he was put in jail.'
      CH_MKW_PSC_MAI_012620_1_Becoming_Christian

5.7.11. SLOT 11: 1st person, 1st person dual/plural inclusive, and 2>1 person

SLOT 11 is dedicated to the marking of 1st person, as seen with the 1st person past marker =*alaŋ* in (690), 1st person inclusive dual marker =*tʌhj* in (691), 1st person inclusive plural marker =*tʌh*, as in (692), and 2nd person singular acting on 1st person marker singular =*tɕi*, as in (693). Note that the past form associated with 1st person is =*alaŋ* in all studied varieties and =*akaŋ ~ kaŋ* in RAP-13 (Polkim, Sarling, Syamrang, Yuiling, Santhali), as shown in (694).

(690) *baba   si=ti,   i   tahŋ=o   ɾihŋ=ma*
      father die=SEQ1   PROX be.huge_be.like=NMZ:REL   drum=ADD

      juin=alaŋ   ŋa=i.
      buy=1.PST   1SG=ERG

      'When my father passed, I bought a drum like this one.'
      CH_MKW_BLC_SIL_113019_2_Becoming_Shaman

(691)   *djah*   *ŋi-tɕi=kaj*   *pulis* *aʔl=na=tʌhj=tɕʌ !*
          now   1DU=DAT   police take.away=NPST=3>1DU.INCL/INV=1/3DU

          'Now the police will take the two of us away!'
          CH_MKW_DBC_MAI_1_020320_The two brothers

(692)   *o*   *bʰasa*   *pʰe=aktiko*   *ŋi*   *Nepali bʰasa*
          DIST   language   leave=SEQ2   1PL   Nepali language

          sik=ʌ=sa   *boŋ=na=tʌh=i,*
          learn=LN=NMZ1   look.for=NPST=1PL.INCL=PL

          'Having abandon that language, we look for learning the Nepali language,'
          CH_MKW_DKC_MAI_012720_Chepang_Language

(693)   *lʌw*   *al=ʌ,*   *pahj=ʌ*   *didi,*   *djah=paj*
          well   go=2SG.IMP.INTR   leave=2SG.IMP.INTR   elder.sister   now=DIS

          ŋa=kaj   *srʌjk*   *sat=te=ka=tɕi.*
          1SG=DAT   louse   kill=2=2/3.PST=2SG>1SG

          'Well, go, leave elder sister, now that you killed my lice,'
          CH_CTW_BBC_POL_111720_5_Two Sisters

(694)   *ŋai*   *i=tɕjuk*   *juin-ɾaj*   *lʰak=kaŋ*
          1SG=ERG   PROX=QTY   story   recount=1.PST

          'I told you this many stories,'
          CH_CTW_JBC_BHR_102420_2_Agreement

5.7.12. SLOT 12: Number and directionality

Person number marking occurs in slot 12. This is illustrated in (695) to (697). Number combines with directionality, that is, they may indicate that the person marked for number is acted upon or is acting on another participant. Some morphemes marking number may be parsed, and some may not, being portmanteau. I therefore do not present them as holding different syntagmatic positions.

541

(695) ʌ,   ŋi=kʰe,   juin   sat=laŋ   bʌn=taŋ,
       uh   1PL=DIS   bat   kill=PUR   jungle=ALL

       al=na=ŋ=su,   ja-diʔŋ.
       go=NPST=1=1PL.EXCL   night

       'Uh, we go to the junge to hunt bats, by night.'
       CH_MKW_MRNDC_SIL_081818_5_Bats_Whistle

(696) tɕit-nʌm   niŋ-dʑi   pʰon   dʑahŋ=te=ka=dʑ=u.
       two.days.ago   2DU   phone do_make=2=2/3.PST=2DU=3O/DIR

       'Two days ago, you two called.'
       CH_MKW_SC_SIL_010220_3_Life

(697) dwi   din=ko   lagi   kʰe=ja=kaj   ki
       two   day=GEN   for   COP=COND=DAT   or

       bihibaɾ=taŋ   pahj=i,
       Thursday=ATT   leave=1PL.IMP.INTR

       'If it's for two days, then let's leave Thursday.'
       CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation

5.7.13. SLOT 13: Reflexive

The reflexive marker =sʌ ~ si occurs in SLOT 13. It is illustrated in (698) and (699).

(698) dʑʌmma   dwi   din=taŋ   naj   law=ka=sʌ
       all   two   day=ATT   clothes   apply_wear=2/3.PST=REFL

       suɾwal   ɾʌ   dʌwɾa.
       trouser   and   upper.garment

       'In all, for only two days, (my grandfather) wore the trousers and the coat.'
       CH_MKW_BBC_SIL_032920_2_My grandfather

(699) | *Chepang=lʌm* | *pɾaj* | *suɾu=haŋ=kʰe,* | *laŋgʌte* |
|---|---|---|---|
| Chepang=PL | mainly | beginning=LOC1=DIS | loin.cloth |

law=na=i=si,		*dʰoti*	*law=na=i=si.*
apply_wear=NPST=PL=REFL		loin.cloth	apply_wear=NPST=PL=REFL

'The Chepangs, mainly, at the beginning, they wear the loin cloth.'

CH_MKW_SCBKC_SIL_081918_3_Chepang_culture

5.7.14. SLOT 14: Negation, remote past, imperative negation

The last SLOT 14 hosts the negation morpheme =*lʌ* ~ *li* ~ *lu*, as in (700), the remote past marker =*to*, as in (701), and the negative imperative marker =*ljam* (LOTHAR) ~ *lam* (MANAHARI), as in (702).

(700) | *naŋ=paj* | *no=ma* | *saʔj=te=u=lu.* |
|---|---|---|
| 2SG=DIS | ear=ADD | hear=2=3O/DIR=NEG |

'You don't hear well.'

CH_MKW_BMB_BAN_103119_9_Mother and daughter

(701) | *Chepang* | *kuɾa=sʌj=le* | *hot=ŋʌ=to,* |
|---|---|---|
| Chepang | language=ABL=DIS | question=1=REM.PST |

'I used to ask questions in the Chepang language.'

CH_MKW_GBC_CYO_120119_Conversation_with_Bipana

(702) | *ŋi=kaj* | *pʰe=ljam.* |
|---|---|
| 1PL=DAT | leave=NEG.IMP |

'Don't leave us.'

CH_MKW_BMB_BAN_103119_8_Jungle

5.8. Inflectional morphology

The preceding section examined verbal morphology in terms of structural position relative to the verb root. This section focuses on the description of Chepang verbal inflectional morphology in terms of functional domains. I describe the domains of negation (§ 5.8.1), verbal argument indexation (§ 5.8.2, § 5.8.3), basic tense, aspect and modality (§ 5.8.4).

Chepang exhibits one of the most complex verbal argument indexation systems attested in TH languages. In § 5.8.2, I introduce previous accounts of Chepang verbal argument indexation (§ 5.8.2.1) and show that the attested patterns can be analyzed as forming a non-canonical direct-inverse system, compared to direct-inverse systems found in other TH languages and cross-linguistically (§ 5.8.3.3). I show how the divergent patterns result from historical changes. The argument indexation patterns further sociopragmatic forces rather than a pre-existing referential hierarchy.

In § 5.8.3, I describe in detail the forms used to mark verbal argument indexation which raises the question of the validity of the functionality of such hierarchy (DeLancey 1981; Nichols 1992a; Siewierska 1998; Siewierska 2004; Zuñiga 2006; Lockwood & Macaulay 2012) to describe direct-inverse systems, and illustrates that direct-inverse systems are the result of historical changes in indexation forms (Filimonova 2005; Bickel 2008; Cristofaro 2013; Rose 2015; Gildea & Zúñiga 2016; Arkadiev 2020). I describe another construction considered archaic in all studied varieties, which is formed with the morpheme =*bʌt* and indexes the possessor of a possessed noun-phrase functioning as a benefactive/malefactive (§ 5.8.3.6).

Finally, section § 5.8.4 describes the marking of tense, aspect, and modality. These descriptions are accompanied with examples of intransitive, transitive, ditransitive, middle/reflexive/reciprocal paradigms.

5.8.1. Negation =*ma* and =*lʌ*

Two negation enclitic morphemes are attested on the verb, the morpheme =*lʌ* ~ *li* ~ *lu*, and the morpheme =*ma*.

The morpheme =*lʌ* ~ *li* ~ *lu* developed three allomorphs through regressive assimilation with the vowel of the preceding morpheme, which is either the plural marker =*i* for =*li*, or the direct or object marker =*u* or 1st person plural morpheme =*su* for =*lu*. This negation marker is illustrated in (703) to (704).

It is likely an innovation at the level of PCB, from the grammaticalization of a copula *le* still used in Chepang and in Magar (Grunow-Hårsta 2008). Bhujel features the cognate negation morpheme -*l* (Regmi 2007: 243).

(703) *na=lʌ,*   *Newar na=lʌ,*   *Tamang*   *na=lʌ,*
       COP=NEG   Newar COP=NEG   Tamang   COP=NEG

       'There were no, there were no Newar, there were no Tamangs,
       CH_CTW_BBC_POL_102520_1_Polkim_Archive

(704) *git*   *tɕiʔ=n=i=li,*   *juin-ɾaj*   *tɕiʔ=n=i=li*
       song   know=DIR/TR=PL=NEG   story   know=DIR/TR=PL=NEG

       o-mʌj,
       DIST-PL.H

       'Them, they don't know any songs, they don't know stories.'
       CH_MKW_SC_SIL_010220_3_Life

(705) *ma,*   *naŋ=i*   *tɕiʔ=te=u=lu,*   *o=kʰe.*
       yes   2SG=ERG   know=2=3O/DIR=NEG   DIST=DIS

       'Yeah, you don't know that one.'
       CH_MKW_STC_SIL_120619_4_E_2

The original negation marker was the morpheme =*ma*, reconstructible as *ma= at the proto-level of the common ancestor of Proto-Magar (PM) and Proto-Chepang-Bhujel (PCB), attaching to the left edge of the verbal complex.

It is attested as a prefix across a variety of TH languages, closely related to Chepang, like Magar (Grunow-Hårsta 2008: 189), Kham (Watters 2002: 2003), and beyond; it is reconstructed as *ma (Benedict 1972: 97; Matisoff 2003: 488).

The morpheme =*ma* may or may not occur in a negative sentence, as seen in (706) and (707). As shown in § 5.7.2, the morpheme =*ma* can occur twice; it may attach the first element preceding it before the verbal root, and again follow it, as in (708) and (709).

(706) *pʌɾ=ʌ=na=ŋ,*   *ŋa*   *degi*   *biha*   *dʑahŋ=sa*
      study=LN=NPST=1   1SG   now   wedding   do_make=NMZ1

      mʌn=ma   *na=ŋʌ=lʌ.*
      heart=NEG   COP=1=NEG

      'I study, I don't have the desire to get married.'
      CH_MKW_BMB_BAN_103119_9_Mother and daughter

(707) *majtʌ=ko*   *Chepang*   *bʰasa=pa,*
      maternal.side=GEN   Chepang   language=DIS

      ʌ   *ŋa=ko*   *ba-ama=taŋ*   *noʔ=i=to,*
      uh   1SG=GEN   parents=ATT   speak=PL=REM.PST

      ŋa=i=pa   *saʔj=ŋʌ=lʌ.*
      1SG=ERG=DIS   listen=1=NEG

      'As for the Chepang language of my maternal side, uh, my parents used to speak it, but I didn't listen to it.'
      CH_MKW_KKBP_LPK_101917_3_Chepang_Language

(708) *doh*   *pʌtɕas,*   *pʌjsa=ma*   *na=ma=l=o*   *kʰʌlti=haŋ,*
      what   fifty   money=NEG   COP=NEG=COP=NMZ:REL   pocket=LOC1

      'What fifty, I don't even have money in my pocket.'
      CH_MKW_SC_BGR_101619_4_Songs

(709)　*ku=ma=u=lu,*　　　　　　*manta=ma*　　*sat=ma=u=lu,*
　　　　steal=NEG=3O/DIR=NEG　　person=NEG　　kill=NEG=3O/DIR=NEG

　　　　dege=tɕʌhe　　*pe=to*　　　　　　　　*sjaw=a.*
　　　　now=DIS　　　　be.nice=NMZ:ADV2　　　become=PST

'He doesn't steal, he doesn't kill people, as for now, he became nice.'
CH_CTW_SRP_GUN_102620_4_Local_justice

The morpheme =*ma* developed the function of additive marker. This development triggered the reanalysis of the morpheme =*lʌ* ~ *li* ~ *lu* as a marker of negation from its function of copula. The morpheme =*ma* can be thus found in positive statements such as that in (710) with the meaning of 'too, also.' This particular function is correlated with pre-verbal position of the morpheme =*ma*. In a negative construction, it carries the meaning of 'not even,' as in (711), or keeps its negative meaning, as in (712). It is possible that the morpheme =*ma* further spread its additive function to its use in the post verbal position, as in (713). Very scarce examples are attested for such construction.

In Bhujel, a cognate additive morpheme -*m* is reported (Regmi 2007: 128); it seems to have completely grammaticalized into an additive since only the morpheme -*l*, cognate with the Chepang negation marker =*lʌ* ~ *li* ~ *lu*, is left to mark negation (Regmi 2007: 243).

(710)　*ŋi=ma*　　　　*dʑe=alaŋ=su,*　　　*bʌdi*　*tahŋ=to*
　　　　1PL=ADD　　　eat=1.PST=1PL.EXCL　very　　be.huge_be.like=NMZ:ADV2

　　　　kʰe=to　　　　*tak-si,*　　　　*hʌj !*
　　　　COP=REM.PST　mango-tree　　PART

'We too ate (its mangoes), it was super huge the mango tree, hey!
CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

(711) *tɕʰjo,*   *hʌw=i=ma*   *doh=ma*
enough   younger.sister=ERG=ADD   what=ADD

bʌj=ta=ŋ=lʌ,   *ane...*
give=INV=1=NEG   PART

'Enough, my sister doesn't even give me anything, so...'
CH_CTW_BBC_POL_111720_5_Two Sisters

(712) *naŋ=ko*   *kʰet=ma*   *suk=ŋʌ=lʌ*   *dahj=ti*
2SG=GEN   field=NEG   plant=1=NEG   say=SEQ1

to=alaŋ.
tell_say=1.PST

''I won't plant your field," I told the land-owner.'
CH_MKW_BMB_BAN_103119_4_Meeting with Sahu_Archive

(713) *tɕini-pak=ko=le*   *dʑahŋ=o*   *kʰe=dik*   *nʌbʰʌ=a*
sugar-pack=GEN=DIS   do_make=NMZ:REL   COP=COND.PST   not.be=COND

ŋa=kaj   *ʌ*   *to=ma=tɕja=n=i*   *na=hʌ*   *dʌjʰa.*
1SG=DAT   uh   tell_say=ADD=IRR=DIR/TR=PL   COP=COP   maybe

'If it were for having used sugar to make honey, uh, otherwise, they would have told me too maybe.'
CH_CTW_BBC_POL_102420_2_Tu'm

5.8.2. Introduction to Chepang verbal argument indexation

In this section, I introduce Chepang verbal indexation morphology. I give an account of earlier descriptions (§ 5.8.2.1) found in Caughley (1978; 1982), DeLancey (1981), Thompson (1990), Rutgers (1993) and Watters & Regmi (2008). I show that Chepang verbal indexation can be analyzed as forming a direct-inverse system as suggested by DeLancey (1981) but that this system is highly non-canonical, and give a typological overview of how direct-inverse systems are typologically described in TH and beyond (§ 5.8.3.3).

5.8.2.1. Earlier accounts of Chepang verbal argument indexation

Previous analyses of verbal indexation morphology were conducted by Caughley (1971), Bauman (1975), Caughley (1978; 1982) (§ 5.8.2.1.3 and § 5.8.2.1.5, respectively), in addition to DeLancey (1981) (§ 5.8.2.1.4), Thompson (1990) (§ 5.8.2.1.6), Rutgers (1993) (§ 5.8.2.1.7), and Watters & Regmi (2008) (§ 5.8.2.1.8), based on language data provided in Caughley's analyses (1978; 1982). In this section, I give a summary of each of these analyses.

5.8.2.1.1. Caughley 1971

The first analysis of Chepang verbal morphology is found in Caughley (1971). He describes the complexity of Chepang verbal indexation, looking at the morphology found with both intransitive and transitive verbs. He introduces the notion of "goal" that he also uses in his subsequent analyses (1978, 1982). He compares the independent pronouns and verbal indexation forms along the line of the idea, which goes back to Hodgson (1857, 1874) and Konow (1909) who qualify languages featuring such morphology as "complex pronominalized languages," that verbal morphology arose from the affixation of independent pronouns on the verbal root (Caughley 1971:1):

> The term "pronominalised" refers to the feature of distinguishing in the verb "the person subject by means of pronominal affixes". By this [Konow] presumably means that subject and object agreement is indicated in the verb by repetition of the free subject pronoun, though perhaps in a modified form, as a verb affix. It is this similarity of form between the free pronoun and the affix therefore, which distinguishes pronominalization from general verb agreement.

He describes the use of the morpheme -*taang* as a "goal" in "focus," which corresponds to the modern description of =*ta=ŋ* as an inverse marker followed by the 1st person morpheme =*ŋ*.

He concludes that the "pronominalised" system found in Chepang is similar to that found in the Munda language.

5.8.2.1.2.   Bauman (1975)

Based on the data published in Caughley (1971), Bauman (1975) suggests, following Shafer (1974), that Chepang verbal structure is Tibeto-Burman (TB), and that its morphology and lexical items are comparable to that found in other TB languages spoken in remote areas, rather than to Munda (Bauman 1975: 69–71). Bauman (1975) finds that Chepang verbal morphology is similar to Rawang, Jyarung (rGyalrong), and Limbu (1975: 96–97) and classifies Chepang and Vayu (Hayu) together in a West Central Himalayish sub-branch of TB (1975: 73), following Shafer's (1974) classification. Finally, Bauman (1975: 273–275) provides the first paradigm tables of Chepang. Following Caughley (1971), he concludes that Chepang has two sets of transitive morphology, i.e., a "goal focus" and "object focus" marking.

5.8.2.1.3.   Caughley (1978)

Caughley (1978) uses the term "cross-reference" to basically refer to verbal indexation, with the meaning that participants are marked through indexation on the verb, cross-referencing the roles that the arguments hold in the sentence. For Caughley (1978), these roles are that of "agent" and "goal." These syntactic roles are basically A and O{P/R} arguments, respectively. These arguments can be expressed as noun-phrases that receive case marking (Caughley 1978: 167). When only one participant is indexed on the verb, Caughley (1978: 174) calls it a "topic." Caughley (1978: 167) suggests that "surface case determines," i.e., governs, which argument will be marked or cross-referenced on the verb. Nevertheless, he contradictorily further shows that Chepang features two indexation patterns whose presence does not necessarily correlate with the type of "surface case" appearing on the argument noun-phrase. He gives for instance examples (714) and (715) which show that in both cases, the A and O arguments are

marked the same way, while verbal indexation is different: in (714) the verb indexes the "agent" (A), and in (715), it indexes the "goal" (P/R).

(714)   *sumcaak-ʔi*   *coʔ*   *ʔaamaapaa-nis-kaayʔ*
        three-person-agent   child   parent-two-goal

        waanʔ-naa-n-i.
        bring-nonpast-3rd-person-agent-plural

        'Three people bring a child to two parents.'
        (Caughley 1978: 167)

(715)   *sumcaak-ʔi*   *coʔ*   *ʔaamaapaa-nis-kaayʔ*
        three-person-agent   child   parent-two-goal

        waanʔ-naa-thaa-ca.
        bring-nonpast-3rd-person-goal-dual

        'Three people bring a child to two parents.'
        (Caughley 1978: 167)

The patterns that Caughley (1978) observes are thus not per se determined or governed by the type of case marking present on the argument noun-phrases, that is, not by grammatical or syntactic relations. This raises of course the question of what triggers the choice of one pattern over the other. In the same paper, Caughley (1978: 168) proposes that Chepang's "cross-reference," i.e., indexation, is the result of participant ranking, where the "social role played by the participant, or his thematic importance [at the level of discourse]" determines the presence of the marking of the "agent" (A) or that of the "goal" (P/R): "the higher ranked participant in the clause is the one selected for cross-reference." Caughley (1978: 168) proposes that the "highest ranked individual is the one with the greatest authority or ability to initiate the action," and that this participant ranking is based on "cultural ordering." Caughley (1978: 168) gives four different hierarchies that illustrate the trigger to the presence of one or the other participant. These are reported in Figure 75. They are based on his observations of the different scenarios occurring in narrative texts, each of which is tied to a specific context.

Figure 75. Social cross-reference hierarchy (Caughley 1978: 168)

> king > minister > subject (of kingdom)
>
> husband > wife > sister-in-law
>
> father > mother > older sibling > younger sibling
>
> human > spirit > nonpersonal > animate > inanimate object

When the P/R argument is expressed through verbal indexation, there are certain configurations where its marking is correlated with the use of a special additional morpheme on the verb. In Caughley (1978), it is transcribed *-taa ~ -thaa ~ thaay*. While this morpheme has allomorphs (§ 5.8.3.5), the allomorph *thaay* found in Caughley (1978) is in fact to be further parsed as $=t^ha=i \sim t^h\Lambda=i$, where $=i$ corresponds to the marking of 3rd person. Caughley (1978) does not explain further the presence of this morpheme but considers it as part of the construction that indexes the P/R argument. He glosses this morpheme "goal" in reference to the term "goal" used to define the marking of the P/R argument.

Caughley (1978: 174) notes that both the A and P/R arguments may also be indexed on the verb, calling this pattern a "dual topic." He stresses that dual and plural number can determine which of the two arguments is indexed, in particular when it comes to scenarios where both the speaker and the hearer are involved, typically referred to as Speech Act Participants (SAP).

Finally, as for the analysis of the morphology found in Caughley (1978), a number of forms he proposes turn out to be erroneous. For instance, the form *-naang* is analyzed as a 2nd person marker in 1>2 scenarios (Caughley 1978: 175), while it is the combination of the non-past morpheme *=na* and the 1st person form *=ŋ*. In Caughley (1982), the form *-naŋ* in the same scenario is further inconsistently analyzed as either *-naŋ* to mark 2nd person or as *-na* to mark 2nd person suffixed by *-ŋ* to mark 1st person. These erroneous analyses have been repeated in different papers where Chepang verbal morphology is illustrated or discussed (DeLancey 1981; Thompson 1990; Rutgers 1993; Watters & Regmi 2008; Regmi 2009).

Nevertheless, through his analysis, Caughley (1978) points towards what is typologically described as direct-inverse systems in later literature. Such systems have been traditionally argued to be based on pre-existing referential hierarchies (grounded in prominence, topicality or animacy). Caughley (1978) does not use the terms "direct" or "inverse," but shows that what he calls the "unmarked case" relates to the marking of the A on the verb, i.e., the "agent case," while what he calls the "goal case" is expressed by the presence of the forms *-taa ~ thaa ~ thaay* in association with person and number marking (Caughley 1978: 176–177).

5.8.2.1.4.   DeLancey (1981)

After Caughley (1971; 1978) published his analysis of Chepang verbal indexation, DeLancey (1981: 87–89) shows that Chepang can be described as a form of direct-inverse system, such as that found in other TH languages such as rGyalrong, Nocte and Rawang. DeLancey (1981) analyzes the Chepang forms *-taa* and *-thaa* as inverse markers and the form *-u* as a direct marker.

DeLancey (1981: 88) notices an "odd distribution" where some constructions found in the 2>1 scenario are presented as combining both inverse and direct forms (Caughley 1978: 175). These constructions, reported for 2DU>1DU/PL and 2PL>1DU/PL, are erroneous (§ 5.8.3.5) and were not kept in Caughley (1982). They are reported in (716) along with simplified glosses. The way they are presented is in fact quite confusing since Caughley (1978: 175) uses parentheses around certain morphemes without explaining what these mean. Clearly, the forms as presented do not correspond to any syntagmatic string attested in Chepang verbal indexation morphology.

DeLancey's (1981) analysis of Chepang as exhibiting a form of direct-inverse system is accurate as we will see in § 5.8.3.3.

(716)	2DU>1DU	*(teʔ)-j-u(-taa-ng-ca)*	2-DU-DIR-INV-1-DU
	2DU>1PL	*(teʔ)-j-u(-taa-ng-i)*	2-DU-DIR-INV-1-PL
	2PL>1DU	*(teʔ)-s-u(-taa-ng-ca)*	2-PL-DIR-INV-1-DU
	2pl>1pl	*(teʔ)-s-u(-taa-ng-i)*	2-PL-DIR-INV-1-PL

5.8.2.1.5.   Caughley (1982)

Caughley (1982) offers a more in-depth analysis of verbal indexation than
Caughley (1978). Nevertheless, his approach to the analysis remains nevertheless the
same as Caughley (1978), using the same terminology. Some of the indexation forms
reported in Caughley (1978) do not occur in Caughley (1982), and as shown in § 5.8.3.5,
several are erroneous forms.

Caughley (1982) comprises two chapters that strictly concern Chepang verbal
indexation morphology (717). In the remainder of this section, I will summarize the main
points of Caughley's analysis found in each of these chapters.

(717)   Chapter 2: The verb in relation to content
        Chapter 3: The verb in relation to context

5.8.2.1.5.1. Chapter 2: The verb in relation to content

In his Chapter 2, Caughley (1982: 53–82) introduces Chepang verbal
morphology, and the framework he uses to describe it. Similarly to Caughley (1978), he
uses the terms "cross-reference," "agent," and "goal" (§ 5.8.2.1.3).

In this section, I will only focus on some of his main points regarding verbal
indexation morphology: introduction of "basic" indexation forms (§ 5.8.2.1.5.1.1);
description of indexation reduplication (§ 5.8.2.1.5.1.2); description of a construction
combining "agent" and "goal" (§ 5.8.2.1.5.1.3); description of possessor indexation in
=*bʌt* construction (§ 5.8.2.1.5.1.4); description of a benefactive construction with
intransitive verbs (§ 5.8.2.1.5.1.5).

5.8.2.1.5.1.1. "Pronominal affixes (basic forms only)"

Caughley (1982: 54) introduces a first table that summarizes indexation forms.
The forms that are reported are nevertheless devoid of the context within which they
occur. A few of the forms presented in this table are found in all tenses, and most of them
belong to constructions that express negation or that relate to a specific verbal valency. In

addition, as mentioned in § 5.8.2.1.3, some of these forms are erroneously parsed and glossed, such as those involving 2nd person: *-naŋ* is not a 2nd person form but the combination of the non-past tense morpheme *=na* and the 1st person form *=ŋ* in a construction expressing 1>2 in combination with the presence of a 1>2 morpheme *=ne* or *=tɕe* preceding the non-past morpheme and following the verbal stem.

In Table 137, I report the forms given by Caughley (1982: 54) and comment on their specific occurrence.

Table 137. Caughley (1982: 54) "Pronominal affixes (basic forms only)"

Caughley's gloss	forms	missing details on their specific use	
1ST (EXCL) SG	*-ŋə*	1SG negation (intr./(di)tr.)	
1ST (EXCL) DU	*-ŋə-cə*	1DU negation (intr./(di)tr.)	
1ST (EXCL) PL	*-ŋ-i/sə*	1PL (intr.) / 1PL negation (intr./(di)tr.)	
1ST AND 2ND (INCL) DUAL	*-təyh-cə*	1DU.INCL (intr.)	
1ST AND 2ND (INCL) PLURAL	*-təyh-ʔi*	1PL.INCL (intr.)	
2ND SG	*-naŋ*	1SG>2SG (npst)	suffixal forms that go with the
2ND DU	*-naŋ-jə*	1SG>2DU (npst)	1>2 form *=ne/tɕe* preceding the
2ND PL	*-naŋ-sə*	1SG>2PL (npst)	non-past tense form *=na* such as
			∑*=ne/tɕe=na=ŋ*
3RD SG	Ø	3SG (intr.)	
3RD DU	*-cə*	3DU (intr.)	
3RD PL	*-ʔi/sə*	3PL (intr.)	

In his chapter 2, Caughley (1982: 55) presents the 2nd person form *=te* (noted *teʔ*)[49]. His analysis of *=te* is that it occurs in contexts where 2nd person is involved as a participant and that it is "strictly part of the Information Flow marking system" he

---

[49] Note that no glottal stop occurs on the 2nd person morpheme *=te* in Chepang.

describes in Chapter 3. For Caughley (1982), the morpheme *=te* functions as some sort of focus marker (§ 5.8.2.1.5.2).

5.8.2.1.5.1.2. "Reduplication of indexation"

Caughley (1982: 55) argues that the verbal form *=na*[50], that he mistakenly takes for a 2nd person form in the 1>2 scenario, can be reduplicated and he gives the example reported in (718). This type of example is not attested in our data and may just correspond to a stuttered form or a misunderstanding during elicitation. In addition, note that in this example, the form *=ne* (noted *-neʔ*)[51] is erroneously analyzed as a non-past form; instead, it carries the meaning of 1>2.

(718)   *niŋ-kay   ŋa-ʔi   bəyʔ-neʔ-na-na-na-ŋ-sə.*
        you-GL   1-AG   give-NPT-2-2-2-1E-PL

        'I will give to you all.'
        (Caughley 1982: 55)

5.8.2.1.5.1.3. "Agent and goal combined construction"

Caughley (1982: 55) further mentions a construction used in 3PL>1PL and 3PL>3DU scenarios that would combine the marking of the "agent" and the "goal" (direct and inverse forms respectively in our analysis). He gives two examples, reported in (719) and (720). This type of construction is not attested and may again just correspond to a misunderstanding during elicitation. Caughley (1982: 55) himself notes: "I have never found any examples [like these] in unelicited speech or text." These constructions look like the constructions mentioned in § 5.8.2.1.4 that show the "odd distribution" observed by DeLancey (1981).

---

[50] This form is also transcribed *-nə* in the same kind of example in Caughley (1982: 159). This form is not found in Chepang, as there is no such thing as a 2nd person morpheme *-na*.

[51] Note that like with *=te*, the morpheme *=ne* does not feature a final glottal stop.

(719) *ŋi-kay ʔowʔ-məyʔ-ʔi ghan-naʔ-s-u-naʔ-ta-ŋ-i.*
       we-GL that-CP1-AG   beat-NPT-PL-AG-NPT-GL-1E-PL

       'They beat us.'

       (Caughley 1982: 55)

(720) *ʔi-nis-kay   ʔowʔ-məyʔ-ʔi ghan-naʔ-s-u-naʔ-tha-cə.*
       this-DL-GL   that-CP1-AG   beat-NPT-PL-AG-NPT-GL-DL

       'They beat these two.'

       (Caughley 1982: 55)

5.8.2.1.5.1.4. "Possessor cross-reference"

Caughley (1982: 56) introduces a construction that indexes the possessor of a noun-phrase through the use of a morpheme *=bʌt* (noted -*bət*). This construction is attested in Chepang with or without co-occurrence of the inverse form *=ta* (§ 5.8.3.6). Caughley (1982: 56) gives seven examples, reported in (721) to (727) with Caughley's glosses. Five of these examples present several problems at different levels, and none of them is accepted as a possible form by the Chepang speakers I have worked with.

First, the examples provided by Caughley (1982: 56) are inconsistent in terms of the position of the tense marker. As we can see in (721), (722), and (723), the non-past tense form *=na* (noted -*naʔ*)[52] and the past form *=a* (noted -*ʔa*)[53] follow the verbal root while in (725), and the non-past *=na* (noted -*naʔ*) follows the morpheme *=bʌt* (noted -*bət*). In fact, in this construction, the tense markers only follow the morpheme *=bʌt* (§ 5.8.3.6). Second, the morpheme *=bʌt* is yet not attested in our data with the perfect aspect marker *=o* (§ 5.8.3.6), contradictorily to (724). Third, this type of construction is not attested with inanimate possessed entities unless the possessed noun-phrase is an O argument in presence of the expression of the A argument (§ 5.8.3.6). Examples (722) and (723) are therefore not recognized by the Chepang speakers I worked with. Fourth,

---

[52] Note that the non-past morpheme *=na* does not feature a final glottal stop.

[53] Note that the past morpheme *=a* does not feature an initial glottal stop.

the =bʌt construction is only attested with a 2nd person possessor when the verb is intransitive (§ 5.8.3.6), making (725) an unreceivable example as well.

Caughley (1982: 72, 97) observes that the use of this construction expresses that the possessor is emotionally affected by the process. This corresponds to what I observed as well with the speakers I worked with. It conveys a benefactive/malefactive meaning.

Finally, the morpheme =bʌt is reported to occur in nominalized clauses (Caughley 1982: 42), as illustrated in (728). This construction is yet not attested in our data.

The =bʌt construction is described in § 5.8.3.6. It should be noted that this construction is considered archaic in all studied varieties.

(721) *ŋa-koʔ coʔ   ryaʔ-naʔ-bət-ta-ŋʔ.*
       I-GEN   child   cry-NPT-POS-DL-1E

   'My child is crying.'

   (Caughley 1982: 56)

(722) *yam   way-ʔa-bət-ta-ŋʔ.*
       rice   destroy-PT-POS-GL-1E

   'My rice is destroyed.'

   (Caughley 1982: 56)

(723) *ŋa-koʔ məkəy   duŋ-na-bət-ta-ŋʔ.*
       1-GEN corn   grow-NPT-POS-GL-1E

   'My corn is growing (for me).'

   (Caughley 1982: 51)

(724) *ŋa-koʔ coʔ   waŋ-ʔo-bət-ta-ŋʔ.*
       I-GEN   child   come-RN-POS-GL-1E

   'My child has come.'

   (Caughley 1982: 56)

(725) *naŋ-koʔ*   *kuyʔ-ʔi*   *ram-kay*   *jəyk-bət-naʔ-thəy.*
       you-GEN   dog-AG   Ram-GL   bite-POS-NPT-GL

*'Your dog bit Ram.'

'Your dog is going to bite Ram.'

(Caughley 1982: 56)

(726) *ŋi-ci-koʔ*   *coʔ*   *si-ʔa-bət-ta-ŋʔ-cə.*
       1-DL-GEN   child   die-PT-POS-GL-1E-DL

'Our child has died.'

(Caughley 1982: 72)

(727) *eheʔ*   *ŋa-koʔ*   *coʔ-jaʔ*   *si-ʔa-bət-ta-ŋʔ*   *ba.*
       EXCL   1-GEN   child-EV   die-PT-POS-GL-1E-DL   CERT

'Oh my child has died!'

(Caughley 1982: 97)

(728) *yom-ʔi*   *coʔ*   *sat-bət-o*   *manta.*
       bear-AG   child   kill-POS-RN   person

'The person whose child was killed by a bear.'

(Caughley 1982: 42)

Based on the examples found in Caughley (1982), this construction is analyzed as a Prominent Internal Possessor (PIP) construction by Nikolaeva, Bárány, & Bond (2019). They observe that the possessor may only be the subject or object argument of the verb (Bárány, Bond & Nikolaeva 2019: 22), that is, an A, S, or O argument. More on PIP constructions and how they are defined can be found in § 5.8.3.6.

5.8.2.1.5.1.5. "Benefactive goal"

Caughley (1982: 69) describes a benefactive construction that he calls "benefactive goal," used with the intransitive verbs *waŋ-* 'come' and *al-* 'go' suffixed with the 1>2 imperative form *=tɕi*. He presents the only example he has with the verb

waŋ- 'come,' reported in (729). This construction is not possible according to the Chepang speakers I worked with.

(729) *dyahmay   jhya-laŋ   waŋ-ci.*[54]
      tonight   drumbeat-PUR come

      'Tonight, come to drumbeat for me.'

      (Caughley 1982: 69)

Benefactive constructions may be formed instead with the verb *bʌj-* 'give,' through calque from the Nepali benefactive construction. The native possessive *=bʌt* construction also conveys a benefactive/malefactive meaning (§ 5.8.3.6).

5.8.2.1.5.2. Chapter 3: The verb in relation to context

In Chapter 3, Caughley (1982: 83–113) introduces the notion of "information flow" that he uses to describe what are, in our analysis, two discourse markers *=paj* (noted *-pay*) and *=taŋ* (noted *-taŋʔ*)[55], and the 2nd person morpheme *=te* (noted *-teʔ*).

He gives a first set of illustrative examples reported in (730) to (732).

(730) *ŋa-ʔi-pay   bəyʔ-neʔ-naŋ.*[56]
      I-AG-DIF   give-NPT-2-1E

      'I will give it to you.'

      (Caughley 1982: 84)

(731) *naŋ-kay   bəyʔ-teʔ-ʔa.*
      you-GL   give-CIF-PT

      'He gave it to you.'

      (Caughley 1982: 84)

---

[54] The form *-ci* is not glossed in this example (Caughley 1982: 69).

[55] Note that the attentional/hearsay marker *=taŋ* does not feature a final glottal stop.

[56] The two forms *-neʔ* and *-naŋ* are glossed with three functions (Caughley 1982: 84).

(732) *ram-koʔ*   *ʔama*   *sitə-taŋʔ-leʔ.*
      Ram-GEN   mother   Sita-IIF-REM

'Ram's mother is Sita (he said).'

(Caughley 1982: 84)

Caughley (1982: 84–88) argues that these three forms are part of a system that expresses "the direction of flow of information." He describes these morphemes as having two types of functions that he calls "primary functions" and "secondary functions." These "primary functions" have to do with the "direction" of the information expressed, in an evidential sense (Caughley 1982: 84).

He describes =*paj* as a "direct information flow marker," a morpheme which expresses an information "originating from the speaker," i.e., the speaker is at the source of the information.

He analyzes the form =*taŋ* as an "indirect information flow" marker, a reported speech marker which, by contrast with =*paj*, expresses that the information does not originate from the speaker but that it was heard from someone else.

As for =*te*, Caughley (1982: 84) emphasizes the fact that =*te* is somewhat "anomalous" in this system since it is found correlating with 2nd person participants. He considers that the morpheme =*te* does not fundamentally belong to the verbal indexation forms since he considers that indexation forms historically come from independent pronouns, and that =*te* is not found as a 2nd person independent pronoun. Another argument put forward by Caughley for not considering =*te* as a 2nd person indexation form, is the fact that he analyzes the form -*naŋ* as a 2nd person indexation form (though it is not, but is the combination of the non-past morpheme =*na* and the first person form =*ŋ* in 1>2 scenario (§ 5.8.3)), in line with his idea that =*naŋ* would originally come from the independent 2nd person pronoun *naŋ*, and that only one 2nd person indexation form is enough: "Moreover, there is already a perfectly good 2nd person pronoun [=*naŋ*] which does occur as a free form" (Caughley 1982: 84). Caughley (1982: 85) thus considers that =*te* mainly carries an evidential function. He (Caughley 1982: 85) notes: "The third form -*teʔ* therefore indicates a flow of information which is contrary to the expected direction, since a person is normally expected to be the source, not the recipient of information

561

concerning his own actions. In effect it is alerting the addressee to the fact that the utterance concerns him, not someone else."

Finally, based on the fact that the forms =*paj*, =*taŋ*, and =*te* can occur more than once in a sentence, Caughley (1982: 85) suggests that their secondary function is that of indicating the status of the information as new (=*taŋ* and =*te*) or given (=*paj*), i.e., focus and topic, respectively.

The hypothesis that =*paj*, =*taŋ*, and =*te* form an evidential system that indicates the status of the information expressed as new or given does not hold: =*paj* is a discourse marker that indicates a contrast, =*taŋ* is also a discourse marker that carries both an attentional and reported speech function, and =*te* is a 2nd person indexation form.

The rest of Chapter 3 covers negation, tense, aspect, and modality, which I will not comment on here.

5.8.2.1.6.   Thompson (1990)

Thompson **(1990)** offers a study of the direct-inverse system of Chepang based on the data found in Caughley (1982) and a collection of Chepang texts (1970c). Thompson's **(1990)** study does not intend to describe the Chepang verbal indexation system as a whole. Instead, it seeks to understand the topicality of direct object arguments in discourse and whether the construction marking the object argument on the verb in presence of the morpheme =*ta* (and allomorphs) should be analyzed as an inverse versus a passive construction. He intends to measure the "predictability and importance" of an argument **(1990: 407)** using Givón's (1983) methodology, which consists in measuring an argument's referential distance through referent persistence and referent quotient, i.e., respectively the importance of an argument at the local level (within a range of ten sentences) and at the level of discourse.

Thompson **(1990: 408)** for the first time acknowledges that the morphemes =*i* (noted -*'i*) and =*kaj* (noted -*kay*) that Caughley (1982) respectively calls "agent" and "goal" are in fact ergative and dative case markers, respectively.

While Thompson **(1990: 416)** recognizes that =*u* is a direct morpheme, as initially proposed by DeLancey (1981), he is the first to analyze the morpheme =*n* as a direct

morpheme as well. This latter analysis holds in the sense that *=u* and *=n* appear in direct constructions, and their distribution is complementary, since they do not share the same tense/person configurations. For example, in 2SG>3 scenarios, the morpheme *=u* occurs in the non-past tense, as in (733) and the morpheme *=n* in the past tense, as in (734).

(733) gʌ=tɕjuk=ko   naŋ=i   mʌhj=te=na=u
       INT=QTY=GEN   2SG=ERG   recite.mantra=2=NPST=3O/DIR

       u=tɕjuk=ko   lek-si-siŋ   bɾaw=na.
       ANA=QTY=GEN   flame-tree-tree   be.big=NPST

       'How much you recite the mantras, that much the flame tree becomes big.'
       CH_CTW_JMC_PYK_101920_Cing_Lan_Archive

(734) lʌw,   tɕiŋ-lan   ɾʌjsʌ   naŋ,
       well   Cing-spirit   COP.MIR   2SG

       ŋa=ko   tɕaɾ   bʌhini-tɕoʔ
       1SG=GEN   four   younger.sister-child

       mak=te=ka=n.
       devour_eat.completly=2=2/3.PST=DIR/TR

       'Well, it turns out that you are the Cing spirit, you devoured my four little girls.'
       CH_CTW_JMC_PYK_101920_Cing_Lan_Archive

Thompson **(1990: 416)** follows Caughley in analyzing the inverse morpheme *=ta* as having three allomorphs *-ta ~ -tha ~ -thəy*, while it has the three allomorphs *=ta ~ tʰa ~ tʰʌ*, the third one to be further parsed as *=tʰa=i* with *=i* marking 3rd person singular O{P/R} argument.

Thompson **(1990: 417–418)** suggests that the inverse morpheme *=ta* may either originate from the 2nd person marker *=te* on the grounds that with 2nd person, the morpheme *=te* occurs and not *=ta*, or from an old perfective marker. The latter hypothesis is favored by Thompson **(1990: 418)**. He cites Caughley's (1982: 185) example, reported in (735), which shows that a past perfective morpheme *-tə* occurs in Bhujel (then referred

to as Western Chepang). However, Thompson **(1990: 418)** does not transcribe this morpheme as *-tə* but as *-ta*. In fact, Thompson's hypothesis is not substantiated, since Bhujel (Regmi 2007: 214) indeed features a remote past morpheme *-t* ~ *-ta* which is cognate with Chepang remote past morpheme *=to*. The origin of *=ta* remains unclear, but it is more likely that it comes from a verb.

(735) *ŋiʔ-tə-cə.*
       laugh-PT.PF-DL

     'The two laughed.'

     (Caughley 1982: 185)

Thompson **(1990: 421)** concludes that Chepang features *=ta* constructions that are better analyzed as inverse than passive: "Inverse constructions, on the other hand, promote a non-agent, but do not necessarily demote an agent. By this criterion, it is clear that the Chepang construction is an inverse and not a passive. What is striking about the text counts for Chepang discussed above is that they indicate a topical non-agent, more than a non-topical agent."

5.8.2.1.7.   Rutgers (1993)

Rutgers (1993) offers another reanalysis of Chepang verbal indexation morphology, seeking to compare it to the verbal indexation system posited by van Driem (1990; 1991a; 1991b; 1992) for Proto-Kiranti. Through this study, Rutgers (1993) aims at better understanding the position of Chepang within TH.

The analysis that Rutgers (1993) provides focuses on the verbal indexation forms in their paradigmatic and syntagmatic distribution in the verbal string. He does not intend to describe the system as a whole.

Rutgers (1993) bases his analysis on the Chepang data found in Caughley (1978; 1982) and a collection of Chepang texts (1970c). Rutgers (1993: 110–111) notes that he mostly relied on the paradigms presented in the tables in appendices of Caughley (1982:

218–219) since the examples found in texts (1970c; 1982) did not show enough diversity in the forms expressed.

Rutgers (1993: 111) notes that the paradigms given in Caughley (1978) and Caughley (1982) present some differences, in particular when it comes to local scenarios (1>2 and 2>1).

Since Caughley (1982) does not include these forms in the paradigms, Rutgers (1993: 110) does not take them into consideration in his analysis. For instance, the forms used to index 1>2 that are reported in Caughley (1978) are precisely those recognized as correct by the Chepang speakers I worked with, while those in Caughley (1982) are not: the form used for 1DU.EXCL>2SG in the non-past tense can be *=ne=na=ŋ=dʑʌ* (noted *neʔnaaŋjə*) 1>2=NPST=1=DU; and not *-neʔ-nə-naŋ-dʑʌ*, where there is no such thing as a morpheme *-nə* and where *-naŋ* is to be further parsed as *=na=ŋ* NPST=1. In doing so, Rutgers (1993: 112–113, 115) reports the same erroneous forms regarding 1>2 and misses the point in analyzing *=nə* and *=na* as 2nd person morphemes in 1>2 and 2>1 scenarios, respectively.

Rutgers' (1993) analysis does not recognize that Chepang features a form of direct-inverse system (Rutgers 1993: 123–124) as suggested by DeLancey (1981). He considers that the form *=ta=ŋ* (noted *-taŋ*) marks a 1st person exclusive patient (Rutgers 1993: 116), that the morpheme *=u* (noted *-u*) marks a 3rd person patient (Rutgers 1993: 119), and that the morpheme *=n* indexes 2SG/3SG>3 in the past tense.

Finally, Rutgers (1993: 127) concludes that Chepang may not be considered a Kiranti language and recognizes that the morpheme *-na* should not be analyzed as a 2nd person marker in 2>1 scenario.

5.8.2.1.8.   Watters & Regmi (2008)

When describing the direct-inverse system found in Bhujel, the most closely related language to Chepang, Watters & Regmi (2008) briefly illustrate the direct-inverse system of Chepang using examples from Caughley (1982).

While they do not cite them, Watters & Regmi (2008) follow DeLancey (1981) in analyzing *=u* as a direct morpheme and Thompson (1990) in analyzing the morpheme *=n* as a direct morpheme.

5.8.3. Chepang verbal argument indexation

In this section, I start by describing verbal argument indexation attested in Chepang in intransitive (§ 5.8.3.1) and reflexive/reciprocal constructions, along with providing their complete paradigms (§ 5.8.3.2). I introduce the Chepang non-canonical direct-inverse system with regard to TH and beyond (§ 5.8.3.3) and summarize Chepang argument indexation system with regard to such non-canonicality (§ 5.8.3.4). I then provide a description of transitive and ditransitive constructions (§5.8.3.5) for non-past and past tenses (§ 5.8.3.5.1), remote past (§ 5.8.3.5.1), and optative/imperative (§ 5.8.3.5.3), based on full paradigms.

Note that negative constructions are similar in non-past and past tenses, since no overt non-past or past morphology appears in these constructions. The negation morpheme is *=lʌ ~ lu ~ li,* and a facultative morpheme *=ma* can occur following the verbal stem or preceding it (§ 5.8.1). This morpheme further conveys the meaning of an additive, i.e., 'even,' but is the original negation marker, while *=lʌ* was a copula. The morpheme *=lʌ* underwent progressive vowel harmony, where the quality of the vowel /ʌ/ changed into /i/ when following the plural morpheme *=i,* giving the allomorph *=li,* and into /u/ when following the 3rd person object or direct marker *=u,* giving the allomorph *=lu.*

Beyond the morphosyntactic complexity of the non-canonical direct-inverse system of Chepang, I describe additional layers of complexity that lie in the speakers' usage of direct and inverse constructions triggered by socio-pragmatic effects. I suggest that referential hierarchies cannot fully grasp the way Chepang speakers use such a system, and that the sociopragmatic effects driving the use of one possible construction over another possible construction in identical scenarios can shed light on the dynamicity of a system which could well be the reflection of a typological trait that existed at the

level of PTH, that could further explain the diversity of the verbal systems that developed from it in different clades.

5.8.3.1.   Intransitive verbal argument indexation

Intransitive verbs morphosyntactically index a single argument on the verb, i.e., S argument, as in (736) with the 3ʳᵈ person plural morpheme =*i*. When overtly expressed in a sentence, this argument occurs in an absolute form. Intransitive verbs include as well stative verbs, as in (737) with *bʰam-* 'be white,' which may have the property of being both dynamic and stative (§ 5.4). In addition, as shown in (§ 3.4.5) an array of constructions formed with lexialized intransitive verbs, as illustrated in (738).

Examples of intransitive verbs are given in Table 138. The following sub-sections presenting the attested complete paradigms are as follows:

- § 5.8.3.1.1 Intransitive non-past and past paradigm
- § 5.8.3.1.2 Intransitive remote past paradigm
- § 5.8.3.1.3 Intransitive irrealis paradigm
- § 5.8.3.1.4 Intransitive optative paradigm
- § 5.8.3.1.5 Intransitive imperative paradigm

(736)   *satʰi=lʌm*   *waŋ=ka=i,*   *sʌhʌjog*   *dzahŋ=laŋ.*
       friend=PL   come=2/3.PST=PL   help   do_make=PUR

       'Friends came, to help.'

       CH_MKW_SC_SIL_120619_2_E_5

(737)   *rʌ*   *Dagu*   *radza=ko*   *kliʔ*   *bʰam=a.*
       and   Dadu   king=GEN   feces   be.white=PST

       'And the feces of the king Dagu whitened.'

       CH_CTW_BBC_POL_102420_3_Chepang_Kings

(738) o   bela=haŋ=paj   sikaɾ   ges=i=to,   sikaɾ.
       DIST   moment=LOC1=DIS   hunt   play=PL=REM.PST   hunt

'At that time people would hunt.'

CH_CTW_SBPC_DHM_111220_2_Kusunda

Table 138. Intransitive verbs

intransitive verb	meaning
al-	'go'
bah-	'crawl'
blajk-	'feel lazy, bored'
eʔn-	'sleep'
ke-	'sing (song)'
kɾʌw-	'flee'
ku-	'smoke (fire)'
noʔ-	'speak'
pahj-	'leave, go home'
pe-	'be nice, good'
ɾja-	'cry'
si-	'die'
tɕjuŋ-	'sit'
wah-	'walk'
waŋ-	'come'

5.8.3.1.1.   Intransitive non-past and past paradigm

Verbal argument indexation forms associated with intransitive verbs in past and non-past tenses are presented in Table 139. In this table, the symbol ∑ indicates the position of the verbal stem and 'TM' that of the tense and modality markers, that is for non-past, the morpheme =*na* and for past, the morpheme =*a* with 2nd and 3rd person singular and =*aka* ~ *ka* with 2nd and 3rd person non-singular. The past form associated with 1st person is portemanteau, i.e., combining past and 1st person functions; it is =*alaŋ*

568

in all studied varieties and =*akaŋ* ~ *kaŋ* in RAP-13 (Polkim, Sarling, Syamrang, Yuiling, Santhali).

Table 139. Intransitive verbal argument indexation in non-past and past

Intransitive (PST/NPST)		
	POS.	**NEG.**
1SG	∑-TM-**ŋ**	∑(-ma)-**ŋʌ**-lʌ
1DU.INCL	∑-TM-**tʌhj-tɕʌ**	∑(-ma)-**tʌhj-tɕʌ**-lʌ[57]
1DU.EXCL	∑-TM-**ŋ-tɕʌ**	∑(-ma)-**ŋʌ-tɕʌ**-lʌ
1PL.INCL	∑-TM-**tʌh-i**	∑(-ma)-**tʌh-i**-li[58]
1PL.EXCL	∑-TM-**ŋ-su** (<tr.) ~ **-ŋ-i** (*arch.* intr.)	∑(-ma)-**ŋʌ-su**-lu ~ (-ma)-**ŋ-i**-li
2SG	∑-**te**-TM	∑(-ma)-**te**-lʌ
2DU	∑-**te**-TM-**dʑʌ**	∑(-ma)-**te-dʑʌ**-lʌ
2PL	∑-**te**-TM-**i**	∑(-ma)-**te-i**-li
3SG	∑-TM	∑(-ma)-lʌ
3DU	∑-TM-**tɕʌ**	∑(-ma)-**tɕʌ**-lʌ
3PL	∑-TM-**i**	∑(-ma)-**i**-li

The detailed analysis of each indexation form (person and number) associated with intransitivity and non-past and past tenses is given in Table 140.

---

[57] Negation for 1DU.INCL in past tense; in non-past, the jussive/imperative negative form applies.

[58] Negation for 1DU.INCL in past tense; in non-past, the jussive/imperative negative form applies.

Table 140. Person and number indexation forms – Intransitive, non-past, past

1SG.NPST	**-ŋ**
1SG.PST	**-alaŋ ~ akaŋ ~ kaŋ**
1NEG	**-ŋʌ**
1DU.INCL	**-tʌhj**
1PL.INCL	**-tʌh**
1/3.DU	**-tɕʌ**
1/2/3.PL	**-i**
1PL.EXCL	**-su** (<tr.) ~ **i** (*arch*. intr.)
2	**-te**
2DU	**-dʑʌ**

In remote past tense, irrealis and optative, as we will see, basically the same argument indexation forms are used, but there are a few differences to be noted.

The 1st person form in remote past (§ 5.8.3.1.2) is *=ŋʌ* and not *=ŋ*, like with negation, and the negative construction has an additional morpheme *=jak* which specifies that the process took place before, in a remote past or anterior to time of speech, and often suggests that such process does not occur anymore.

With 1st person plural exclusive, only the construction with the form *=su* is attested in remote past tense, and with irrealis (§ 5.8.3.1.3) and optative (§ 5.8.3.1.4) modalities, i.e., *=ŋʌ=su* (REM.PST) and *=ŋ=su* (IRR, OPT).

Negative forms in optative (§ 5.8.3.1.4) constructions are different from that found in past, non-past, remote past and irrealis constructions. They are similar to the negative forms used with the imperative (§ 5.8.3.1.5), i.e., *=ljam* (LOTHAR) ~ *lam* (MANAHARI).

With 2nd person, irrealis forms express negation (§ 5.8.3.1.3)

Finally, with dual and plural inclusive forms, the negative construction only occurs in past tense. In non-past, the jussive/imperative negative form applies (§ 5.8.3.1.5), i.e., *=ljam* (LOTHAR) ~ *lam* (MANAHARI).

5.8.3.1.2. Intransitive remote past paradigm

Verbal argument indexation forms associated with intransitive verbs in remote past are presented in Table 141. In this table, the symbol ∑ indicates the position of the verbal stem. In this table, I include the forms marking the remote past tense, which are =*to* with the positive forms and =*jak* with the negative forms. When an attested variety shows a different form, it is noted.

Table 141. Intransitive verbal argument indexation in remote past

Intransitive (REMOTE PAST)		
	POS.	**NEG.**
1SG	∑**-ŋʌ**-to	∑(-ma)-jak-**ŋʌ**-lʌ
1DU.INCL	∑**-tʌhj-tɕʌ**-to	∑(-ma)-jak-**tʌhj-tɕʌ**-lʌ
1DU.EXCL	∑**-ŋʌ-tɕʌ**-to	∑(-ma)-jak-**ŋʌ-tɕʌ**-lʌ
1PL.INCL	∑**-tʌh-i**-to	∑(-ma)-jak-**tʌh-i**-li
1PL.EXCL	∑**-ŋʌ-su**-to (<tr.)	∑(-ma)-jak-**ŋʌ-su**-lu
2SG	∑**-te**-to	∑(-ma)-jak-**te**-lʌ
2DU	∑**-te-dʑʌ**-to	∑(-ma)-jak-**te-dʑʌ**-lʌ
2PL	∑**-te-i**-to [tejto~teito]	∑(-ma)-jak-**te-i**-li
2PL	∑**-t-i**-to (POL)	∑(-ma)-jak-**t-i**-li (POL)
3SG	∑-to	∑(-ma)-jak-lʌ
3DU	∑**-tɕʌ**-to	∑(-ma)-jak-**tɕʌ**-lʌ
3PL	∑**-i**-to	∑(-ma)-jak-**i**-li

The detailed analysis of each indexation form (person and number) associated with intransitivity and remote past tense is given in Table 142.

Table 142. Person and number indexation forms – Intransitive, remote past

1SG.REM.PST	**-ŋʌ**
1NEG	**-ŋʌ**
1DU.INCL	**-tʌhj**
1PL.INCL	**-tʌh**
1/3.DU	**-tɕʌ**
1/2/3.PL	**-i**
1PL.EXCL	**-su** (<tr.)
2	**-te**
2DU	**-dzʌ**

5.8.3.1.3.   Intransitive irrealis paradigm

Verbal argument indexation forms associated with intransitive verbs in irrealis are presented in Table 143. In this table, the symbol ∑ indicates the position of the verbal stem. In this table, I include the form marking the irrealis, which is *=tɕja* in the Lothar varieties and *=tɕa* in some Manahari varieties (RAK-8).

Note that the 2nd person positive irrealis forms also function as negative imperative. This is illustrated in (739) and (740).

Table 143. Intransitive verbal argument indexation in irrealis

Intransitive (IRREALIS)		
	POS.	NEG.
1SG	∑-tɕja-**ŋ**	∑(-ma)-**ŋʌ**-lʌ
1DU.INCL	∑-tɕja-**tʌhj-tɕʌ**	∑[59]
1DU.EXCL	∑-tɕja-**ŋ-tɕʌ**	∑(-ma)-**ŋʌ-tɕʌ**-lʌ
1PL.INCL	∑-tɕja-**tʌh-i**	∑[60]
1PL.EXCL	∑-tɕja-**ŋ-su** (<tr.)	∑(-ma)-**ŋʌ-su**-lu
2SG	∑-**te**-tɕja (IMP.NEG)	∑(-ma)-**te**-lʌ
2DU	∑-**te**-tɕja-**dʑʌ** (IMP.NEG)	∑(-ma)-**te-dʑʌ**-lʌ
2PL	∑-**te**-tɕja-**i** (IMP.NEG)	∑(-ma)-**te-i**-li
2PL		∑(-ma)-**t-i**-li (POL)
3SG	∑-tɕja	∑(-ma)-lʌ
3DU	∑-tɕja-**tɕʌ**	∑(-ma)-**tɕʌ**-lʌ
3PL	∑-tɕja-**i**	∑(-ma)-**i**-li

The detailed analysis of each indexation form (person and number) associated with intransitivity and irrealis is given in Table 144.

Table 144. Person and number indexation forms – Intransitive, irrealis

1	**-ŋ**
1NEG	**-ŋʌ**
1DU.INCL	**-tʌhj**
1PL.INCL	**-tʌh**
1/3.DU	**-tɕʌ**
1/2/3.PL	**-i**
1PL.EXCL	**-su** (<tr.)
2	**-te**
2DU	**-dʑʌ**

(739) - *naŋ krʌw=te=tɕja da   ane,*
       2SG   flee=2=IRR   PART   PART

       - *krʌw=ŋʌ=lʌ   krʌw=ŋʌ=lʌ !*

       flee=1=NEG   flee=1=NEG

       '- Then you would flee,

       - I won't flee, I won't flee!'

       CH_CTW_KMC_TAP_102520_2_The bat and the crab

(740) *ama,   pir=te   dʑahŋ=tɕja=u*
       mother   worry=2   do_make=IRR=3O/DIR

       ʌ   "ŋa   si=na=ŋ"   mʰʌr=te=tɕja,

       uh   1SG   die=NPST=1   think=2=IRR

       si=te=tɕja,   si=sa   kʰe=lʌ,   pahj=ti   waŋ=ʌ.

       die=2=IRR   die=NMZ1   COP=NEG   leave=SEQ1   come=2SG.IMP.INTR

       'Mother, don't worry, uh, you might think "I will die," don't die, dying is not the
       solution, leave and come (live with me).
       CH_MKW_SC_SIL_010220_3_Life

5.8.3.1.4.   Intransitive optative paradigm

Verbal argument indexation associated with intransitive verbs in optative mood
are presented in Table 145. In this table, the symbol ∑ indicates the position of the verbal
stem. In this table, I include the form marking the optative, which is *=pa.*

Table 145. Intransitive verbal argument indexation in optative

Intransitive (OPTATIVE)	
	POS. [61]
1SG	∑-pa-**ŋ**
1DU.INCL	∑-pa-**tʌhj-tɕʌ**
1DU.EXCL	∑-pa-**ŋ-tɕʌ**
1PL.INCL	∑-pa-**tʌh-i**
1PL.EXCL	∑-pa-**ŋ-su** (<tr.)
2SG	∑-**te**-pa
2DU	∑-**te**-pa-**dʑʌ**
2PL	∑-**te**-pa-**i**
3SG	∑-pa
3DU	∑-pa-**tɕʌ**
3PL	∑-pa-**i**

The detailed analysis of each indexation form (person and number) associated with intransitivity and optative is given in Table 146.

Table 146. Person and number indexation forms – Intransitive, optative

1	**-ŋ**
1DU.INCL	**-tʌhj**
1PL.INCL	**-tʌh**
1/3.DU	**-tɕʌ**
1/2/3.PL	**-i**
1PL.EXCL	**-su** (<tr.)
2	**-te**
2DU	**-dʑʌ**

---

[61] For negation with all persons, jussive/imperative negative forms apply.

5.8.3.1.5. Intransitive imperative paradigm

Verbal argument indexation forms associated with intransitive verbs in imperative are presented in Table 147. In this table, the symbol ∑ indicates the position of the verbal stem.

Note that 1st person jussive uses irrealis in the form of a question (to self or others) and non-past morphology in statement. Inclusive and exclusive distinctions are not maintained in the imperative. With 3rd person hortative, the optative forms are used.

Imperative negation is marked with the morpheme =*ljam* (LOTHAR) ~ *lam* (MANAHARI) for all persons.

Table 147. Intransitive verbal argument indexation in optative

Intransitive (IMPERATIVE)	
	POS.
1SG	∑-tɕja-ŋ (Q)
1SG	∑-na-ŋ (A)
1DU	∑-tɕʌ
1DU	∑-tɕja-ŋ-tɕʌ (Q)
1PL	∑-i
2SG	∑-ʌ
2DU	∑-dʑʌ
2PL	∑-nʌ
2PL.H	∑-sa-sjaw-pa
3SG	∑-pa
3DU	∑-pa-tɕʌ
3PL	∑-pa-i

The detailed analysis of each indexation form (person and number) associated with intransitivity and imperative is given in Table 148.

Table 148. Person and number indexation forms – Intransitive, imperative

1	-ŋ
1/3.DU	-tɕʌ
1/3.PL	-i
2SG	-ʌ
2DU	-dzʌ
2PL	-nʌ

5.8.3.2. Reflexive, and reciprocal verbal argument indexation

The differences between intransitive and middle/reflexive/reciprocal constructions lie in the presence of the morpheme =sʌ ~ si in middle/reflexive constructions and the morpheme =kaj in reciprocal constructions. The morpheme =sʌ underwent progressive vowel harmony, where the mid-central vowel /ʌ/ of the suffix =sʌ changed into /i/, copying the vowel of the preceding plural marker =i, giving the allomorph =si.

Verbs marked with reflexive morphology are morphosyntactically indexed for a single S argument, as in (741). Like intransitive constructions, the S argument occurs in the absolutive. Reflexive constructions express a process self-initiated, and self-directed by the participant that occurs as an S argument. In fact, this type of construction is only attested when the process is physically directed towards and affecting the body of the participant. Therefore, cognitive verbs expressing meaning like 'ask self,' or 'look' do not occur with the reflexive/reciprocal form.

Note that the morpheme used to mark non-past is similar, i.e., =na, but the one marking past tense is different. It is not =a but =aka ~ ka, as in (742), similarly to the past morphology associated with 2nd and 3rd person non-singular with intransitive verbs or 2nd and 3rd person with (di)transitive verbs. Finally, some speakers may express a reflexive construction without a reflexive marker, as in (743). This is considered an acceptable form, but it is also recognized as somewhat incomplete by the speakers.

While reflexive morphology is found with a certain type of verb that we can call reflexive, note that reflexive morphology is not solely dedicated to such specific verb.

For more details on the type of argument structures found with reflexive verbs, see § 3.4.5.

Examples of verbs that can take reflexive morphology are given in Table 149. The following sub-sections presenting the attested complete paradigms are as follows:

- § 5.8.3.2.1 Reflexive/reciprocal non-past and past paradigm

- § 5.8.3.2.2 Reflexive/reciprocal remote past paradigm

- § 5.8.3.2.3 Reflexive/reciprocal irrealis paradigm

- § 5.8.3.2.4 Reflexive/reciprocal optative paradigm

- § 5.8.3.2.5 Reflexive/reciprocal imperative paradigm

(741)   *sjaʔn,*   *o=haŋ=le*   *boɾ=na=sʌ.*
       insect   DIST=LOC1=DIS   huddle=NPST=REFL

'The caterpillar, there, huddles.'

CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling

(742)   *ʌh*   *ŋa=ko*   *badʑa,*   *naj*   *law=o=ko*
       well   1SG=GEN   grandfather   clothes   apply_wear=NMZ:REL=GEN

       kuɾa   *dahj=sa,*   *dʑʌmma*   *naj*   *dwi*   *din*   *mʌtɾʌj=taŋ*
       thing   say=NMZ1   all   clothes   two   day   only=ATT

       law=ka=sʌ,   *dʌwɾa-suɾwal.*
       apply_wear=2/3.PST=REFL   upper.garment-trouser

'Well, my grandfather, regarding the kind of clothes he was wearing, in all he wore clothes only two days, the traditional Nepalese men outfit.'

CH_MKW_BBC_SIL_032920_2_My grandfather

(743)   *ŋa*   *mjaŋ*   *ɾʌʔj=na=ŋ.*
       1SG   hair   wash=NPST=1

'I wash my hair.'

CH_MKW_1_28-31_BMB_BAN_100917_1_E

Table 149. Examples of verbs found with reflexive morphology

verb with reflexive	meaning
boɾ-	'coil self up'
ʥik-	'hurt'
lʰis-	'come off (nail, cuticle)'
dut kʰas-	'leak (breast milk)'
klju-~klu-	'fall off (hair, hairs)'
ʥis-	'wash (body, body parts)'
nuhl-~nuh-	'rub (body, body parts)'
mjaŋ tis-	'plait hair'
mjaŋ ɾʌʔj-	'wash hair'
mjaŋ sjal-	'tidy hair'
pit-	'pinch'
kɾjuŋ-	'fold (arm, leg)'
ŋoltum gɾjus-~gɾus-	'sit on knees'
mʰeʔ ahŋ-	'heat self-up by fire'
tuk bljʌw-	'be nauseated'
lʰuŋ blʌn-	'have indigestion'
sos-	'be itchy'

5.8.3.2.1.   Reflexive/reciprocal non-past and past paradigm

The argument indexation forms found in reflexive and reciprocal constructions in non-past are presented in Table 150, and that associated with past are in Table 151. The non-past morpheme is *=na* and the past morpheme *=aka ~ ka*.

Note that the past form associated with 1st person is *=alaŋ* in all studied varieties and *=akaŋ ~ kaŋ* in RAP-13 (Polkim, Sarling, Syamrang, Yuiling, Santhali). I show the different constriction types attested in each scenario or configuration.

In this table, the symbol ∑ indicates the position of the verbal stem. The middle/reflexive morpheme =sʌ ~ si is bolded and underlined like the reciprocal form. The symbol {x} indicates that the morpheme may occur in either of these positions.

Table 150. Reflexive/reciprocal verbal argument indexation in non-past

Reflexive / Reciprocal (Non-Past)		
	pos.	**neg.**
1SG	∑-na-**ŋ**-**sʌ**	∑(-ma)-**ŋʌ**-**sʌ**-lʌ
1DU.INCL.REFL[62]		∑-ljam~lam
1DU.INCL.RECIP[63]		∑-**kaj**-ljam~lam
1DU.EXCL.REFL	∑-na-**ŋ**-**tɕʌ**	∑(-ma)-**ŋʌ**-**tɕʌ**-lʌ
1DU.EXCL.RECIP	∑-**kaj**-na-**ŋ**-**tɕʌ**	∑{-ma}-**kaj**{-ma}-**ŋʌ**-**tɕʌ**-lʌ
1PL.INCL.REFL[64]		∑-ljam~lam
1PL.INCL.RECIP[65]		∑-**kaj**-ljam~lam
1PL.EXCL.REFL	∑-na-**ŋ**-**su**	∑(-ma)-**ŋʌ**-**su**-lu
1PL.EXCL.RECIP	∑-**kaj**-na-**ŋ**-**su**	∑{-ma}-**kaj**{-ma}-**ŋʌ**-**su**-lu
2SG	∑-**te**-na-**sʌ**	∑(-ma)-**te**-**sʌ**-lʌ
2DU.REFL	∑-**te**-na-**dzʌ**	∑(-ma)-**te**-**dzʌ**-lʌ
2DU.RECIP	∑{-**te**}-**kaj**{-**te**}-na-**dzʌ**	∑{-ma}{-**te**}-**kaj**{-ma}{-**te**}-**dzʌ**-lʌ
2PL.REFL	∑-**te**-na-**i**-**si**	∑(-ma)-**te**-**i**-**si**-li
	∑-**te**-na-**i** (<intr.)	
2PL.RECIP	∑{-**te**}-**kaj**{-**te**}-na-**i**-**si**	∑{-ma}-**kaj**{-ma}-**te**-**i**-**si**-li
		∑{-ma}-**kaj**{-ma}-**te**-**si**-li (RAP-13)
3SG	∑-na-**sʌ**	∑(-ma)-**sʌ**-lʌ
3DU.REFL	∑-na-**tɕʌ**	∑(-ma)-**tɕʌ**-lʌ
3DU.RECIP	∑-**kaj**-na-**tɕʌ**	∑{-ma}-**kaj**{-ma}-**tɕʌ**-lʌ
3PL.REFL	∑-na-**i**-**si**	∑(-ma)-**i**-**si**-li
		∑(-ma)-**i**-li (<intr.)
3PL.RECIP	∑-**kaj**-na-**i**-**si**	∑{-ma}-**kaj**{-ma}-**i**-**si**-li
		∑(-ma)-**i**-li (<intr.)

---

[62] For 1st person dual inclusive, jussive imperative form applies, like with negation.

[63] For 1st person dual inclusive, jussive imperative form applies, like with negation.

[64] For 1st person plural inclusive, jussive imperative form applies, like with negation.

[65] For 1st person plural inclusive, jussive imperative form applies, like with negation.

Table 151. Reflexive/reciprocal verbal argument indexation in past

Reflexive / Reciprocal (Past)		
	pos.	**neg.**
1SG	∑-**alaŋ**-<u>**sʌ**</u>	∑(-ma)-**ŋʌ**-<u>**sʌ**</u>-lʌ
	∑-**akaŋ~kaŋ**-<u>**sʌ**</u> (POL)	
1DU.INCL.REFL	∑-a-**tʌhj**-<u>**tɕʌ**</u>	∑(-ma)-**tʌhj**-<u>**tɕʌ**</u>-lʌ
1DU.INCL.RECIP	∑-<u>**kaj**</u>-a-**tʌhj**-<u>**tɕʌ**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**tʌhj**-<u>**tɕʌ**</u>-lʌ
1DU.EXCL.REFL	∑-**alaŋ**-<u>**tɕʌ**</u>	∑(-ma)-**ŋʌ**-<u>**tɕʌ**</u>-lʌ
	∑-**akaŋ~kaŋ**-<u>**tɕʌ**</u> (POL)	
1DU.EXCL.RECIP	∑-<u>**kaj**</u>-**alaŋ**-<u>**tɕʌ**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**ŋʌ**-<u>**tɕʌ**</u>-lʌ
	∑-<u>**kaj**</u>-**akaŋ~kaŋ**-<u>**tɕʌ**</u> (POL)	
1PL.INCL.REFL	∑-a-**tʌh**-<u>**i**</u>	∑(-ma)-**tʌh**-**i**-li
1PL.INCL.RECIP	∑-<u>**kaj**</u>-a-**tʌh**-<u>**i**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**tʌh**-**i**-li
1PL.EXCL.REFL	∑-**alaŋ**-<u>**su**</u> (<intr.)	∑(-ma)-**ŋʌ**-<u>**su**</u>-lu (<intr.)
	∑-**alaŋ**-i-<u>**si**</u> (*arch.*)	∑(-ma)-**ŋ**-i-<u>**si**</u>-li (*arch.*)
	∑-**akaŋ~kaŋ**-i-<u>**si**</u> (*arch.*) (POL)	
1PL.EXCL.RECIP	∑-<u>**kaj**</u>-**alaŋ**-<u>**su**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**ŋʌ**-<u>**su**</u>-lu
	∑-<u>**kaj**</u>-**akaŋ~kaŋ**-<u>**su**</u> (POL)	
2SG	∑-**te**-ka-<u>**sʌ**</u>	∑(-ma)-**te**-<u>**sʌ**</u>-lʌ
2DU.REFL	∑-**te**-ka-<u>**dzʌ**</u>	∑(-ma)-**te**-<u>**dzʌ**</u>-lʌ
2DU.RECIP	∑{-**te**}-<u>**kaj**</u>{-**te**}-ka-<u>**dzʌ**</u>	∑{-ma}{-**te**}-<u>**kaj**</u>{-ma}{-**te**}-**dzʌ**-lʌ
2PL.REFL	∑-**te**-ka(-**i**)-<u>**si**</u>	∑(-ma)-**te**-i-<u>**si**</u>-li
	∑-**te**-ka-**i** (<intr.)	∑(-ma)-**te**-i-li (<intr.)
		∑(-ma)-**t**-i-li (<intr.) (POL)
2PL.RECIP	∑{-te}-<u>**kaj**</u>{-**te**}-ka(-**i**)-<u>**si**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**te**(-**i**)-<u>**si**</u>-li
	∑-<u>**kaj**</u>-**te**-ka-**i** (<intr.)	
3SG	∑-ka-<u>**sʌ**</u>	∑(-ma)-<u>**sʌ**</u>-lʌ
3DU.REFL	∑-ka-<u>**tɕʌ**</u>	∑(-ma)-<u>**tɕʌ**</u>-lʌ
3DU.RECIP	∑-<u>**kaj**</u>-ka-**tɕʌ**	∑{-ma}-<u>**kaj**</u>{-ma}-**tɕʌ**-lʌ
3PL.REFL	∑-ka-**i**-<u>**si**</u>	∑(-ma)-**i**-<u>**si**</u>-li
		∑(-ma)-**i**-li (<intr.)
3PL.REFL	∑-<u>**kaj**</u>-ka-**i**-<u>**si**</u>	∑{-ma}-<u>**kaj**</u>{-ma}-**i**-<u>**si**</u>-li
		∑(-ma)-**i**-li (<intr.)

5.8.3.2.2. Reflexive/reciprocal remote past paradigm

The argument indexation forms found in reflexive and reciprocal constructions in remote past are presented in Table 152. The remote past marker is *=to*.

In this table, the symbol ∑ indicates the position of the verbal stem. The middle/reflexive morpheme *=sʌ ~ si* is bolded and underlined like the reciprocal form. The symbol {*x*} indicates that the morpheme may occur in either of these positions.

Table 152. Reflexive/reciprocal verbal argument indexation in remote past

Reflexive / Reciprocal (Remote Past)		
	pos.	**neg.**
1SG	∑-**ŋʌ**-**s̲ʌ**-to	∑(-ma)-jak-**ŋʌ**-**s̲ʌ**-lʌ
1DU.INCL.REFL	∑-**tʌhj**-**t̲ɕ̲ʌ**-to	∑(-ma)-jak-**tʌhj**-**t̲ɕ̲ʌ**-lʌ
1DU.INCL.RECIP	∑-**kaj**-**tʌhj**-**t̲ɕ̲ʌ**-to	∑{-ma}-jak-**kaj**{-ma}-tʌhj-**t̲ɕ̲ʌ**-lʌ
1DU.EXCL.REFL	∑-**ŋʌ**-**t̲ɕ̲ʌ**-to	∑(-ma)-jak-**ŋʌ**-**t̲ɕ̲ʌ**-lʌ
1DU.EXCL.RECIP	∑-**kaj**-**ŋʌ**-**t̲ɕ̲ʌ**-to	∑{-ma}-jak-**kaj**{-ma}-**ŋʌ**-**t̲ɕ̲ʌ**-lʌ
1PL.INCL.REFL	∑-**tʌh**-**i**-**s̲i**-to	∑(-ma)-jak-**tʌh**-**i**-**s̲i**-li
1PL.INCL.RECIP	∑-**kaj**-**tʌh**-**i**-to	∑{-ma}-jak-**kaj**{-ma}-**tʌh**-**i**-li
1PL.EXCL.REFL	∑-**ŋ**-**i**-**s̲i**-to	∑(-ma)-jak-**ŋ**-**i**-**s̲i**-li (arch.)
1PL.EXCL.RECIP	∑-**kaj**-**ŋʌ**-**s̲u**-to (<tr.)	∑{-ma}-jak-**kaj**{-ma}-**ŋʌ**-**s̲u**-lu
	∑-**kaj**-**ŋ**-**i**-**s̲i**-to (<1pl.excl)	
2SG	∑-**te**-**s̲ʌ**-to	∑(-ma)-jak-**te**-**s̲ʌ**-lʌ
2DU.REFL	∑-**te**-**d̲z̲ʌ**-to	∑(-ma)-jak-**te**-**d̲z̲ʌ**-lʌ
2DU.RECIP	∑-**kaj**-**te**-**d̲z̲ʌ**-to	∑{-ma}-jak-**kaj**{-ma}{-te}-**d̲z̲ʌ**-lʌ
2PL.REFL	∑-**te**-**i**-**s̲i**-to	∑(-ma)-jak-**te**-**i**-**s̲i**-li
	∑-**te**-**s̲i**-to (POL)	∑(-ma)-jak-**te**-**s̲i**-li (POL)
2PL.RECIP	∑-**kaj**-**te**-**i**-to	∑{-ma}-jak-**kaj**{-ma}-**te**-**i**-li
	∑-**kaj**-**t**-**i**-to (POL)	∑{-ma}-jak-**kaj**{-ma}-**t**-**i**-li (POL)
3SG	∑-**s̲ʌ**-to	∑(-ma)-jak-**s̲ʌ**-lʌ
3DU.REFL	∑-**t̲ɕ̲ʌ**-to	∑(-ma)-jak-**t̲ɕ̲ʌ**-lʌ
3DU.RECIP	∑-**kaj**-**t̲ɕ̲ʌ**-to	∑{-ma}-jak-**kaj**{-ma}-**t̲ɕ̲ʌ**-lʌ
3PL.REFL	∑-**i**-**s̲i**-to	∑(-ma)-jak-**i**-**s̲i**-li
3PL.RECIP	∑-**kaj**-**i**-to	∑{-ma}-jak-**kaj**{-ma}-**i**-li

5.8.3.2.3. Reflexive/reciprocal irrealis paradigm

The argument indexation forms found in reflexive and reciprocal constructions in irrealis are presented in Table 153. The irrealis marker is *=tɕja* in Lothar and *=tɕa* in Manahari.

In this table, the symbol ∑ indicates the position of the verbal stem. The middle/reflexive morpheme *=sʌ ~ si* is bolded and underlined like the reciprocal form. The symbol {*x*} indicates that the morpheme may occur in either of these positions.

As one can see in Table 153, only 1st person exclusive and 3rd person forms may be used as an assertion or a question, and other forms either express a question, noted (Q), or an assertation, noted (A), but not both.

Except with 1st person singular, questions are formed with the non-past tense morpheme *=na* and followed by the morpheme *ja* 'or' which developped the function of interrogative marker.

Assertions with 1st person inclusive carry a negative imperative or jussive function, as is the case with 2nd person. Such constructions, if used with negative morphology featuring the inclusive marker in its construction, the negative expression would only refer to the past tense, by contrast with other negative constructions which keep their irrealis modality.

Table 153. Reflexive/reciprocal verbal argument indexation in irrealis

Reflexive / Reciprocal (Irrealis)		
	pos.	**neg.**
1SG	∑-tɕja-**ŋ**-**sʌ** (Q)	∑(-ma)-**ŋʌ**-**sʌ**-lʌ
1DU.INCL	∑-tɕja-**tʌhj**-**tɕʌ** (A) (IMP.NEG)	n/a
1DU.EXCL	∑-tɕja-**ŋ**-**tɕʌ**	∑(-ma)-**ŋʌ**-**tɕʌ**-lʌ
1PL.INCL	∑-tɕja-**tʌh**-**i** (A) (IMP.NEG)	n/a
1PL.EXCL	∑-tɕja-**ŋ**-**su**	∑(-ma)-**ŋʌ**-**su**-lu
2SG	∑-**te**-na-**sʌ** ja (Q)	∑(-ma)-**te**-**sʌ**-lʌ
	∑-**te**-tɕja-**sʌ** (A) (IMP.NEG)	
2DU.REFL	∑-**te**-na-**dzʌ** ja (Q)	∑(-ma)-**te**-**dzʌ**-lʌ
	∑-**te**-tɕja-**dzʌ** (A) (IMP.NEG)	
2DU.RECIP	∑-**kaj**-**te**-na-**dzʌ** ja (Q)	∑(-ma)-**te**-**dzʌ**-lʌ
	∑-**te**-tɕja-**dzʌ** (A) (IMP.NEG)	
2PL.REFL	∑-**te**-na(-**i**)-**si** ja (Q)	∑(-ma)-**te**-**i**-**si**-li
	∑-**te**-na ja (Q) (<intr.)	∑(-ma)-**te**-**si**-li (POL)
	∑-**te**-tɕja-**i**-**si** (A) (IMP.NEG)	
	∑-**te**-tɕja-**i** (A) (IMP.NEG)	
2PL.RECIP	∑-**kaj**-**te**-na-**i**-**si** ja (Q)	∑(-ma)-**te**-**i**-**si**-li
	∑-**kaj**-**te**-na-**si** ja (Q)	∑(-ma)-**te**-**si**-li (POL)
	∑-**kaj**-**te**-na ja (Q) (<intr.)	
	∑-**te**-tɕja-**i**-**si** (A) (IMP.NEG)	
	∑-**te**-tɕja-**i** (A) (IMP.NEG)	
3SG	∑-tɕja-**sʌ**	∑(-ma)-**sʌ**-lʌ
3DU	∑-tɕja-**tɕʌ**	∑(-ma)-**tɕʌ**-lʌ
3PL	∑-**i**-**si**	∑(-ma)-**i**-**si**-li

5.8.3.2.4.   Reflexive/reciprocal optative paradigm

The argument indexation forms found in reflexive and reciprocal constructions in optative are presented in Table 154. The optative marker is =*pa*.

In this table, the symbol ∑ indicates the position of the verbal stem. The middle/reflexive morpheme =*sʌ* ~ *si* is bolded and underlined like the reciprocal form. The symbol {*x*} indicates that the morpheme may occur in either of these positions.

Table 154. Reflexive/reciprocal verbal argument indexation in optative

Reflexive / Reciprocal (Optative)	
	pos.
1SG	∑-pa-ŋ-**s̲ʌ̲**
1DU.INCL.REFL	∑-pa-tʌhj-**t̲ɕ̲ʌ̲**
1DU.EXCL.REFL	∑-pa-ŋ-**t̲ɕ̲ʌ̲**
1DU.RECIP	∑-**k̲a̲j̲**-pa-ŋ-**t̲ɕ̲ʌ̲**
1PL.INCL.REFL	∑-pa-tʌh-i
1PL.EXCL.REFL	∑-pa-ŋ-**s̲u̲**
1PL.RECIP	∑-**k̲a̲j̲**-pa-ŋ-**s̲u̲**
2SG	∑-**t̲e̲**-pa-**s̲ʌ̲**
2DU.REFL	∑-**t̲e̲**-pa-**d̲z̲ʌ̲**
2DU.RECIP	∑-**k̲a̲j̲**-**t̲e̲**-pa-**d̲z̲ʌ̲**
2PL.REFL	∑-pa(-i)-**s̲i̲**
2PL.RECIP	∑-**k̲a̲j̲**-**t̲e̲**-pa(-i)-**s̲i̲**
3SG	∑-pa-**s̲ʌ̲**
3DU.REFL	∑-pa-**t̲ɕ̲ʌ̲**
3DU.RECIP	∑-**k̲a̲j̲**-pa-**t̲ɕ̲ʌ̲**
3PL.REFL	∑-pa(-i)-**s̲i̲**
3PL.RECIP	∑-**k̲a̲j̲**-pa(-i)-**s̲i̲**

5.8.3.2.5.   Reflexive/reciprocal imperative paradigm

The argument indexation forms found in reflexive and reciprocal constructions in imperative are presented in Table 155.

In this table, the symbol ∑ indicates the position of the verbal stem. The middle/reflexive morpheme =sʌ ~ si is bolded and underlined like the reciprocal form. The symbol {x} indicates that the morpheme may occur in either of these positions.

Table 155. Reflexive/reciprocal verbal argument indexation in imperative

	Reflexive / Reciprocal (Imperative)
	pos.
1SG	∑-tɕja-ŋ-<u>**sʌ**</u> (Q)
	∑-pa-ŋ-<u>**sʌ**</u>
1DU.REFL	∑-<u>**tɕʌ**</u>
1DU.RECIP	∑-<u>**kaj**</u>-<u>**tɕʌ**</u>
1PL.REFL	∑-i-<u>**si**</u>
1PL.RECIP	∑-<u>**kaj**</u>-i
2SG	∑-<u>**sʌ**</u>
	∑-<u>**te**</u>-pa-<u>**sʌ**</u>
2DU.REFL	∑-<u>**dzʌ**</u>
2DU.RECIP	∑-<u>**kaj**</u>-<u>**dzʌ**</u>
2PL.REFL	∑-<u>**n**</u>-i-<u>**si**</u>
2PL.RECIP	∑-<u>**kaj**</u>-<u>**nʌ**</u>
3SG	∑-pa-<u>**sʌ**</u>
3DU.REFL	∑-pa-<u>**tɕʌ**</u>
3DU.RECIP	∑-<u>**kaj**</u>-pa-<u>**tɕʌ**</u>
3PL.REFL	∑-pa-i-<u>**si**</u>
3PL.RECIP	∑-<u>**kaj**</u>-pa-i

5.8.3.2.6. Observations on reflexive =sʌ ~ si and dual forms

Intransitive and reflexive/reciprocal constructions exhibit the same person markers, and the marking of dual number shows a complementary distribution with the reflexive morpheme =sʌ ~ si, as presented in Table 156. That is, the morpheme =sʌ ~ si does not occur with the dual forms, only with singular and plural forms. This may suggest either the loss of the reflexive form =sʌ in dual configurations, in which case, the dual morphemes =tɕʌ 1/3DU and =dzʌ 2DU became portmanteau marking both a dual and reflexive function, or that one can posit a common functional origin for reflexive and dual markers, namely reflexivity. In this case the 2nd and 3rd person reflexive markers =dzʌ and =tɕʌ would have developed the function of dual. Since these dual forms have

cognates across TH languages, this development would have occurred at a higher-level clade.

Table 156. Argument indexation with intransitive and middle/reflexive/reciprocal

S		INTRANSITIVE		REFL/RECIP	REFL
1(SG)		-ŋ		-ŋ	**-sʌ**
1(SG).NEG		-ŋʌ		-ŋʌ	**-sʌ**
1/3DU		-ŋ-**tɕʌ** 1DU / -**tɕʌ** 3DU		-ŋ-**tɕʌ** 1DU / -**tɕʌ** 3DU	
1PL.EXCL	∑-TNS	-su (<tr.)	∑(-**kaj**)-TNS	-su RECIP (<tr.)	
		-i (arch. intr.)		-i REFL/RECIP	**-si**
2/3PL		-i		-i	**-si**
2(SG)		te-		te-	**-sʌ**
2DU		**-dʑʌ**		**-dʑʌ**	
1DU.INCL		**-tʌhj-tɕʌ**		**-tʌhj-tɕʌ**	
1PL.INCL		-tʌh-i		-tʌh-i	**-si**

5.8.3.3. Chepang regarding direct-inverse systems in TH and beyond

Direct-inverse systems are cross-linguistically described as marking both the syntactic role and hierarchical ranking of the participant indexed on (di-)transitive verbs through the presence of indexation and of a special morpheme called "inverse." In an inverse configuration, the argument indexed is hence indicated being acted upon and ranking higher than the A argument (Comrie 1980; DeLancey 1981; Nichols 1992a; Givón 1994; Zuñiga 2006; Bickel & Nichols 2007; Zúñiga 2014; Jacques & Antonov 2014). Such a system is traditionally described as functionally arising from a referential hierarchy (Silverstein 1976; DeLancey 1981; Comrie 1980; Comrie 1981; Klaiman 1991; Nichols 1992a; Siewierska 1998; Corbett 2000; Song 2001; Siewierska 2004; Zuñiga 2006; Bickel 2008; Lockwood & Macaulay 2012). Direct-inverse systems are also categorized as a subtype of hierarchical alignment (Nichols 1992a).

Jacques and Antonov (2014: 302–303) define canonical direct-inverse systems based on three criteria:

First, in such a system, all person–number markers are neutral with regard to syntactic roles (S, A, and O). Second, the ambiguity which this entails (especially in mixed scenarios) is resolved by way of obligatory (and mutually exclusive) markers, called *direct* (in the case of 1→2, SAP→3, 3'→3) and *inverse* (in the case of 2→1, 3→SAP, and 3'→3), respectively. These markers do not appear on intransitive verbs. This property is generally described in terms of referential hierarchies. Third, inverse verb forms, unlike verb forms in a passive construction, do not undergo valency changes (the verb does not become intransitive when an inverse marker is added), and the arguments keep the same syntactic properties (such as case marking and pivot accessibility) as in the direct construction.

These criteria are summarized as follows: (1) person and number markers do not syntactically or morphologically distinguish their semantic or syntactic roles (S, A, or O argument); (2) an additional morpheme (*direct* or *inverse*) clarifies the syntactic role of the arguments based on an assumed referential hierarchy (1>2>3'>3); (3) an inverse construction does not affect the syntactic valency of the construction.

Direct-inverse systems are attested for instance in the Americas (Algonquian, Mapudungun, Sahaptian, Mixe-Zoquean, Cariban), the Caucasus (Northeast and Northwest Caucasian), and in the Himalayas (TH languages).

In TH, direct-inverse systems occur in rGyalrongic (DeLancey 1981; Sun & Shidanluo 2002; Jacques 2010; Gong 2014; Jacques & Antonov 2014); Nocte (Das Gupta 1971; DeLancey 1981); a number of Eastern Kiranti languages, like Bantawa (Ebert 1997a; Doornenbal 2009), Athpare (Ebert 1991; Ebert 1997a), Camling (Ebert 1997b), Belhare (Ebert 1991; Bickel 1995), and Dungmali (Ebert 1991); Western Kiranti, like Khaling (Toba 1989; Driem 1993b; Jacques et al. 2012; Jacques & Antonov 2014) and Dumi (Driem 1993b); Rawang (Barnard 1934; DeLancey 1981; Driem 1993b; LaPolla 2007) and Dulong (Trung) (Driem 1993b).

Chepang may be described as typologically exhibiting a non-canonical direct-inverse system. By non-canonical, I mean that it shows differences regarding the distribution of direct and inverse morphosyntactic features from what is expected in what is described as a canonical direct-inverse system, as described for instance by Jacques & Antonov (2014: 302–303).

The non-canonicality of Chepang's direct-inverse system is observed at the level of the morphosyntactic constructions and the multiplicity of possible constructions in identical scenarios. The inverse morpheme *=ta ~ tʌ ~ tʰʌ* occurs in the local (2SG>1NSG) scenario with the marking of both the A for person, and the O for person and non-singular number, as in (744); in the mixed (3>1) scenario with the marking of the O for person and non-singular number, as in (745); and in the non-local (3>3) scenario with the marking of the O for number, as in (746).

Morphosyntactically, the presence of the inverse marker and the indexation of both the A and O in 2SG>1NSG scenario is not expected, as in (744); and inverse constructions are not expected either in 3SG>3SG scenario, as in (746). In addition, in 2>1SG, 2NSG>1 and 3>2 scenarios different constructions are used. In the 2SG>1SG scenario for example, both the marking of A for person and a morpheme indexing 2SG>1SG are used, as in (747). In 3>2SG, only the marking of 2nd person is present, as expected in a canonical hierarchical alignment system, as shown in (748).

(744)   *aʔl=te=na=ta=ŋ=tɕʌ.*
        bring=2=NPST=INV=1=1/3DU

        'You will bring the two of us (there).'
        CH_MKW _PC_SIL_ E

(745)   *aʔl=na=ta=ŋ=tɕʌ.*
        bring=NPST=INV=1=1/3DU

        'S/he/they will bring the two of us (there).'
        CH_MKW _PC_SIL_ E

(746) *aʔl=na=ta=i.*
bring=NPST=INV=3>3SG

'S/he/they will bring him (there).'
CH_MKW _PC_SIL_ E

(747) *aʔl=te=na=tɕi.*
bring=2=NPST=2SG>1SG

'You will bring me (there).'
CH_MKW _PC_SIL_ E

(748) *aʔl=te=na.*
bring=2=NPST

'S/he/they will bring you (there).'
CH_MKW _PC_SIL_ E

In addition, the construction in (744) with the 2SG>1NSG scenario, although considered archaic, is possible, along with two other constructions: one indexing the O for person and number in absence of inverse marker, as shown in (749); and a direct construction indexing the A for person, as in (750). These latter two occur in questions and assertions, but only the construction in (750) can be negated. They are distinguished through the attitude of the speaker towards the process expressed. In (749), the speaker expresses a request, their desire for the process to happen, or their belief that it will indeed happen. In (750), the speaker acknowledges the intention of the addressee for the process to happen. As questions, (751) and (752) respectively ask for the speaker's request to be fulfilled already, and for the addressee's confirmation regarding their intention.

(749) *aʔl=na=ŋ=tɕʌ.*
bring=NPST=1=1/3DU

'You will bring the two of us (there).'
CH_MKW _PC_SIL_ E

(750)   *aʔl=te=na=u.*
bring=2=NPST=3O/DIR

'You will bring me/the two of us/us (there).'

CH_MKW _PC_SIL_ E

(751)   *aʔl=na=ŋ=tɕʌ ?*
bring=NPST=1=1/3DU

'Will you bring the two of us (there)?'

CH_MKW _PC_SIL_ E

(752)   *aʔl=te=na=u ?*
bring=2=NPST=3O/DIR

'Will you bring me/the two of us/us (there)?'

CH_MKW _PC_SIL_ E

Such a complex direct-inverse system is not found in other TH languages. In TH, a system closer to a canonical direct-inverse system is found in rGyalrongic (Jacques & Antonov 2014: 305). In Zbu rGyalrong (Gong 2014), for instance, the inverse marker occurs in 3>1 scenarios where the O argument is indexed for person and/or number, and in 3>2 and 2>1 scenarios where the O arguments are indexed for person and/or number, in addition to marking the A argument as well in 2>1. These later two scenarios show the marking of 2nd person regardless of its syntactic role in combination with the inverse marker. Despite some discrepancies, this is representative of a canonical system where inverse morphology occurs in 2/3>1 and 3>2 scenarios.

Another fairly canonical direct-inverse system is found in Nocte with the presence of an inverse marker in expected scenarios, i.e., in 3>1/2 and 2>1 with indexation of the O argument (DeLancey 1981).

The Eastern Kiranti languages, Bantawa, Camling, Athpare, Belhare, and Dungmali are analyzed as featuring direct-inverse systems (Ebert 1991; 1997a; 1997b; Bickel 1995). However, Doornenbal (2009) does not consider that Bantawa exhibits a direct-inverse system on the grounds that for instance, inverse morphology is lacking in

expected scenarios, such as 2>1, and that direct and inverse morphology combine in 3DU>3. Such a system if analyzed as direct-inverse, may be considered non-canonical.

Finally, Khaling, a Western Kiranti language, also displays a non-canonical direct-inverse system (Toba 1989; Driem 1993b; Jacques et al. 2012; Jacques & Antonov 2014) featuring a morpheme marking both 2nd person in intransitive and (di)transitive configurations and inverse. The same type of identical marking of 2nd person and inverse is also present in Rawang (Barnard 1934; DeLancey 1981; Driem 1993b; LaPolla 2007), Dumi, and Dulong (Trung) (Driem 1993b).

Chepang's direct-inverse system questions once again (Filimonova 2005; Bickel 2008; Cristofaro 2013; Gildea & Zúñiga 2016; Witzlack-Makarevich et al. 2016) the functional validity of referential hierarchies (DeLancey 1981; Nichols 1992b; Siewierska 1998; Siewierska 2004; Zuñiga 2006; Lockwood & Macaulay 2012). Specifically, when it comes to argument indexation in local scenarios, diverging traits can for the most part be explained as the result of historical changes, through internal and external comparison, and of sociopragmatic forces at play in the speaker's choice to use such or such construction.

5.8.3.4.   Summary of Chepang's indexation system

As introduced in § 5.8.3.3, the morphosyntactic constructions used to express a (di)transitive process may be multiple within a single scenario or may differ from expected patterns of a direct-inverse argument indexation system. Chepang transitive and ditransitive verbal argument indexation can be broken down into five main construction types, as follows:

- direct   > indexation of A +/- direct marker
- inverse   > indexation of O/R + inverse marker
- inverse/mix   > indexation of A, O/R + inverse marker
- mix   > indexation of A and O/R (variation in number)
- object/anticausative   > indexation of O/R

These construction types are presented within the scenarios where they apply in Table 157.

The higher number of construction types is observed for 2nd person acting on 1st person in local scenario (2>1), i.e., four of the five construction types: two mix constructions, an object construction, a direct construction, and an inverse/mix construction.

In 1st person acting on 2nd person in local scenario (1>2), two construction types are used: mix and direct. These constructions further show the use of distinct constructions.

In 2nd person and 1st person acting on 3rd person in mixed scenarios (1/2>3), only direct constructions are observed.

In 3rd person acting on 1st and 2nd persons in mixed scenario (3>1/2), several construction types are observed. The inverse construction type occurs in 3>1 and conforms to a referential hierachy where 1st person ranks higher than 3rd person. However, when it comes to 1st person dual or plural inclusive, only inclusive forms are used and not inverse morphology. In 3>2, object/anticausative and direct construction types can be used. The variety of constructions in the 3>1/2 scenario shows that a referential hierarchy is insufficient to explain such patterns.

Finally, in 3rd person acting on 3rd person, both inverse and direct construction types are attested. The inverse construction distribution conforms with a referential hierarchy where inverse marking occurs when a human referent is acted upon by a non-human referent. The direct construction may, however, be used regardless of humanness of the referent, which is unexpected and underlies the fact that the motive behind the use of inverse morphology is not a priori based on an assumed referential hierarchy.

Chepang's non-canonical direct-inverse system can be seen as dynamic in the sense that it is clearly attached to pragmatic forces behind which the different constructions are chosen to be used or not by the speaker within the frame of the scenarios to which they are attached.

This has several implications in the way we can envisage to reconstructing the indexation system at the level of Proto-Chepang (PC), Proto-Chepang-Bhujel (PCB), and beyond, since such a system has so often been described as non-dynamic. The idea that

pragmatic forces are at play in the use of indexation forms is very recent (Heath 1991; Stirling & Manderson 2011; Gast et al. 2015; DeLancey 2018; Konnerth & Sansò 2021; Konnerth 2021) and Chepang is the first TH language described as exhibiting such a system. It is possible that the presence of a variety of constructions within a single scenario exists in other TH languages, and that earlier descriptions of indexation systems have not recognized this pattern. Indeed, such constructions could be perceived as the result of a simplification by the speaker rather than meaningful alternative expressions.

In Chepang for instance, Caughley (1978) reports mix forms in 1>2 local scenarios that were not kept in his final description of the paradigms (1982), and does not mention the direct forms that can also be used alongside inverse forms in 3>2 mix scenarios or alongside the mix and inverse forms used in 2>1 local scenarios.

Table 157. Summary of Chepang (di)transitive indexation system

	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
1SG	REFL RECIP			**I – MIX** 1>2 / 1A / 2O.NUM 1>2 / 1A / 1A.NUM.DIR 1>2 / 1A / 1A.NUM **II – DIRECT** 1A 1A / 1A.NUM.DIR			**I – DIRECT** 1A 1A / 1A.NUM.DIR		
1DU.INCL (2)									
1DU.EXCL (TO OTHER)									
1PL.INCL (3+GROUP)									
1PL.EXCL (TO OTHER)									

	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
2SG	**I – MIX (2SG>1SG) [NPST/PST]** 2A / 2>1 **II – MIX (2DU>1NPL; 2PL>1) [NPST]** 1O / 2A.NUM **III – OBJECT (2SG>1NSG; 2DU>1PL) [NPST]** 1O / 1O.NUM **IV – DIRECT [NPST/PST]**			REFL RECIP			**I – DIRECT** 2A / 2A.NUM.DIR		
2DU									
2PL									

594

	2A / 2A.NUM.DIR **V – INVERSE/MIX (2SG>1NSG) [NPST/PST]** 2A / INV *ta* / 1O.NSG / 1O.NUM								
	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
3SG REFL 3DU REFL RECIP 3PL REFL RECIP							REFL RECIP		
	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
3SG 3DU 3PL	**I – INVERSE** INV *ta* / 1O INV *ta* / 1O / 1O.NUM INV.INCL *tʌhj, tʌh* / 1O.NUM (REF. HIER. 1>3)			**I – OBJECT** 2O 2O / 2O.NUM **II – DIRECT** 3A.NUM.DIR			**I – INVERSE** INV *ta ~ tʌ ~ tʰʌ* / 3O.NUM 3hum>3hum; 3nhum>3hum *3hum>3nhum (REF. HIER. 3HUM>3NHUM) **II – DIRECT** 3A.NUM.DIR 3hum/3nhum>3hum/3nhum		

5.8.3.5.   Transitive and ditransitive verbal argument indexation

In this section, I describe argument indexation attested with transitive and ditransitive paradigms. All the paradigms are presented according to the scenario within which they are found, organized by which person is acting upon another.

Examples of verbs that can take (di)transitive morphology are given in Table 158. The following sub-sections presenting the attested complete paradigms are as follows:

Table 158. Examples of verbs found with (di)transitive morphology

transitive verb	meaning
aʔl-	'take away'
boŋ-	'look for'
ʣahŋ-	'do, make'
ʣe-	'eat'
kʌn-	'look at'
lat-	'carry'
leʔ-	'take, buy'
lʰo-	'chop off'
pʰe-	'leave, abandon'
ɾa-	'cut (weed), saw'
sat-	'kill'
waʔn-	'bring'

ditransitive verb	meaning
bʌj-	'give'
kas-	'feed'
tan-	'show'
te-	'ask for, beg'

5.8.3.5.1. (Di)transitive non-past and past paradigm

The argument indexation forms found in (di)transitive constructions in non-past and past are presented together in the following sub-sections.

The non-past morpheme is =*na* and the past morpheme attested with (di)transitive verbs is =*ak ~ aka ~ ka*. Note that the past form associated with 1st person is =*alaŋ* in all studied varieties and =*akaŋ ~ kaŋ* in RAP-13 (Polkim, Sarling, Syamrang, Yuiling, Santhali). I show the different constriction types attested in each scenario or configuration.

In the following tables, the symbol ∑ indicates the position of the verbal stem and TM that of the tense markers. All indexation forms are bolded. The symbol {*x*} indicates that the morpheme may occur in either of these positions, and the symbol (*x*) that the morpheme remains optional. The negative forms occur below the tilde ~. When the forms are not attested because they do not correspond to any possible scenario, it is noted by n/a.

5.8.3.5.1.1. 1SG > 2 / 3

1SG > 2/3	2SG	2DU	2PL	3SG	3DU	3PL
1SG	I − MIX ∑-**ne**-TM-**ŋ**[66] ∑-**tɕe**-TM-**ŋ** ~ ∑(-ma)-**ne**-**ŋ**ʌ-lʌ ∑(-ma)-**tɕe**-**ŋ**ʌ-lʌ II − DIRECT ∑-TM-**ŋ** ~ ∑(-ma)-**ŋ**ʌ-lʌ	I − MIX ∑-**ne**-TM-**ŋ**-**dʑ**ʌ ∑-**tɕe**-TM-**ŋ**-**dʑ**ʌ ~ ∑(-ma)-**ne**-**ŋ**ʌ-**dʑ**ʌ-lʌ ∑(-ma)-**tɕe**-**ŋ**ʌ-**dʑ**ʌ-lʌ II − DIRECT ∑-TM-**ŋ** ~ ∑(-ma)-**ŋ**ʌ-lʌ	I − MIX ∑-**ne**-TM-**ŋ**-**s**ʌ ∑-**tɕe**-TM-**ŋ**-**s**ʌ ~ ∑(-ma)-**ne**-**ŋ**ʌ-**s**ʌ-lʌ ∑(-ma)-**tɕe**-**ŋ**ʌ-**s**ʌ-lʌ II − DIRECT ∑-TM-**ŋ** ~ ∑(-ma)-**ŋ**ʌ-lʌ	I − DIRECT ∑-TM-**ŋ** ~ ∑(-ma)-**ŋ**ʌ-lʌ		

---

[66] The 1>2 morpheme =*ne* is not attested in RAP-13: Polkim, Sarling, Syamrang, Yuiling, Santhali.

5.8.3.5.1.2. 1DU > 2 / 3

1DU > 2/3	2SG	2DU	2PL	3SG	3DU	3PL
1DU.INCL (2)	n/a	n/a	n/a	I – DIRECT ∑-TM-**tʌhj-tɕ-u** ~ ∑(-ma)-**tʌhj-tɕ-u**-lu[67]		
1DU.EXCL (TO OTHER)	I – MIX ∑-**ne**-TM-**ŋ-tɕ-u** ∑-**tɕe**-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-**tɕe-ŋʌ-tɕ-u**-lu II – DIRECT ∑-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ŋʌ-tɕ-u**-lu ∑(-ma)-**ŋʌ-tɕʌ**-lʌ (<intr.)	I – MIX ∑-**ne**-TM-**ŋ-tɕ-u** ∑-**tɕe**-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-**tɕe-ŋʌ-tɕ-u**-lu I – MIX ∑-**ne**-TM-**ŋ-dzʌ** ∑-**tɕe**-TM-**ŋ-dzʌ** ~ ∑(-ma)-**ne-ŋʌ-dzʌ**-lʌ ∑(-ma)-**tɕe-ŋʌ-dzʌ**-lʌ I – MIX ∑-**ne**-TM-**ŋ-tɕʌ** ∑-**tɕe**-TM-**ŋ-tɕʌ** ~ ∑(-ma)-**ne-ŋʌ- tɕʌ**-lʌ ∑(-ma)-**tɕe-ŋʌ- tɕʌ**-lʌ II – DIRECT ∑-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ŋʌ-tɕ-u**-lu ∑(-ma)-**ŋʌ-tɕʌ**-lʌ (<intr.)	I -MIX ∑-**ne**-TM-**ŋ-tɕ-u** ∑-**tɕe**-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-**tɕe-ŋʌ-tɕ-u**-lu I – MIX ∑-**ne**-TM-**ŋ-sʌ** ∑-**tɕe**-TM-**ŋ-sʌ** ~ ∑(-ma)-**ne-ŋʌ-sʌ**-lʌ ∑(-ma)-**tɕe-ŋʌ-sʌ**-lʌ II – DIRECT ∑-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ŋʌ-tɕ-u**-lu ∑(-ma)-**ŋʌ-tɕʌ**-lʌ (<intr.)	I – DIRECT ∑-TM-**ŋ-tɕ-u** ~ ∑(-ma)-**ŋʌ-tɕ-u**-lu		

---

[67] With 1DU.INCL acting on 3rd person, negation is only attested in past tense; jussive/imperative negation applies in non-past.

5.8.3.5.1.3. 1PL > 2 / 3

1PL > 2/3	2SG	2DU	2PL	3SG	3DU	3PL
1PL.INCL (3+GROUP)	n/a	n/a	n/a	I – DIRECT ∑-TM-**tʌhj-n-i** ~ ∑(-ma)-**tʌhj-n-i**-li[68]		
1PL.EXCL (TO OTHER)	I – MIX ∑-**ne**-TM-**ŋ-s-u** ∑-**tɕe**-TM-**ŋ-s-u** ~ ∑(-ma)-**ne-ŋʌ-s-u**-lu ∑(-ma)-**tɕe-ŋʌ-s-u**-lu II – DIRECT ∑-TM-**ŋ-s-u** ~ ∑(-ma)-**ŋʌ-s-u**-lu	I – MIX ∑-**ne**-TM-**ŋ-s-u** ∑-**tɕe**-TM-**ŋ-s-u** ~ ∑(-ma)-**ne-ŋʌ-s-u**-lu ∑(-ma)-**tɕe-ŋʌ-s-u**-lu I – MIX ∑-**ne**-TM-**ŋ-dʑʌ** ∑-**tɕe**-TM-**ŋ-dʑʌ** ~ ∑(-ma)-**ne-ŋʌ-dʑʌ**-lʌ ∑(-ma)-**tɕe-ŋʌ-dʑʌ**-lʌ II – DIRECT ∑-**ŋ-s-u** ~ ∑(-ma)-**ŋʌ-s-u**-lu	I – MIX ∑-**ne**-TM-**ŋ-s-u** ∑-**tɕe**-TM-**ŋ-s-u** ~ ∑(-ma)-**ne-ŋʌ-s-u**-lu ∑(-ma)-**tɕe-ŋʌ-s-u**-lu I – MIX ∑-**ne**-TM-**ŋ-sʌ** ∑-**tɕe**-TM-**ŋ-sʌ** ~ ∑(-ma)-**ne-ŋʌ-sʌ**-lʌ ∑(-ma)-**tɕe-ŋʌ-sʌ**-lʌ II – DIRECT ∑-**ŋ-s-u** ~ ∑(-ma)-**ŋʌ-s-u**-lu	I – DIRECT ∑-TM-**ŋ-s-u** ~ ∑(-ma)-**ŋʌ-s-u**-lu		

---

[68] With 1PL.INCL acting on 3rd person, negation is only attested in past tense; jussive/imperative negation applies in non-past.

5.8.3.5.1.4. 2sg > 1 / 3

2SG > 1/3	1SG	1DU	1PL	3SG	3DU	3PL
2SG	I – MIX (NPST/NPST) request/confirmation/question ∑-**te**-TM-**tɕi** ~ ∑(-ma)-**te**-**tɕi**-li II – DIRECT (NPST) realization/comment/question ∑-**te**-TM-**u** ~ ∑(-ma)-**te**-**u**-lu II – DIRECT (PST) realization/comment/question ∑-**te**-TM-**n** ~ ∑(-ma)-**te**-**u**-lu	I – OBJECT (NPST) request/confirmation/question ∑-TM-**ŋ**-**tɕʌ** ~ (*NEG)[69] II – DIRECT (NPST) realization/comment/question ∑-**te**-TM-**u** ~ ∑(-ma)-**te**-**u**-lu II – DIRECT (PST) realization/comment/question ∑-**te**-TM-**n** ~ ∑(-ma)-**te**-**u**-lu III – INVERSE/MIX (NPST/PST) ∑-**te**-TM-**ta**-**ŋ**-**tɕʌ** (*arch.*) ~ ∑(-ma)-**ta**-**ŋ**-**tɕʌ**-**lʌ**	I – OBJECT (NPST) request/confirmation/question ∑-TM-**ŋ**-**sʌ** ~ (*NEG)[70] II – DIRECT (NPST) realization/comment/question ∑-**te**-TM-**u** ~ ∑(-ma)-**te**-**u**-lu II – DIRECT (PST) realization/comment/question ∑-**te**-TM-**n** ~ ∑(-ma)-**te**-**u**-lu III – INVERSE/MIX (NPST/PST) ∑-**te**-TM-**ta**-**ŋ**-**i** (*arch.*) ~ ∑(-ma)-**ta**-**ŋ**-**i**-li	I – DIRECT (NPST) ∑-**te**-TM-**u** ~ ∑(-ma)-**te**-**u**-lu I – DIRECT (PST) ∑-**te**-TM-**n** ~ ∑(-ma)-**te**-**u**-lu		

---

[69] No specific negation occurs with this construction.

[70] No specific negation occurs with this construction.

5.8.3.5.1.5. 2DU > 1 / 3

2DU > 1/3	1SG	1DU	1PL	3SG	3DU	3PL
2DU	I – MIX (NPST) request/confirmation/question $\sum$-TM-**ŋ-dʑʌ** ~ (*neg)[71] II – DIRECT (NPST) realization/comment/question $\sum$-**te**-TM-**dʑ-u** ~ $\sum$(-ma)-**te-dʑ-u**-lu II – DIRECT (PST) realization/comment/question $\sum$-**te**-TM-**dʑ-u** ~ $\sum$(-ma)-**te-dʑ-u**-lu		I – OBJECT (NPST) request/confirmation/question $\sum$-TM-**ŋ-sʌ** ~ (*neg)[72] II – DIRECT (NPST) realization/comment/question $\sum$-**te**-TM-**dʑ-u** ~ $\sum$ (-ma)-**te-dʑ-u**-lu II – DIRECT (PST) realization/comment/question $\sum$-**te**-TM-**dʑ-u** ~ $\sum$(-ma)-**te-dʑ-u**-lu	I – DIRECT $\sum$-**te**-TM-**dʑ-u** ~ $\sum$(-ma)-**te-dʑ-u**-lu		

---

[71] No specific negation occurs with this construction.

[72] No specific negation occurs with this construction.

5.8.3.5.1.6. 2PL > 1 / 3

2PL > 1/3	1SG	1DU	1PL	3SG	3DU	3PL
2PL	I – MIX (NPST) request/confirmation/question $\sum$-TM-**ŋ-sʌ** ~ (*neg)[73] II – DIRECT (NPST) realization/comment/question $\sum$-**te**-TM-**n-i** ~ $\sum$(-ma)-**te-n-i**-li II – DIRECT (PST) realization/comment/question $\sum$-**te**-TM-**n-i** ~ $\sum$(-ma)-**te-n-i**-li			I – DIRECT $\sum$-**te**-TM-**n-i** ~ $\sum$(-ma)-**te-n-i**-li		

---

[73] No specific negation occurs with this construction.

5.8.3.5.1.7. 3SG > 1 / 2

	1SG	1DU	1PL	2SG	2DU	2PL
3SG						
	I – INVERSE	I – INVERSE	I – INVERSE	I – OBJECT	I – OBJECT	I – OBJECT
	∑-TM-**ta-ŋ**	∑-TM-**ta-ŋ-tɕʌ**	∑-TM-**ta-ŋ-su** (<tr.)	∑-**te**-TM	∑-**te**-TM-**dzʌ**	∑-**te**-TM-**i**
	~	∑-TM-**tʌhj-tɕʌ** (INCL.)	∑-TM-**ta-ŋ-i** (*arch.*)	~	~	~
	∑(-ma)-**ta-ŋ**-lʌ	~	∑-TM-**tʌh-i** (INCL.)	∑(-ma)-**te**-lʌ	∑(-ma)-**te-dzʌ**-lʌ	∑(-ma)-**te-i**-li
		∑(-ma)-**ta-ŋ-tɕʌ**-lʌ	~			∑(-ma)-**t-i**-li (POL)
		∑(-ma)-**tʌhj-tɕʌ**-lʌ [74]	∑(-ma)-**ta-ŋ-su**-lu			
			∑(-ma)-**ta-ŋ-i**-li (*arch.*)			
			∑(-ma)-**tʌh-i**-li [75]			
	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 2>3	REF. HIER. 2>3	REF. HIER. 2>3
				II – DIRECT (NPST)	II – DIRECT (NPST)	II – DIRECT (NPST)
				∑-TM-**u**	∑-TM-**u**	∑-TM-**u**
				~	~	~
				∑(-ma)-**u**-lu	∑(-ma)-**u**-lu	∑(-ma)-**u**-lu
				II – DIRECT (PST)	II – DIRECT (PST)	II – DIRECT (PST)
				∑-TM-**n**	∑-TM-**n**	∑-TM-**n**
				~	~	~
				∑(-ma)-**u**-lu	∑(-ma)-**u**-lu	∑(-ma)-**u**-lu

[74] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

[75] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

5.8.3.5.1.8. 3SG > 3

	3SG	3DU	3PL
3SG	I – INVERSE ∑-TM-**tʌ**~**tʰʌ**-**i** ~ ∑(-ma)-**ta**-lʌ ∑(-ma)-**tʌj**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT (NPST) ∑-TM-**u** ~ ∑(-ma)-**u**-lu II – DIRECT (PST) ∑-TM-**n** ~ ∑(-ma)-**u**-lu 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta**-**tɕʌ** ~ ∑(-ma)-**ta**-**tɕʌ**-lʌ ∑(-ma)-**ta**~**tʌj**-**tɕʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT (NPST) ∑-TM-**u** ~ ∑(-ma)-**u**-lu II – DIRECT (PST) ∑-TM-**n** ~ ∑(-ma)-**u**-lu 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta**-**sʌ** ~ ∑(-ma)-**ta**-**sʌ**-lʌ ∑(-ma)-**ta**~**tʌj**-**sʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT (NPST) ∑-TM-**u** ~ ∑(-ma)-**u**-lu II – DIRECT (PST) ∑-TM-**n** ~ ∑(-ma)-**u**-lu 3hum/nhum>3hum/nhum

5.8.3.5.1.9. 3ᴅᴜ > 1 / 2

	1sɢ	1ᴅᴜ	1ᴘʟ	2sɢ	2ᴅᴜ	2ᴘʟ
3ᴅᴜ	I – INVERSE ∑-TM-**ta-ŋ** ~ ∑(-ma)-**ta-ŋ**-lʌ	I – INVERSE ∑-TM-**ta-ŋ-tɕʌ** ∑-TM-**tʌhj-tɕʌ** (INCL.) ~ ∑(-ma)-**ta-ŋ-tɕʌ**-lʌ ∑(-ma)-**tʌhj-tɕʌ**-lʌ [76]	I – INVERSE ∑-TM-**ta-ŋ-su** ∑-TM-**ta-ŋ-i** (*arch.*) ∑-TM-**tʌh-i** (INCL.) ~ ∑(-ma)-**ta-ŋ-su**-lu ∑(-ma)-**ta-ŋ-i**-li (*arch.*) ∑(-ma)-**tʌh-i**-li [77]	I – OBJECT ∑-**te**-TM ~ ∑(-ma)-**te**-lʌ	I – OBJECT ∑-**te**-TM-**dzʌ** ~ ∑(-ma)-**te-dzʌ**-lʌ	I – OBJECT ∑-**te**-TM-**i** ~ ∑(-ma)-**te-i**-li ∑(-ma)-**t-i**-li (POL)
	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 2>3 II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu	REF. HIER. 2>3 II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu	REF. HIER. 2>3 II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu

---

[76] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

[77] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

5.8.3.5.1.10. 3DU > 3

	3SG	3DU	3PL
3DU	I – INVERSE ∑-TM-**tʌ~tʰʌ-i** ~ ∑(-ma)-**ta**-lʌ ∑(-ma)-**tʌ-i**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta-tɕʌ** ~ ∑(-ma)-**ta-tɕʌ**-lʌ ∑(-ma)-**ta~tʌ-i-tɕʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta-sʌ** ~ ∑(-ma)-**ta-sʌ**-lʌ ∑(-ma)-**ta~tʌ-i-sʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum <u>REF. HIER. 3HUM>3NHUM</u> II – DIRECT ∑-TM-**tɕ-u** ~ ∑(-ma)-**tɕ-u**-lu 3hum/nhum>3hum/nhum

5.8.3.5.1.11. 3PL > 1 / 2

	1SG	1DU	1PL	2SG	2DU	2PL
3PL	I – INVERSE ∑-TM-**ta-ŋ** ~ ∑(-ma)-**ta-ŋ**-lʌ	I – INVERSE ∑-TM-**ta-ŋ-tɕʌ** ∑-TM-**tʌhj-tɕʌ** (INCL.) ~ ∑(-ma)-**ta-ŋ-tɕʌ**-lʌ ∑(-ma)-**tʌhj-tɕʌ**-lʌ [78]	I – INVERSE ∑-TM-**ta-ŋ-su** ∑-TM-**ta-ŋ-i** (*arch.*) ∑-TM-**tʌh-i** (INCL.) ~ ∑(-ma)-**ta-ŋ-su**-lu ∑(-ma)-**ta-ŋ-i**-li (*arch.*) ∑(-ma)-**tʌh-i**-li [79]	I – OBJECT ∑-**te**-TM ~ ∑(-ma)-**te**-lʌ	I – OBJECT ∑-**te**-TM-**dzʌ** ~ ∑(-ma)-**te-dzʌ**-lʌ	I – OBJECT ∑-**te**-TM-**i** ~ ∑(-ma)-**te-i**-li ∑(-ma)-**t-i**-li (POL)
	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 1>3	REF. HIER. 2>3 II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n**-i-li	REF. HIER. 2>3 II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n**-i-li	REF. HIER. 2>3 II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n**-i-li

---

[78] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

[79] Inclusive negation only used in past tense; in non-past, jussive/imperative forms apply.

5.8.3.5.1.12. 3PL > 3

	3SG	3DU	3PL
3PL	I – INVERSE ∑-TM-**tʌ~tʰʌ-i** ~ ∑(-ma)-**ta**-lʌ ∑(-ma)-**tʌ-i**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum REF. HIER. 3HUM>3NHUM II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n-i**-li 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta-tɕʌ** ~ ∑(-ma)-**ta-tɕʌ**-lʌ ∑(-ma)-**ta~tʌ-i-tɕʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum REF. HIER. 3HUM>3NHUM II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n-i**-li 3hum/nhum>3hum/nhum	I – INVERSE ∑-TM-**ta-sʌ** ~ ∑(-ma)-**ta-sʌ**-lʌ ∑(-ma)-**ta~tʌ-i-sʌ**-lʌ (POL) 3hum>3hum; 3nhum>3hum *3hum>3nhum REF. HIER. 3HUM>3NHUM II – DIRECT ∑-TM-**n-i** ~ ∑(-ma)-**n-i**-li 3hum/nhum>3hum/nhum

5.8.3.5.2. (Di)transitive remote past paradigm

The argument indexation forms found in (di)transitive constructions in remote past are presented together in the following sub-sections. The remote past morpheme is =*to*. With negation, remote past is marked by the morpheme =*jak*.

In the following tables, the symbol ∑ indicates the position of the verbal stem and TM that of the tense markers. All indexation forms are bolded. The symbol {*x*} indicates that the morpheme may occur in either of these positions, and the symbol (*x*) that the morpheme remains optional. The negative forms occur below the tilde ~. When the forms are not attested because they do not correspond to any possible scenario, it is noted by n/a.

5.8.3.5.2.1. 1 > 2

	2SG	2DU	2PL
1SG	I − MIX ∑-**ne-ŋʌ**-to [80] ∑-**tɕe-ŋʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ**-lʌ	I − MIX ∑-**ne-ŋʌ-dzʌ**-to ∑-**tɕe-ŋʌ-dzʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-dzʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-dzʌ**-lʌ	I − MIX ∑-**ne-ŋʌ-sʌ**-to ∑-**tɕe-ŋʌ-sʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-sʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-sʌ**-lʌ
1DU.INCL (2)	n/a	n/a	n/a
1DU.EXCL (TO OTHER)	I − MIX ∑-**ne-ŋʌ-tɕ-u**-to ∑-**tɕe-ŋʌ-tɕ-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-tɕ-u**-lu	I − MIX ∑-**ne-ŋʌ-tɕ-u**-to ∑-**tɕe-ŋʌ-tɕ-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-tɕ-u**-lu II − MIX ∑-**ne-ŋʌ-dzʌ**-to ∑-**tɕe-ŋʌ-dzʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-dzʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-dzʌ**-lʌ	I − MIX ∑-**ne-ŋʌ-tɕ-u**-to ∑-**tɕe-ŋʌ-tɕ-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-tɕ-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-tɕ-u**-lu II − MIX ∑-**ne-ŋʌ-sʌ**-to ∑-**tɕe-ŋʌ-sʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-sʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-sʌ**-lʌ
1PL.INCL (3+GROUP)	n/a	n/a	n/a
1PL.EXCL (TO OTHER)	I − MIX ∑-**ne-ŋʌ-s-u**-to ∑-**tɕe-ŋʌ-s-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-s-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-s-u**-lu	I − MIX ∑-**ne-ŋʌ-s-u**-to ∑-**tɕe-ŋʌ-s-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-s-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-s-u**-lu II − MIX ∑-**ne-ŋʌ-dzʌ**-to ∑-**tɕe-ŋʌ-dzʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-dzʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-dzʌ**-lʌ	I − MIX ∑-**ne-ŋʌ-s-u**-to ∑-**tɕe-ŋʌ-s-u**-to ~ ∑(-ma)-jak-**ne-ŋʌ-s-u**-lu ∑(-ma)-jak-**tɕe-ŋʌ-s-u**-lu II − MIX ∑-**ne-ŋʌ-sʌ**-to ∑-**tɕe-ŋʌ-sʌ**-to ~ ∑(-ma)-jak-**ne-ŋʌ-sʌ**-lʌ ∑(-ma)-jak-**tɕe-ŋʌ-sʌ**-lʌ

---

[80] The 1>2 morpheme =*ne* does not occur in RAP-13: Polkim, Sarling, Syamrang, Yuiling, Santhali

5.8.3.5.2.2. 1 > 3

	3SG	3DU	3PL
1SG	I – DIRECT ∑-**ŋʌ**-to ~ ∑(-ma)-jak-**ŋʌ**-lʌ		
1DU.INCL (2)	I – DIRECT ∑-**tʌhj-tɕ-u**-to ~ ∑(-ma)-jak-**tʌhj-tɕ-u**-lu		
1DU.EXCL (TO OTHER)	I – DIRECT ∑-**ŋʌ-tɕ-u**-to ~ ∑(-ma)-jak-**ŋʌ-tɕ-u**-lu		
1PL.INCL (3+GROUP)	I – DIRECT ∑-**tʌhj-n-i**-to ~ ∑(-ma)-jak-**tʌhj-n-i**-li		
1PL.EXCL (TO OTHER)	I – DIRECT ∑-**ŋʌ-s-u**-to ~ ∑(-ma)-jak-**ŋʌ-s-u**-lu		

5.8.3.5.2.3. 2 > 1

	1SG	1DU	1PL
2SG	I – MIX ∑-**te**-**tɕi**-to ~ ∑(-ma)-jak-**te**-**tɕi**-li	I – DIRECT/MIX ∑-**te**-**u**-to ~ ∑(-ma)-jak-**te**-**u**-lu	I – DIRECT/MIX ∑-**te**-**u**-to ~ ∑(-ma)-jak-**te**-**u**-lu
2DU	I – DIRECT/MIX ∑-**te**-**dʑ**-**u**-to ~ ∑(-ma)-jak-**te**-**dʑ**-**u**-lu		
2PL	I – DIRECT/MIX ∑-**te**-**n**-**i**-to ~ ∑(-ma)-jak-**te**-**n**-**i**-li		

5.8.3.5.2.4. 2 > 3

	3SG	3DU	3PL
2SG	I – DIRECT/MIX ∑-**te**-**u**-to ~ ∑(-ma)-jak-**te**-**u**-lu		
2DU	I – DIRECT/MIX ∑-**te**-**dz**-**u**-to ~ ∑(-ma)-jak-**te**-**dz**-**u**-lu		
2PL	I – DIRECT/MIX ∑-**te**-**n**-**i**-to ~ ∑(-ma)-jak-**te**-**n**-**i**-li		

5.8.3.5.2.5. 3 > 1

	1sɢ	1ᴅᴜ	1ᴘʟ
3sɢ	I – INVERSE ∑-**ta-ŋ**-to ~ ∑(-ma)-jak-**ta-ŋ**-lʌ	I – INVERSE ∑-**ta-ŋ-tɕʌ**-to ~ ∑(-ma)-jak-**ta-ŋ-tɕʌ**-lʌ -- ∑-**tʌhj-tɕʌ**-to (INCL.) ~ ∑(-ma)-jak-**tʌhj-tɕʌ**-lʌ (INCL.)	I – INVERSE ∑-**ta-ŋ-su**-to (<tr.) ∑-**ta-ŋ-i**-to (*arch.*) ~ ∑(-ma)-jak-**ta-ŋ-su**-lu ∑(-ma)-jak-**ta-ŋ-i**-li (*arch.*) -- ∑-**tʌh-i**-to (INCL.) ~ ∑(-ma)-jak-**tʌh-i**-li (INCL.)
3ᴅᴜ	I – INVERSE ∑-**ta-ŋ**-to ~ ∑(-ma)-jak-**ta-ŋ**-lʌ	I – INVERSE ∑-**ta-ŋ-tɕʌ**-to ~ ∑(-ma)-jak-**ta-ŋ-tɕʌ**-lʌ -- ∑-**tʌhj-tɕʌ**-to (INCL.) ~ ∑(-ma)-jak-**tʌhj-tɕʌ**-lʌ (INCL.)	I – INVERSE ∑-**ta-ŋ-su**-to (<tr.) ∑-**ta-ŋ-i**-to (*arch.*) ~ ∑(-ma)-jak-**ta-ŋ-su**-lu ∑(-ma)-jak-**ta-ŋ-i**-li (*arch.*) -- ∑-**tʌh-i**-to (INCL.) ~ ∑(-ma)-jak-**tʌh-i**-li (INCL.)
3ᴘʟ	I – INVERSE ∑-**ta-ŋ**-to ~ ∑(-ma)-jak-**ta-ŋ**-lʌ	I – INVERSE ∑-**ta-ŋ-tɕʌ**-to ~ ∑(-ma)-jak-**ta-ŋ-tɕʌ**-lʌ -- ∑-**tʌhj-tɕʌ**-to (INCL.) ~ ∑(-ma)-jak-**tʌhj-tɕʌ**-lʌ (INCL.)	I – INVERSE ∑-**ta-ŋ-su**-to (<tr.) ∑-**ta-ŋ-i**-to (*arch.*) ~ ∑(-ma)-jak-**ta-ŋ-su**-lu ∑(-ma)-jak-**ta-ŋ-i**-li (*arch.*) -- ∑-**tʌh-i**-to (INCL.) ~ ∑(-ma)-jak-**tʌh-i**-li (INCL.)

5.8.3.5.2.6. 3 > 2

	2SG	2DU	2PL
3SG	I − OBJECT ∑-**te**-to ~ ∑(-ma)-jak-**te**-lʌ	I − OBJECT ∑-**te**-**dzʌ**-to ~ ∑(-ma)-jak-**te**-**dzʌ**-lʌ	I − OBJECT ∑-**te**-**i**-to ∑-**t**-**i**-to (POL) ~ ∑(-ma)-jak-**te**-**i**-li ∑(-ma)-jak-**t**-**i**-li (POL) I − DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu
	I − DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu	I − DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu	
3DU	I − OBJECT ∑-**te**-to ~ ∑(-ma)-jak-**te**-lʌ	I − OBJECT ∑-**te**-**dzʌ**-to ~ ∑(-ma)-jak-**te**-**dzʌ**-lʌ	I − OBJECT ∑-**te**-**i**-to ∑-**t**-**i**-to (POL) ~ ∑(-ma)-jak-**te**-**i**-li ∑(-ma)-jak-**t**-**i**-li (POL) I − DIRECT ∑-**tɕ**-**u**-to ~ ∑(-ma)-jak-**tɕ**-**u**-lu
	I − DIRECT ∑-**tɕ**-**u**-to ~ ∑(-ma)-jak-**tɕ**-**u**-lu	I − DIRECT ∑-**tɕ**-**u**-to ~ ∑(-ma)-jak-**tɕ**-**u**-lu	
3PL	I − OBJECT ∑-**te**-to ~ ∑(-ma)-jak-**te**-lʌ	I − OBJECT ∑-**te**-**dzʌ**-to ~ ∑(-ma)-jak-**te**-**dzʌ**-lʌ	I − OBJECT ∑-**te**-**i**-to ∑-**t**-**i**-to (POL) ~ ∑(-ma)-jak-**te**-**i**-li ∑(-ma)-jak-**t**-**i**-li (POL) I − DIRECT ∑-**n**-**i**-to ~ ∑(-ma)-jak-**n**-**i**-li
	I − DIRECT ∑-**n**-**i**-to ~ ∑ (-ma)-jak-**n**-**i**-li	I − DIRECT ∑-**n**-**i**-to ~ ∑(-ma)-jak-**n**-**i**-li	

5.8.3.5.2.7. 3 > 3

	3SG	3DU	3PL
3SG	I – INVERSE ∑-**tʰʌ-i**-to ∑-**tʌ-i**-to (POL) ~ ∑(-ma)-jak-**ta**-lʌ ∑(-ma)-jak-**tʌj**-lʌ (POL) I – DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu	I – INVERSE ∑-**ta-tɕʌ**-to ~ ∑(-ma)-jak-**ta-tɕʌ**-lʌ I – DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu	I – INVERSE ∑-**ta-sʌ**-to ~ ∑(-ma)-jak-**ta-sʌ**-lʌ I – DIRECT ∑-**u**-to ~ ∑(-ma)-jak-**u**-lu
3DU	I – INVERSE ∑-**tʰʌ-i**-to ∑-**tʌ-i**-to (POL) ~ ∑(-ma)-jak-**ta**-lʌ ∑(-ma)-jak-**tʌj**-lʌ (POL) I – DIRECT ∑-**tɕ-u**-to ~ ∑(-ma)-jak-**tɕ-u**-lu	I – INVERSE ∑-**ta-tɕʌ**-to ~ ∑(-ma)-jak-**ta-tɕʌ**-lʌ I – DIRECT ∑-**tɕ-u**-to ~ ∑(-ma)-jak-**tɕ-u**-lu	I – INVERSE ∑-**ta-sʌ**-to ~ ∑(-ma)-jak-**ta-sʌ**-lʌ I – DIRECT ∑-**tɕ-u**-to ~ ∑(-ma)-jak-**tɕ-u**-lu
3PL	I – INVERSE ∑-**tʰʌ-i**-to ∑-**tʌ-i**-to (POL) ~ ∑(-ma)-jak-**ta**-lʌ ∑(-ma)-jak-**tʌj**-lʌ (POL) I – DIRECT ∑-**n-i**-to ~ ∑ (-ma)-jak-**n-i**-li	I – INVERSE ∑-**ta-tɕʌ**-to ~ ∑(-ma)-jak-**ta-tɕʌ**-lʌ I – DIRECT ∑-**n-i**-to ~ ∑(-ma)-jak-**n-i**-li	I – INVERSE ∑-**ta-sʌ**-to ~ ∑(-ma)-jak-**ta-sʌ**-lʌ I – DIRECT ∑-**n-i**-to ~ ∑(-ma)-jak-**n-i**-li

5.8.3.5.3. (Di)transitive imperative/optative paradigm

The argument indexation forms found in (di)transitive constructions in imperative and optative are presented together in the following sub-sections. The optative morpheme is =pa and the imperative is expressed through the encliticization of argument indexation markers. Imperative and optative negation is marked with the morpheme =ljam (LOTHAR) ~ =lam (MANAHARI) for all persons.

In the following tables, the symbol ∑ indicates the position of the verbal stem and TM that of the tense markers. All indexation forms are bolded. The symbol {x} indicates that the morpheme may occur in either of these positions, and the symbol (x) that the morpheme remains optional. The negative forms occur below the tilde ~. When the forms are not attested because they do not correspond to any possible scenario, it is noted by n/a.

5.8.3.5.3.1. 1 > 2 / 3

	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
1SG		REFL			∑-tɕja-ŋ (Q) ∑-ne-pa-ŋ (OPT) ∑-tɕe-pa-ŋ (OPT)			∑-tɕja-ŋ (Q) ∑-pa-ŋ (OPT)	
1DU		RECIP REFL			∑-tɕja-ŋ-tɕ-u (Q) ∑-ne-pa-ŋ-tɕ-u (OPT) ∑-tɕe-pa-ŋ-tɕ-u (OPT)			∑-tɕ-u ∑-pa-ŋ-tɕ-u (OPT)	
1PL		RECIP REFL			∑-tɕja-ŋ-s-u (Q) ∑-ne-pa-ŋ-s-u (OPT) ∑-tɕe-pa-ŋ-s-u (OPT)			∑-n-i ∑-pa-ŋ-n-i (OPT)	

5.8.3.5.3.2. 2 > 1 / 3

	1SG	1DU	1PL	2SG	2DU	2PL	3SG	3DU	3PL
2SG	∑-tɕi -- ∑-te-pa-tɕi (OPT)			REFL	n/a	n/a	∑-u -- ∑-te-pa-u (OPT)		
2DU	∑-dʑ-u			n/a	RECIP REFL	n/a	∑-dʑ-u -- ∑-te-pa-dʑ-u (OPT)		
2PL	∑-n-u			n/a	n/a	RECIP REFL	∑-n-u -- ∑-te-pa-n-i (OPT)		
2HON	∑-sa sjaw-pa			n/a	n/a	n/a	∑-sa sjaw-pa		

5.8.3.5.3.3. 3 > 1 / 2

	1SG	1DU	1PL	2SG	2DU	2PL
3SG 3DU 3PL	∑-pa-ta-ŋ	∑-pa-ta-ŋ-tɕʌ -- ∑-pa-tʌhj- tɕʌ (INCL.)	∑-pa-ta-ŋ-sʌ -- ∑-pa-tʌh-i (INCL.)	∑-te-pa	∑-te-pa-dʑʌ	∑-te-pa-i

5.8.3.5.3.4. 3 > 3

	3SG	3DU	3PL
3SG	∑-pa-tʰʌ-i ∑-pa-tʌ-i (POL) -- ∑-pa-**u**	∑-pa-**ta-tɕʌ** -- ∑-pa-**u**	∑-pa-**ta-sʌ** -- ∑-pa-**u**
3DU	∑-pa-tʰʌ-i ∑-pa-tʌ-i (POL) -- ∑-pa-**tɕ-u**	∑-pa-**ta-tɕʌ** -- ∑-pa-**tɕ-u**	∑-pa-**ta-sʌ** -- ∑-pa-**tɕ-u**
3PL	∑-pa-tʰʌ-i ∑-pa-tʌ-i (POL) -- ∑-pa-**n-i**	∑-pa-**ta-tɕʌ** -- ∑-pa-**n-i**	∑-pa-**ta-sʌ** -- ∑-pa-**n-i**

5.8.3.6.   Possessor indexation in =*bʌt* construction

The =*bʌt* construction indexes the possessor on the verb with animate possessed referents expressed as S and P/R arguments, and inanimate possessed referents functioning as P arguments. However, we will see that the morpheme =*bʌt*, which correlates with the presence of an inverse form (or object construction type in the case of 2nd person) when the verb is transitive or ditransitive, does not only trigger the indexation of the possessor.

This construction has been characterized as Prominent Internal Possessor (PIP) (Bárány, Bond & Nikolaeva 2019). Note that the use of this construction is considered archaic in all studied varieties, and that no natural example was found in our corpus.

In Chepang, the PIP construction conveys a benefactive/malefactive meaning, expressing that the speaker is somewhat emotionally affected by the situation expressed, as in (753), or that the situation directly benefits or affects the speaker, as in (754) and (755). In all cases, the speaker is the possessor and seeks the interlocutor's compassion or understanding. When the speaker is not the possessor but 2nd person, the construction conveys the speaker's compassion, understanding or excitement to the interlocutor; in this case, the 2nd person possessor is indexed. This is illustrated in (756) to (757). Note that the ergative may or may not occur when the verb is intransitive.

The possessed referent may be animate and function as S or P/R argument; when inanimate, for the PIP construction to be used, it has to hold the function of SP argument. In this case, the A is not expressed, forming an anticausative construction, as in (758).

(753) ŋa=ko   tɕoʔ(=i)   rja=bʌt=na=ta=ŋ.
       1SG=GEN   child(=ERG)   cry=PIP=NPST=INV=1

       'My child cries.'
       CH_MKW _PC_SIL_082020_E_1

(754) ŋa=ko   tɕoʔ=i   tuʔm=bʌt=ka=ta=ŋ.
       1SG=GEN   child=ERG   kiss=PIP=2/3.PST=INV=1

       'My child kissed me.'
       CH_MKW _PC_SIL_082020_E_1

(755) ŋa=ko   tɕoʔ=i   ŋiʔ=bʌt=ka=ta=ŋ.
       1SG=GEN   child=ERG   laugh=PIP=2/3.PST=INV=1

       'My child laughed at me.'
       CH_MKW _PC_SIL_082020_E_1

(756) ŋaŋ=ko   tɕoʔ(=i)   rja=bʌt=te=a.
       2SG=GEN   child(=ERG)   cry=PIP=2=PST

       'Your child cried.'
       CH_MKW _PC_SIL_082020_E_1

(757) ŋaŋ=ko   tɕoʔ(=i)   waŋ=bʌt=te=a.
       2SG=GEN   child(=ERG)   come=PIP=2=PST

       'Your child arrived!'

       CH_MKW_PC_SIL_082020_E_1

(758) ŋa=ko   jam   waj=bʌt=ka=ta=ŋ
       1SG=GEN   rice   throw=PIP=2/3.PST=INV=1

       'My rice got thrown away.'

       CH_MKW_PC_SIL_082020_E_1

The PIP construction is also attested in another configuration: the possessor is 2nd person, the possessed is the A argument, and another participant is involved, a P argument; the nature of the process can trigger the presence of the morpheme =bʌt, expressing the concern of the speaker; in this case, the possessor is not indexed on the verb while the P argument of the verb is, as in (759).

(759) ŋaŋ=ko   kwi=i   o=kaj   dzʌjk=bʌt=ka=tʰʌ=i !
       2SG=GEN   dog=ERG   DIST=DAT   bite=PIP=2/3.PST=INV=3>3SG

       'Your dog bit her/him!'

       CH_MKW_PC_SIL_082020_E_1

Finally, when the speaker is the benefactive of a process, regardless of the presence of a possessive construction, the inverse morpheme may occur with intransitive verbs. This is illustrated in (760).

(760) naŋ   waŋ=te=tɕja=ta=ŋ !
       2SG   come=2=IRR=INV=1

       'You would come for me!'

       CH_MKW_PC_SIL_082020_E_1

PIP constructions in Chepang, in addition to the presence of inverse morphology with intransitive verbs, needs further investigation; this would shed light on other patterns

observed with inverse morphology in (di)transitive constructions, and on how the Chepang's non-canonical direct-inverse came to be.

5.8.4. Basic tense, aspect, modality

This section presents the morphology that conveys basic tense, aspect and modality through the presence of a morpheme following the verb stem: non-past (§ 5.8.4.1), past (§ 5.8.4.2), remote past (§ 5.8.4.3), and irrealis (§ 5.8.4.4).

5.8.4.1. Non-past: non-past habitual and future

The morpheme =*na* marks non-past tense; it can express habitual aspect (761) and (762), as well as future tense (763).

(761) *ŋi*   *juin=ma*   *dze=na=ŋ=su.*
1PL   bat=ADD   eat=NPST=1=1PL.EXCL

'We eat also bats.'
CH_MKW_PNC_SIL_081818_1_Life

(762) *naŋ*   *Nepali*   *miks* *dzahŋ=ti*   *noʔ=te=na*   *ja ?*
2SG   Nepali   mix do_make=SEQ1   speak=2=NPST   or

'Do you speak mixing with Nepali?'
CH_MKW_MRC_DAM_112819_Conversation_with_Bipana

(763) *lʌ*   *ama,*   *sjahŋ=ma*   *ŋa=paj*
well   mother   tomorrow=ADD   1SG=DIS

kjan   *boŋ=laŋ*   *al=naŋ.*
dish   look.for=PUR   go=NPST=1SG

'Well mother, I'll go look for food tomorrow too.'
CH_MKW_DBC_MAI_2_020320_Newa_Dung

5.8.4.2.   Past

Past tense expresses past events which happen before the time of speech. The past tense marker associated with 2nd and 3rd person singular is =a with intransitive verbs. The past morphemes =ak ~ aka ~ ka correlate with other specific functions, such as (di)transitivity, directionality, and the marking of number. In addition, only =aka ~ ka is found with reflexive and reciprocal morphology (§ 5.8.3.2.1). Finally, the past morpheme attested with 1st person became portemanteau and two allomorphs are attested: the morpheme =alaŋ, and =akaŋ ~ kaŋ. The former is used in all varieties and the latter in RAP-13: Polkim, Sarling, Syamrang, Yuiling, Santhali.

The distribution of past forms is summarized in Table 159.

Table 159. Distribution of past markers

past forms	intr.2/3sg	intr.2/3pl	tr.2/3	inverse	refl/recip	1st person
=a	=te=a					
	=a					
=ak~a~aka~ka		=te=aka~ka=i	=te=aka~ka=n	=ak~a~ka=ta=ŋ	=aka~ka=sʌ	
		=aka~ka=i	=aka~ka=n	=ak~a~ka=tʌ~tʰʌ=i	=aka~ka=i=si	
=alaŋ~akaŋ~kaŋ						=alaŋ
						=akaŋ
						=kaŋ

The past morphemes =a and =ak ~ aka ~ ka associated with 2nd and 3rd person are illustrated in (764) to (769) in intransitive and transitive constructions, and the morpheme =alaŋ ~ akaŋ ~ kaŋ associated with 1st person is illustrated in (770) and (771).

(764)   *kliʔ=taŋ*   *boŋ=laŋ*   *al=a.*
        shit=ATT   look.for=PUR   go=PST

'(The mouse) went to look for some shit.'
CH_CTW_JBC_BHR_111720_2_Mouse

(765) *Palung=haŋ   dah=aka=i.*
Palung=LOC1   reach=2/3.PST=PL

'They arrived in Palung.'

CH_MKW_PSC_MAI_012720_Local_History_1

(766) *ten=paj   ŋa=kaj   dʌjti   guhl=te=aka=n ?*
today=DIS   1SG=DAT   why   follow=2=2/3.PST=DIR/TR

'Why did you follow me today?'

CH_MKW_PBPC_CHI_110219_4_Sita and the origin of khar

(767) *Chepang bʰasa   mʰeʔ=aka=n=i.*
Chepang language   forget=2/3.PST=DIR/TR=PL

'They forgot the Chepang language.'

CH_MKW_DKC_MAI_012720_Chepang_Language

(768) *tas   kʰel=ʌ=o   bʌlʌ   pʰe=ka=nʌ,*
cards   play=LN=NMZ:REL   little   leave=2/3.PST=DIR/TR

han-rʌksi   pʰe=ka=nʌ,
beer-alcohol   leave=2/3.PST=DIR/TR

'He stopped playing cards a little, he quit drinking,'

CH_MKW_SC_SIL_010220_1_Life

(769) *naŋ=i   ŋa=kaj   aʔl=ak=ta=ŋ   kja !*
2SG=ERG   1SG=DAT   take.away=PST=INV=1   PART

'You took me away / married me, that's what I mean!'

CH_MKW_BMB_BAN_090118_7_Chepang_marriage

(770) *biɾami*　　　*sjaw=ti*　　　　　*i=haŋ*　　　　*pande=kaj*
　　　　sick　　　　　become=SEQ1　　　PROX=LOC1　　shaman=DAT

　　　　ʥʰja=tak=alaŋ.

　　　　practice.shamanism=CAUS=1.PST

　　　　'Having become sick, I made the shaman practice a ceremony here.'

　　　　CH_MKW_PSC_MAI_012620_1_Becoming_Christian

(771) *kʰaŋ=akaŋ !*　　　　*bʰena,*　　　　　　　　*lʌw,*
　　　　cook=1.PST　　　　　elder.sister's.husband　　well

　　　　ʥe=laŋ　　　*pok=ʌ !*

　　　　eat=PUR　　　　enter=2SG.IMP.INTR

　　　　'I cooked! Brother-in-law, well, come in to eat!'

　　　　CH_CTW_JBC_BHR_102420_1_Cing_Lan

5.8.4.3.　　Remote past, habitual past

The morpheme *=to* is marks remote past or habitual past. This is illustrated in (772) and (773). With negation, it is marked by the morpheme *=jak*, as in (774).

(772) *ŋa*　　*o=kʰa*　　　*ʌ,*　　*Kalitar*　　*mu=dʰana*
　　　　1SG　　DIST=LOC2　　uh　　Kalitar　　　COP=SIM

　　　　ŋa=i　　　　*oʔn*　　　　　　　　*ɾa=ŋʌ=to*　　　　*dʌj !*

　　　　1SG=ERG　　　imperata.arundinacia　　cut=1=REM.PST　　PART

　　　　'Me, there, uh, when living in Kalitar, I used to cut imperata arundinacia grass, hey!'

　　　　CH_MKW_PMRC_LAM_081618_3_Kalitar

(773) *ʌni*   *o*   *laʔ-luiʔ,*   *laʔ=i*   *ap=n=i=to,*
     so   DIST   arrow-bow   arrow=INST   shoot=DIR/TR=PL=REM.PST

     ʌni   *ruiŋ-sja=ma*   *si=to,*   *bʌdel=ma*   *si=to,*
     so   hog_sambar-prey=ADD   die=REM.PST   wild.pig=ADD die=REM.PST

     'So, they used to shoot with these arrows and bows, and hogs, sambar deer would die, wild pigs would die,'

     CH_MKW_SCBKC_SIL_081918_5_Hunting

(774) *lʌ,*   *o*   *tɕʌlʌn=tɕʌhe,*   *ŋa=i*   *tɕiʔ=jak=ŋʌ=lʌ,*
     well   DIST   tradition=DIS   1SG=ERG   know=REM.PST=1=NEG

     pʰeɾe,   *tɕiʔ=ŋʌ=lʌ.*
     DIS   know=1=NEG

     'Well, as for those traditions, I never knew them, I don't know them.'

     CH_CTW_SMBC_BBC_GUN_012120_Chepang_Kings

5.8.4.4. Irrealis

Irrealis is expressed using the morpheme =*tɕja* (LOTHAR) and allomorph =*tɕa* (MANAHARI). It is illustrated in (775) and (776).

(775) *i=kʰa*   *waŋ=ti*   *git*   *ke=tɕja=ŋ=tʌ*   *mʰʌɾ=to.*
     PROX=LOC2   come=SEQ1   song   sing=IRR=1=REP   think=REM.PST

     'Arriving here, I had thought that I would sing.'

     CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation

(776) *kwej*   *sabun* *leʔ=o*   *kʰe=tɕja,*
     some   soap   take_buy=NMZ:REL   COP=IRR

     kwej   *sampu*   *leʔ=o*   *kʰe=tɕja,*
     some   shampoo   take_buy=NMZ:REL   COP=IRR

     'Some would buy soap, some would buy shampoo,'

     CH_MKW_STC_SIL_120619_2_E_4

APPENDIX I

METADATA OF RECORDED SPEAKERS

Speakers who participated in the recordings that constitute the Chepang Corpus (transcribed and translated) by June 2022

Table 160. Chepang speakers who participated in the recordings from Lothar

LOTHAR - RAK-6, RAK-8

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Moti Ram Nak Darey Chepang	MRNDC	m	55	2018	Silinge, rak-6	SIL	Nak Darey		Silinge, rak-6
Sita Ram Nak Darey Chepang	SRNDC	m	80	2018	Silinge, rak-6	SIL	Nak Darey		Silinge, rak-6
Sani Cara Bas Kota Chepang Praja	SCBKC	m	~55	2018	Silinge, rak-6	SIL	Bas Kota		Silinge, rak-6
Prashant Ngarba Chepang	PNC	m	34	2018	Silinge, rak-6	SIL	Pakhrin		Silinge, rak-6
Prem Maya Rayda Chepang	PMRC	f	~50	2018	Silinge, rak-6	SIL	Rayda		?
Bipana Chepang	BC	f	22	2019	Silinge, rak-6	SIL	Nak Darey		Silinge, rak-6
Lakshmi Chepang	LC	f	43	2019	Silinge, rak-6	SIL	Pakhin	Nak Darey	Silinge, rak-6
Rita Chepang	RC	f	18	2019	Silinge, rak-6	SIL			Bangrang, rak-6
Jhagari Maya Chepang	JMC	f	74	2019	Silinge, rak-6	SIL			Dambarang, rak-7
Bishu Lal Chepang	BLC	m	?	2019	Silinge, rak-6	SIL			Silinge, rak-6
Sita Chepang	STC	f	39	2019	Silinge, rak-6	SIL			Silinge, rak-6
Susmita Chepang	SC	f	45	2019	Silinge, rak-6	SIL			Silinge, rak-6
Chandri Maya Chepang	CMC	f	30	2019	Silinge, rak-6	SIL			Dhirang, rak-7

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Bakhat Bahadur Chepang	BBC	m	41	2020	Silinge, rak-6	SIL			Silinge, rak-6
Subitra Piti Maya Chepang	SPMC	f	50	2021	Silinge, rak-6	SIL	Pakhin		Silinge, rak-6
Pabitra Chepang	PC	f	34	2021	Silinge, rak-6	SIL	Pakhin	Nak Darey	Silinge, rak-6
Gam Bahadur Chepang	GBC	m	26	2019	Cyorang, rak-6	CYO			Cyorang, rak-6
Devi Lal Chepang	DLC	m	52	2019	Cyorang, rak-6	CYO			?
Buddi Raj Chepang	BRC	m	47	2019	Cyorang, rak-6	CYO			Cyorang, rak-6
Masta Ram Chepang	MRC	m	24	2019	Dambarang, rak-7	DAM			?
Nirmaya Chepang	NC	f	70	2019	Dambarang, rak-7	DAM			?
Shanti Chepang	SC	f	22	2019	Dambarang, rak-7	DAM			?
Saran Kumar Praja	SKP	m	29	2019	Dambarang, rak-7	DAM			?
Tika Ram Chepang	TRC	m	20	2019	Dambarang, rak-7	DAM			?
Sarkini Maya Chepang	SMC	f	65	2019	Aysirang, rak-7	AYS			?

LOTHAR - RAP-13, RAP-11

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Shanta Bahadur Chepang	SBC	m	36	2020	Gundi, rap-13	GUN			Gundi, rap-13
Bhim Bahadur Chepang	BMBC	m	76	2020	Gundi, rap-13	GUN			Gundi, rap-13
Kamala Chepang	KC	f	24	2020	Gundi, rap-13	GUN			Gundi, rap-13
Shanta Ram Praja	SRP	m	42	2020	Gundi, rap-13	GUN			Gundi, rap-13
Bhim Bahadur Chepang	BBC	m	52	2020	Gundi, rap-13	GUN			Brasbang (Dhading)
Yani Maya Chepang	YMC	f	20	2020	Tapang, rap-13	TAP	Gawru		Tapang, rap-13
Kalisman Chepang	KMC	m	59	2020	Tapang, rap-13	TAP	Gawru Khalak		Tapang, rap-13

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Bhim Bahadur Chepang	BBC	m	45	2020	Polkim, rap-13	POL	Thakuri		Polkim, rap-13
Samjana Chepang	SC	f	33	2020	Polkim, rap-13	POL			Kharkanetar rak-6
Santosh Praja	SP	m	19	2020	Polkim, rap-13	POL	Thakuri		Polkim, rap-13
Jit Bahadur Chepang	JBC	m	79	2020	Bhantarang, rap-13	BHR	Rai		Syamrang, rap-13
Kopi Ram Chepang	KRC	m	77	2020	Hatti Sudde, rap-13	HAT			Hatti Sudde, rap-13
Rabi Lal Chepang	RLC	m	58	2020	Jimling, rap-11	JIM	Shahi		Moling, rap-12 (Korak GBS)
Sani Maya Chepang	SMC	f	45	2020	Jimling, rap-11	JIM	Kalinge		Cauki dada, Kaule (Ichyakamana NP)
Ambar Bahadur Chepang	ABC	m	52	2020	Jimling, rap-11	JIM			Jimling, rap-11
Jal Maya Chepang	JMC	f	62	2020	Payak, rap-11	PYK	Bas Puri		Syamrang, rap-13
Rasila Chepang	RC	f	36	2020	Kuccur, rap-11	KCR	Nak Darey		Silinge, rak-6

Table 161. Chepang speakers who participated in the recordings from Manahari

MANAHARI - RAK-8

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Pratap Singh Chepang	PSC	m	70	2020	Maisirang, rak-8	MAI	Sawa		Maisirang, rak-8
Dev Kumar Chepang	DKC	m	45	2020	Maisirang, rak-8	MAI	Sawa		Maisirang, rak-8
Chandra Bahadur Chepang	CBC	m	58	2020	Maisirang, rak-8	MAI			Maisirang, rak-8
Goma Praja	GP	f	35	2020	Maisirang, rak-8	MAI			?
Sarina Chepang	SC	f	28	2020	Maisirang, rak-8	MAI	Sawa		Maisirang, rak-8
Anjali Chepang	AC	f	28	2020	Maisirang, rak-8	MAI	Galsarang		Maisirang, rak-8
Deu Bahadur Chepang	DBC	m	55	2020	Maisirang, rak-8	MAI	Ruing Thorya		?

Table 162. Chepang speakers who participated in the recordings from Manahari

RAPTI - RAP-3, RAP-6

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Bir Bahadur Chepang	BBC	m	69	2020	Pyari Dap	PID			Padampur GBS-1, Kalika NP
Singh Bahadur Praja Chepang	SBPC	m	73	2020	Dhameli / Simaltar	DHM	Bare		Jiru Bas, Talti-1, Dhading

Table 163. Chepang speakers who participated in the recordings from Handikhola

HANDIKHOLA - MAN-4

name	code	biol. gender	age	year	village	code	clan (f)	clan (m)	born
Bishnu Maya Bhaisikhore	BMB	f	54	2017-	Bankarya Tol	BAN	Bhaisikhore		Lamitar
Chamili Panyo Rana Praja	CPR	f	45	2017-	Bankarya Tol	BAN	Kalikote	Potbangmay	Tongra
Sahili Maya Yorangmay Praja	SMYMP	f	~55	2018	Chisopani	CHI	Yorangmay		Khairang
Kishna Kumari Byal Praja	KKBP	f	~75	2017-	Lampakha	LPK	Broso		Athare, Khairang Raksirang
Maya Rana Chepang	MRC	f	~50	2018-	Lampakha	LPK			Gyung, Raksirang
Kishna Bahadur Kurangi Chepang	KBKC	m	~55	2017-	Tongra	TNG			?
Kishna Maya Chepang	KMC	f	56	2018-	Siddha Kali (Kami dada), Handikhola, 3	SK			Dumre Dokothar, Bhakta
Phul Maya Rupokote Chepang	PMRC	f	~50	2018-	Lamitar	LAM	Rupokote		?

Shanta Lal Praja	SLP	m	37	2019-	Ajingare - Damar Naya Basti	AJI		Yado Khola, Raksirang GBS
Ramila Praja	RP	f	25	2019-	Chisopani - Damar Naya Basti	CHI		Dillipur
Samir Praja	SP	m	28	2019-	Chisopani - Damar Naya Basti	CHI		Dillipur
Ayta Ram Rupokote Chepang	ARRKC	m	~45	2019-	Chisopani - Damar Naya Basti	CHI	Rupokote	Gyung, Raksirang GBS
Kishna Bahadur Chepang	KBC	m	~55	2019-	Siddha Kali (Kami dada), Handikhola, 3	SK		?
Phul Bahadur Praja Chepang	PBPC	m	65	2019-	Chisopani	CHI	Chyura Kharka	Durgamchetra, Raksirang GBS

APPENDIX II

METADATA OF CHEPANG RECORDINGS

Recordings that constitute the Chepang Corpus (transcribed and translated) by June 2022

Table 164. Total duration of recordings and number of recordings per type

2017-21		duration	number of recordings
	Elicitation	3:53:45	27
	Texts (Narrative, Expository)	15:56:40	178
	Conversations	3:58:32	26
	Songs	1:16:08	30
Total		**25:05:05**	**261**

Figure 76. Abbreviations for Metadata of Chepang recordings

    (a) Abbreviations used in Table 168 for the type of recording:

        E – Elicitation

        N – Narrative

        Exp – Expository

        C – Conversation

        S – Song

        a – audio recording

        v – video recording

    (b) Recording file name is expressed as follows:

        language_district_(notebook.page)_speaker_village_date_number_title

    (c) Recording file name abbreviations are as follows:

        CH – Chepang

        MKW – Makawanpur district

        CTW – Chitwan district

    (d) Speaker name abbreviations are given in APPENDIX I.

Table 165. Metadata of Chepang recordings

type	recording	speakers	date	type	time
C	CH_MKW_RC_JMC_SIL_120119_Conversation	JMC_RC	12/01/19	a	0:33:21
C	CH_CTW_SBC_BMBC_GUN_012120_Chepang Kings	SCB_BMBC	01/21/20	v	0:05:16
C	CH_CTW_SRP_GUN_102620_1_Gundi_name	SRP	10/26/20	v	0:01:56
C	CH_CTW_SRP_GUN_102620_2_Chepang_kings	SRP	10/26/20	v	0:01:38
C	CH_CTW_SRP_GUN_102620_3_Gundi_settlement	SRP	10/26/20	v	0:02:35
C	CH_CTW_SRP_GUN_102620_4_Local_justice	SRP	10/26/20	v	0:03:37
C	CH_CTW_SRP_GUN_102620_5_Animism_to_Christianity	SRP	10/26/20	v	0:06:25
C	CH_MKW_CMC_BC_SIL_120619_2_Conversation_Dhirang	CMC_BC	12/06/19	a	0:00:58
C	CH_MKW_MRC_DAM_112819_Conversation_with_Bipana	MRC	11/28/19	a	0:09:41
C	CH_MKW_NC_DAM_112819_1_Conversation_with_Bipana	NC	11/28/19	a	0:13:03
C	CH_MKW_SC_DAM_112819_Conversation_with_Bipana	SC	11/28/19	a	0:05:41
C	CH_MKW_TRC_DAM_112819_Conversation_with_Bipana	TRC	11/28/19	a	0:05:41
C	CH_MKW_GBC_CYO_120119_Conversation_with_Bipana	GBC	12/01/19	a	0:37:40
C	CH_MKW_SKP_DAM_112819_Conversation_with_Bipana	SKP	11/28/19	a	0:08:36
C	CH_MKW_BC_JMC_SIL_120619_3_Witches_Monkeys_Conversation	BC_JMC	12/06/19	a	0:32:10
C	CH_MKW_SPMC_LC_SIL_100921_3_Conversation	SPMC_LC	10/09/21	v	0:13:37
C	CH_MKW_BMB_KW_BAN_110619_Conversation_Minuscule_Worms in love	BMB_KW	11/06/19	a	0:02:08
C	CH_MKW_BMB_BAN_103119_6_Conversation_Wife and husband	BMB	10/31/19	a	0:07:50
C	CH_MKW_BMB_BAN_103119_5_On the way to town	BMB	10/31/19	a	0:02:07
C	CH_MKW_BMB_BAN_103119_9_Mother and daughter	BMB	10/31/19	a	0:05:21
C	CH_MKW_RP_SP_CHI_102519_2_Conversation	RP_SP	10/25/19	a	0:02:00
C	CH_MKW_SMC_AYS_1_112819_Conversation_with_Bipana	SMC	11/28/19	a	0:07:27
C	CH_MKW_SLP_MMRBP_AJI_102519_1_Conversation	SLP_MMRBP	10/25/19	a	0:04:29
C	CH_MKW_SC_PB_BGR_101619_1_Conversation	SC_PB	10/16/19	a	0:04:24
C	CH_MKW_BMB_KW_BAN_103119_10_Conversation_Friends	BMB_KW	10/31/19	a	0:07:17
C	CH_MKW_PC_SPMC_LC_SIL_100921_4_Conversation	PC SPMC LC	10/09/21	v	0:13:34
E	CH_MKW_1_16-17_BMB_BAN_100717_1_E	BMB	10/07/17	a	0:01:51
E	CH_MKW_1_17_BMB_BAN_100717_3_E	BMB	10/07/17	a	0:00:15
E	CH_MKW_1_18_BMB_BAN_100717_4_E	BMB	10/07/17	a	0:01:09
E	CH_MKW_1_19-20_BMB_BAN_100717_6_E	BMB	10/07/17	a	0:01:10
E	CH_MKW_1_19-20_BMB_BAN_100717_7_E	BMB	10/07/17	a	0:00:58
E	CH_MKW_1_21-22_CPR_BAN_100817_1_E	CPR	10/08/17	a	0:02:28
E	CH_MKW_1_BMB_BAN_100517_3_E	BMB	10/05/17	a	0:02:24
E	CH_MKW_1_5_BMB_BAN_100517_1_E	BMB	10/05/17	a	0:03:20
E	CH_MKW_1_6-7_BMB_BAN_100517_2_E	BMB	10/05/17	a	0:03:46
E	CH_MKW_1_22-25_CPR_BAN_100817_2_E	CPR	10/08/17	a	0:03:06
E	CH_MKW_1_28-31_BMB_BAN_100917_1_E	BMB	10/09/17	a	0:07:20
E	CH_MKW_1_33-37_BMB_BAN_101017_E	BMB	10/10/17	a	0:22:46
E	CH_MKW_KW_BAN_101217_3_Verb	KW	10/12/17	a	0:00:20
E	CH_MKW_1_40-47_CPR_BAN_101217_1_E	CPR	10/12/17	a	0:14:03
E	CH_MKW_1_48-52_CPR_BAN_101417_1_Verb	CPR	10/14/17	a	0:19:26
E	CH_MKW_1_73-79_CPR_BAN_102417_2_Verb	CPR	10/24/17	a	1:02:24

E	CH_MKW_BMB_BAN_090118_1_Verb	BMB	09/01/18	a	0:05:03
E	CH_MKW_BMB_BAN_090118_2_Verb	BMB	09/01/18	a	0:00:46
E	CH_MKW_STC_SIL_120619_1_E	STC	12/06/19	a	0:02:47
E	CH_MKW_STC_SIL_120619_2_E_1	STC	12/06/19	a	0:09:53
E	CH_MKW_STC_SIL_120619_2_E_2	STC	12/06/19	a	0:10:17
E	CH_MKW_STC_SIL_120619_2_E_3	STC	12/06/19	a	0:12:07
E	CH_MKW_STC_SIL_120619_2_E_4	STC	12/06/19	a	0:11:37
E	CH_MKW_STC_SIL_120619_2_E_5	STC	12/06/19	a	0:10:55
E	CH_MKW_STC_SIL_120619_3_E	STC	12/06/19	a	0:02:00
E	CH_MKW_STC_SIL_120619_4_E_1	STC	12/06/19	a	0:09:07
E	CH_MKW_STC_SIL_120619_4_E_2	STC	12/06/19	a	0:12:27
Exp	CH_MKW_BMB_BAN_100617_1_Born_in_the_jungle	BMB	10/06/17	a	0:01:53
Exp	CH_MKW_SLP_AJI_102519_1_Move to Ajingare	SLP	10/25/19	a	0:01:19
Exp	CH_MKW_SMYMP_CHI_080118_2_Food_resources	SMYMP	08/01/18	a	0:01:05
Exp	CH_MKW_SMYMP_CHI_080118_4_Seasons	SMYMP	08/01/18	a	0:02:54
Exp	CH_MKW_SMYMP_CHI_080118_5_Chepang_Language	SMYMP	08/01/18	a	0:02:20
Exp	CH_MKW_SMYMP_CHI_080118_Agreement	SMYMP	08/01/18	v	0:00:50
Exp	CH_MKW_KKBP_LPK_080918_4_Family	KKBP	08/09/18	a	0:01:32
Exp	CH_MKW_MRC_LPK_080918_3_Chepang_Language	MRC	08/09/18	a	0:03:59
Exp	CH_MKW_PMRC_LAM_081618_2_Chepang_Language	PMRC	08/16/18	a	0:02:02
Exp	CH_MKW_MRNDC_SIL_081818_2_Chepang_Language_Culture	MRNDC	08/18/18	a	0:15:42
Exp	CH_MKW_MRNDC_SIL_081818_3_Bats	MRNDC	08/18/18	a	0:02:32
Exp	CH_MKW_MRNDC_SIL_081818_4_Fishing	MRNDC	08/18/18	a	0:01:03
Exp	CH_MKW_MRNDC_SIL_081818_5_Bats_Whistle	MRNDC	08/18/18	a	0:01:35
Exp	CH_MKW_MRNDC_SIL_081818_6_Dikalak	MRNDC	08/18/18	a/v	0:02:08
Exp	CH_MKW_MRNDC_SIL_081818_7_Chiuri_Tree_Oil	MRNDC	08/18/18	a/v	0:01:18
Exp	CH_MKW_SRNDC_SIL_081818_Life	SRNDC	08/18/18	a	0:02:56
Exp	CH_MKW_PNC_SIL_081818_2_Bat_hunting	PNC	08/18/18	a	0:02:04
Exp	CH_MKW_PNC_SIL_081818_3_Porcupine_hunting	PNC	08/18/18	a	0:00:47
Exp	CH_MKW_PNC_SIL_081818_4_Chepang_Culture_Language	PNC	08/18/18	a	0:03:14
Exp	CH_MKW_KMC_SK_082918_6_Childhood	KMC	08/29/18	a	0:01:47
Exp	CH_MKW_LC_SIL_113019_1_Cave	LC	11/30/19	a	0:11:46
Exp	CH_CTW_BBC_PID_011520_1_Being_Shaman	BBC	01/15/20	v	0:03:38
Exp	CH_CTW_BBC_PID_011520_3_Witches	BBC	01/15/20	v	0:03:33
Exp	CH_CTW_BBC_PID_011520_4_After_death	BBC	01/15/20	v	0:02:12
Exp	CH_CTW_BBC_PID_011520_5_Underworld	BBC	01/15/20	v	0:04:36
Exp	CH_CTW_BBC_PID_011520_6_Tiger_Spirit	BBC	01/15/20	v	0:01:38
Exp	CH_CTW_BBC_PID_011520_7_Chepang_Raute	BBC	01/15/20	v	0:04:08
Exp	CH_CTW_BBC_PID_011520_8_Chepang_Raute	BBC	01/15/20	v	0:01:13
Exp	CH_CTW_BBC_PID_011520_9_Tiger_Spirit	BBC	01/15/20	v	0:05:56
Exp	CH_CTW_BBC_PID_011520_10_Shaman_healing	BBC	01/15/20	v	0:03:00
Exp	CH_MKW_BBC_SIL_032820_3_Lipi	BBC	03/28/20	a	0:04:14
Exp	CH_MKW_BBC_SIL_032820_2_Pambung	BBC	03/28/20	a	0:06:06
Exp	CH_CTW_ABC_JIM_101920_Agreement	ABC	10/19/20	v	0:00:56
Exp	CH_CTW_BBC_POL_102420_2_Tu'm	BBC	10/24/20	v	0:05:56
Exp	CH_CTW_YMC_TAP_102420_3_Agreement	YMC	10/24/20	v	0:00:48

Exp	CH_CTW_BBC_POL_102520_1_Polkim	BBC	10/25/20	a	0:25:13
Exp	CH_CTW_BBC_POL_102520_2_Masks	BBC	10/25/20	a	0:01:44
Exp	CH_CTW_SRP_GUN_102620_6_Agreement	SRP	10/26/20	v	0:00:29
Exp	CH_CTW_SBPC_DHM_111220_2_Kusunda	SBPC	11/12/20	v	0:03:18
Exp	CH_CTW_ABC_JIM_101920_1_Language and Culture	ABC	10/19/20	v	0:03:39
Exp	CH_CTW_BBC_POL_102420_3_Chepang_Kings	BBC	10/24/20	v	0:06:25
Exp	CH_CTW_BBC_PID_011520_2_Underworld	BBC	01/15/20	v	0:02:06
Exp	CH_MKW_BMB_BAN_090118_6_Love_and_marriage	BMB	09/01/18	a	0:02:32
Exp	CH_MKW_DBC_MAI_1_020220_Local_History	DBC	02/02/20	a	0:05:18
Exp	CH_MKW_MRC_LPK_080918_4_Jungle	MRC	08/09/18	a	0:02:05
Exp	CH_MKW_MRC_LPK_080918_6_Clothes	MRC	08/09/18	a	0:01:51
Exp	CH_MKW_MRNDC_SIL_081818_1_Bee_keeping	MRNDC	08/18/18	v	0:03:02
Exp	CH_MKW_PMRC_LAM_081618_4_Porcupine	PMRC	08/16/18	a	0:02:41
Exp	CH_MKW_PMRC_LAM_081618_5_Jungle_resources	PMRC	08/16/18	a	0:07:21
Exp	CH_MKW_PSC_MAI_012620_1_Becoming_Christian	PSC	01/26/20	a	0:32:24
Exp	CH_MKW_PSC_MAI_012620_2_Becoming_Christian	PSC	01/26/20	a	0:06:29
Exp	CH_MKW_SC_SIL_122619_6_Agreement	SC	12/26/19	a	0:00:52
Exp	CH_CTW_ABC_JIM_101920_2_Language and Culture	ABC	10/19/20	v	0:05:35
Exp	CH_CTW_JBC_BHR_102420_2_Agreement	JBC	10/24/20	v	0:00:33
Exp	CH_CTW_RC_KCR_101920_1_Life	RC	10/19/20	v	0:01:58
Exp	CH_CTW_RLC_JIM_101920_Language and Culture	RLC	10/19/20	v	0:08:28
Exp	CH_CTW_RLC_SMC_JIM_101920_Agreement	RLC_SMC	10/19/20	v	0:01:05
Exp	CH_CTW_SMC_JIM_101920_1_Being_Shaman	SMC	10/19/20	v	0:05:05
Exp	CH_CTW_SMC_JIM_101920_2_Syak_Ja'	SMC	10/19/20	v	0:04:50
Exp	CH_CTW_SP_POL_111420_Agreement	SP	10/14/20	v	0:00:59
Exp	CH_MKW_BBC_SIL_032920_1_Chepang people	BBC	03/29/20	a	0:09:30
Exp	CH_MKW_BBC_SIL_032920_2_My grandfather	BBC	03/29/20	a	0:09:47
Exp	CH_MKW_BBC_SIL_032920_3_Origin_Christianity	BBC	03/29/20	a	0:09:04
Exp	CH_MKW_BMB_BAN_090118_3_Mum_to_child_childhood	BMB	09/01/18	a	0:03:02
Exp	CH_MKW_BMB_BAN_090118_4_Mum_to_child_education	BMB	09/01/18	a	0:02:13
Exp	CH_MKW_BMB_BAN_090118_5_Come_back_again	BMB	09/01/18	a	0:05:22
Exp	CH_MKW_DLC_CYO_120119_Shaman_Life	DLC	12/01/19	a	0:13:39
Exp	CH_MKW_KKBP_LPK_101917_1_Life	KKBP	10/19/17	a	0:07:47
Exp	CH_MKW_KKBP_LPK_101917_3_Chepang_Language	KKBP	10/19/17	a	0:04:03
Exp	CH_MKW_MRC_LPK_080918_1_Life	MRC	08/09/18	a	0:02:22
Exp	CH_MKW_SC_SIL_010220_1_Life	SC	01/01/20	a	0:21:29
Exp	CH_MKW_SC_SIL_010220_2_Life	SC	01/01/20	a	0:10:11
Exp	CH_MKW_SC_SIL_010220_3_Life	SC	01/01/20	a	0:16:25
Exp	CH_CTW_BBC_POL_102420_1_Tu'm	BBC	10/24/20	v	0:02:46
Exp	CH_MKW_SC_BGR_101619_2_Parents	SC	10/16/19	a	0:01:04
Exp	CH_MKW_SC_BGR_101619_3_Village_Life	SC	10/16/19	a	0:03:44
Exp	CH_MKW_ARRKC_CHI_102919_1_Life	ARRKC	10/29/19	v	0:13:43
Exp	CH_MKW_SBC_BGR_101719_1_Life	SBC	10/17/19	a	0:06:31
Exp	CH_MKW_SC_BGR_101619_1_Life	SC	10/16/19	a	0:03:01
Exp	CH_CTW_RC_KCR_101920_3_Agreement	RC	10/19/20	a	0:00:17
Exp	CH_MKW_KKBP_LPK_080918_2_Life	KKBP	08/09/18	a	0:01:05

Exp	CH_MKW_KKBP_LPK_080918_3_Goats_Shelter	KKBP	08/09/18	a	0:00:50
Exp	CH_CTW_JBC_BHR_111720_3_Agreement	JBC	11/17/20	v	0:00:30
Exp	CH_CTW_JBC_BHR_111720_4_Tiger_Spirit	JBC	11/17/20	v	0:01:23
Exp	CH_CTW_KRC_HAT_012120_Being_Shaman	KRC	01/21/20	v	0:18:23
Exp	CH_MKW_RP_CHI_102519_1_Life	RP	10/25/19	a	0:20:06
Exp	CH_MKW_KKBP_LPK_080918_1_Tiger	KKBP	08/09/18	a	0:02:40
Exp	CH_MKW_RP_CHI_102519_2_Agreement	RP	10/25/19	a	0:00:26
Exp	CH_MKW_BMB_BAN_100617_2_Jungle_resources_I	BMB	10/06/17	a	0:01:09
Exp	CH_MKW_BMB_BAN_100617_3_Jungle_resources_II	BMB	10/06/17	a	0:01:35
Exp	CH_MKW_BMB_BAN_100717_2_Monkeys	BMB	10/07/17	a	0:02:04
Exp	CH_MKW_BMB_BAN_100717_4_Corn	BMB	10/07/17	a	0:00:44
Exp	CH_MKW_BMB_BAN_102717_1_Asparagus	BMB	10/27/17	a	0:02:50
Exp	CH_MKW_BMB_BAN_102917_3_Nettle	BMB	10/29/17	a	0:01:25
Exp	CH_MKW_BMB_BAN_103119_1_Church	BMB	10/31/19	a	0:11:39
Exp	CH_MKW_BMB_BAN_103119_3_Morning tea	BMB	10/31/19	a	0:01:42
Exp	CH_MKW_BMB_BAN_103119_4_Meeting with Sahu	BMB	10/31/19	a	0:02:25
Exp	CH_MKW_CPR_BAN_102417_3_Morca_bread	CPR	10/24/17	v	0:00:39
Exp	CH_MKW_KKBP_LPK_080918_6_Church	KKBP	08/09/18	a	0:01:21
Exp	CH_MKW_KW_BAN_100917_1_Jhilo	KW	10/09/17	v	0:01:21
Exp	CH_MKW_KW_BAN_102717_2_Cooking_Frogs	KW	10/27/17	a	0:02:16
Exp	CH_MKW_KW_BAN_110217_1_Rice_harvest	KW	11/02/17	a	0:01:42
Exp	CH_MKW_SCBKC_SIL_081918_6_Agreement	SCBKC	08/19/18	a	0:00:51
Exp	CH_MKW_BMB_BAN_100617_4_Jungle_animals	BMB	10/06/17	a	0:01:39
Exp	CH_MKW_BMB_BAN_103119_2_A monkey came by	BMB	10/31/19	a	0:04:22
Exp	CH_MKW_PMRC_LAM_081618_6_Agreement	PMRC	08/16/18	a	0:00:20
Exp	CH_MKW_1_32-33_KW_BAN_100917_2_Corn_porridge	KW	10/09/17	a	0:02:35
Exp	CH_MKW_BMB_BAN_100617_5_Birth	BMB	10/06/17	a	0:08:24
Exp	CH_MKW_BMB_BAN_103119_7_Elephant	BMB	10/31/19	a	0:04:36
Exp	CH_MKW_BMB_BAN_103119_8_Jungle	BMB	10/31/19	a	0:06:03
Exp	CH_MKW_PSC_MAI_012720_Local_History_2	PSC	01/27/20	v	0:02:21
Exp	CH_MKW_PSC_MAI_012720_Local_History_1	PSC	11/27/20	v	0:14:56
Exp	CH_MKW_PSC_MAI_012720_Local_History_3	PSC	11/27/20	v	0:08:20
Exp	CH_MKW_DKC_MAI_012720_Chepang_Language	DKC	11/27/20	a	0:07:03
Exp	CH_MKW_KKBP_LPK_080918_5_Afternoon_Work	KKBP	08/09/18	a	0:00:42
Exp	CH_MKW_KKBP_LPK_080918_Agreement_Archive	KKBP	08/09/18	a	0:00:45
Exp	CH_MKW_KMC_SK_082918_10_Agreement_Archive	KMC	08/29/18	a	0:00:12
Exp	CH_MKW_LC_SIL_112719_1_Life_Archive	LC	11/27/19	a	0:14:14
N	CH_MKW_CPR_BAN_102817_1_Mit_Co'	CPR	10/28/17	a	0:04:34
N	CH_MKW_SMYMP_CHI_080118_1_Life	SMYMP	08/01/18	a	0:01:31
N	CH_MKW_SMYMP_CHI_080118_3_Childhood	SMYMP	08/01/18	a	0:01:25
N	CH_MKW_PNC_SIL_081818_1_Life	PNC	08/18/18	a	0:01:45
N	CH_MKW_SCBKC_SIL_081918_1_Life	SCBKC	08/19/18	a	0:01:46
N	CH_MKW_SCBKC_SIL_081918_2_Chepang_king	SCBKC	08/19/18	a	0:06:36
N	CH_MKW_SCBKC_SIL_081918_3_Chepang_culture	SCBKC	08/19/18	a	0:05:13
N	CH_MKW_SCBKC_SIL_081918_4_Chak_Ja'	SCBKC	08/19/18	a	0:11:13
N	CH_MKW_SCBKC_SIL_081918_5_Hunting	SCBKC	08/19/18	a	0:09:12

N	CH_MKW_BRC_CYO_120119_Yukdhung	BRC	12/01/19	a	0:12:03
N	CH_MKW_JMC_SIL_120619_1_Barbalyak	JMC	12/06/19	a	0:08:03
N	CH_MKW_JMC_SIL_120619_2_Dung_Raja_ra_Dho'ng	JMC	12/06/19	a	0:04:44
N	CH_MKW_SC_SIL_122619_1_Nakko_co'	SC	12/26/19	v	0:13:14
N	CH_MKW_BBC_SIL_032820_1_Lanrang	BBC	03/28/20	a	0:05:16
N	CH_CTW_YMC_TAP_102420_1_Tantula_ra_Meme_Lan	YMC	10/24/20	v	0:13:20
N	CH_CTW_BBC_GUN_102620_1_Cing_Lan	BBC	10/26/20	v	0:17:31
N	CH_MKW_BMB_BAN_090118_7_Chepang_marriage	BMB	09/01/18	a	0:05:06
N	CH_MKW_BMB_BAN_090118_8_Marriage	BMB	09/01/18	a	0:15:13
N	CH_MKW_DBC_MAI_2_020220_Yuk	DBC	02/02/20	a	0:03:27
N	CH_MKW_DBC_MAI_2_020320_Newa_Dung	DBC	02/03/20	v	0:08:12
N	CH_MKW_KMC_SK_082918_2_The_old_Sune_and_the_Tigers	KMC	08/29/18	a	0:04:15
N	CH_MKW_KMC_SK_082918_3_The_Bear	KMC	08/29/18	a	0:04:19
N	CH_MKW_KMC_SK_082918_7_Chak_Ja'	KMC	08/29/18	a	0:02:39
N	CH_MKW_PBPC_CHI_110219_3_Canoe	PBPC	11/02/19	a	0:00:40
N	CH_MKW_PBPC_CHI_110219_4_Sita and the origin of khar	PBPC	11/02/19	a	0:03:29
N	CH_MKW_PBPC_CHI_110219_5_Being_shaman	PBPC	11/02/19	a	0:01:53
N	CH_MKW_PMRC_LAM_081618_1_Life	PMRC	08/16/18	a	0:02:10
N	CH_MKW_PMRC_LAM_081618_3_Kalitar	PMRC	08/16/18	a	0:04:16
N	CH_MKW_PMRC_SIL_081818_1_Life	PMRC	08/18/18	a	0:09:31
N	CH_MKW_PMRC_SIL_081818_2_Life	PMRC	08/18/18	a	0:01:10
N	CH_MKW_PMRC_SIL_081818_3_Grand_Children	PMRC	08/18/18	a	0:01:18
N	CH_MKW_SC_SIL_122619_2_Dhobini_rani	SC	12/26/19	v	0:14:27
N	CH_CTW_BBC_POL_111720_1_Cing_Lan	BBC	11/17/20	v	0:10:05
N	CH_CTW_BBC_POL_111720_2_Cing_Lan	BBC	11/17/20	v	0:16:08
N	CH_CTW_BBC_POL_111720_3_Cing_Lan	BBC	11/17/20	v	0:08:48
N	CH_CTW_BBC_POL_111720_5_Two Sisters	BBC	11/17/20	v	0:12:47
N	CH_CTW_BBC_POL_111720_6_Jogi	BBC	11/17/20	v	0:24:59
N	CH_CTW_JMC_PYK_101920_Cing_Lan	JMC	10/19/20	v	0:21:31
N	CH_CTW_RC_KCR_101920_2_Myth_Origin	RC	10/19/20	v	0:04:21
N	CH_CTW_RC_KCR_101920_2_Sya'n	RC	10/19/20	v	0:04:23
N	CH_CTW_SP_POL_111420_Ream_Tokrak	SP	10/14/20	v	0:10:58
N	CH_MKW_BBC_SIL_042520_1_Minuscules_Without_Shell_Retelling	BBC	04/25/20	a	0:05:21
N	CH_MKW_BMB_BAN_090118_9_Witch_possession	BMB	09/01/18	a	0:08:47
N	CH_MKW_DBC_MAI_1_020320_The two brothers	DBC	02/03/20	v	0:10:35
N	CH_MKW_KBKC_TNG_102617_1_The_Tiger_and_the_Cat	KBKC	10/26/17	a	0:02:42
N	CH_MKW_KBKC_TNG_102617_2_Two_Mangos'_pits	KBKC	10/26/17	a	0:03:12
N	CH_MKW_KBKC_TNG_102817_2_Ram_and_Sita_wedding	KBKC	10/28/17	a	0:04:12
N	CH_MKW_SC_SIL_010120_1_Biram and Sukram	SC	01/01/20	a	0:02:35
N	CH_CTW_JBC_BHR_111720_2_Story	JBC	11/17/20	v	0:03:44
N	CH_CTW_KC_GUN_073120_1_Imagined_story_love_1	KC	07/31/20	a	0:00:36
N	CH_MKW_SBC_BGR_101719_2_Story	SBC	10/17/19	a	0:03:30
N	CH_CTW_KC_GUN_073120_1_Imagined_story_love_2	KC	07/31/20	a	0:00:55
N	CH_CTW_KMC_TAP_102520_2_The bat and the crab	KMC	10/25/20	v	0:05:30
N	CH_CTW_JBC_BHR_111720_1_Warom_Derum	JBC	11/17/20	v	0:03:43
N	CH_CTW_JBC_BHR_102420_1_Cing_Lan	JBC	10/24/20	v	0:16:26

N	CH_CTW_KMC_TAP_102520_3_The frog and the tiger	KMC	10/25/20	v	0:01:45
N	CH_MKW_BMB_BAN_090118_10_Chak_Ja'	BMB	09/01/18	a	0:02:50
N	CH_MKW_SPMC_SIL_100921_1_Tiɾidu'm_BasΛdu'm	SPMC	10/09/21	v	0:07:08
N	CH_MKW_KKBP_LPK_101917_2_Life	KKBP	10/19/17	a	0:01:41
N	CH_MKW_SPMC_SIL_100921_2_Kammar_wa	SPMC	10/09/21	v	0:07:37
S	CH_MKW_BLC_SIL_113019_1_Shaman_Song	BLC	11/30/19	v	0:16:54
S	CH_MKW_BLC_SIL_113019_2_Becoming_Shaman	BLC	11/30/19	v	0:06:02
S	CH_CTW_YMC_TAP_102420_2_Song_Maru_sarasara	YMC	10/24/20	v	0:03:22
S	CH_CTW_SC_POL_102620_1_Song	SC	10/26/20	v	0:03:09
S	CH_MKW_KBKC_TNG_102817_1_Shaman_Song_Witches	KBKC	10/28/17	a	0:02:05
S	CH_MKW_KMC_SK_082918_4_Shaman_Song	KMC	08/29/18	a	0:01:19
S	CH_MKW_KMC_SK_082918_5_Shaman_Song	KMC	08/29/18	a	0:00:42
S	CH_MKW_KMC_SK_110319_1_Shaman_Song	KMC	11/03/19	v	0:01:39
S	CH_MKW_KMC_SK_110319_2_Shaman_Song	KMC	11/03/19	v	0:00:25
S	CH_MKW_KMC_SK_110319_3_Shaman_Song	KMC	11/03/19	v	0:00:30
S	CH_MKW_KMC_SK_110319_4_Shaman_Song	KMC	11/03/19	v	0:01:31
S	CH_MKW_KMC_SK_110319_6_Shaman_Song	KMC	11/03/19	v	0:01:04
S	CH_MKW_SC_SIL_122619_3_My_father's_Songs	SC	12/26/19	a	0:04:42
S	CH_MKW_SC_SIL_122619_3.1_Song	SC	12/26/19	a	0:00:35
S	CH_MKW_SC_SIL_122619_3.2_Song	SC	12/26/19	a	0:00:39
S	CH_MKW_SC_SIL_122619_3.3_Song	SC	12/26/19	a	0:00:33
S	CH_MKW_SC_SIL_122619_4_Song	SC	12/26/19	a	0:03:49
S	CH_MKW_SC_SIL_122619_5_Song	SC	12/26/19	a	0:05:55
S	CH_CTW_SC_SP_POL_102620_2_Raksi_lo'hang_Song	SC_SP	10/26/20	v	0:02:01
S	CH_MKW_BRC_CYO_120119_2_Song	BRC	12/01/19	a	0:00:39
S	CH_MKW_BRC_CYO_120119_3_Song	BRC	12/01/19	a	0:00:24
S	CH_MKW_KBC_SK_1_110319_Shaman_Songs	KBC	11/03/19	a	0:02:15
S	CH_MKW_KMC_SK_110319_7_Shaman_Song	KMC	11/03/19	a	0:01:41
S	CH_MKW_SC_BGR_101619_4_Songs	SC	10/16/19	a	0:04:20
S	CH_MKW_SC_BGR_101619_5_Songs	SC	10/16/19	a	0:05:15
S	CH_MKW_BRC_CYO_120119_1_Song	BRC	12/01/19	a	0:00:39
S	CH_MKW_CPR_BAN_102417_Song_1	CPR	10/24/17	a	0:00:33
S	CH_MKW_KMC_SK_082918_8_Shaman_Song	KMC	08/29/18	a	0:00:51
S	CH_MKW_MK_BAN_102417_Song_1	MK	10/24/17	a	0:00:25
S	CH_MKW_MK_BAN_102417_Song_2	MK	10/24/17	a	0:02:10

GLOSS ABBREVIATIONS

    I provide the gloss abbreviations used in corpus texts along with the morphological forms (with allomorphs) they functionally describe. These are presented under the form of lists sorted alphabetically either for the gloss abbreviations (List 1) or the morphemes (and allomorphs) (List 2).

List 1. Gloss abbreviations to functions and morphemes (and allomorphs)

gloss	morpheme~allomorph	function
1	=ŋ~ŋʌ	1st person
1.PST	=alaŋ~akaŋ~kaŋ	1st person past
1/3DU	=tɕʌ~tɕ	1st or 3rd person dual
1>2	=ne~tɕe	1st person acting upon 2nd person
1DU.IMP.INTR	=tɕʌ	1st person dual imperative intransitive
1DU.IMP.TR	=tɕu	1st person dual imperative transitive
1DU.INCL	=tʌhj	1st person dual inclusive
1PL	ŋi~ni	1st person plural
1PL.EXCL	=su	1st person plural exclusive
1PL.IMP.INTR	=i	1st person plural imperative intransitive
1PL.IMP.TR	=ni	1st person plural imperative (di)transitive
1PL.INCL	=tʌh	1st person plural inclusive
1SG	ŋa	1st person singular
2	=te	2nd person
2/3.PST	=aka~ka	2nd or 3rd person past intransitive non-singular or (di)transitive
2DU	=dzʌ~dz	2nd person dual
2DU.IMP.INTR	=dzʌ	2nd person dual imperative intransitive
2DU.IMP.TR	=dzu	2nd person dual imperative (di)transitive
2PL.IMP.INTR	=nʌ	2nd person plural imperative intransitive
2PL.IMP.TR	=nu	2nd person plural imperative (di)transitive
2SG	naŋ	2nd person singular

2SG.IMP.INTR	=ʌ	2nd person singular imperative intransitive
2SG.IMP.TR	=u	2nd person singular imperative (di)transitive
3>1DU.INCL/INV	=tʌhj	3rd person acting on 1st person dual inclusive or inverse
3>3DU	=tɕʌ	3rd person dual acted upon by 3rd person in inverse construction
3>3PL	=sʌ	3rd person plural acted upon by 3rd person in inverse construction
3>3SG	=i	3rd person singular acted upon by 3rd person in inverse construction
3O/DIR	=u	3rd person object in direct construction
ABL	=sʌj	ablative
ADD	=ma	additive
ANA	u	anaphoric pronoun
ANT	=jak	anteriority with imperative
ANTIEXP	=se~si~sje~sja	emotional anti-experiencer
ASS	=lʌm	associative plural
ATT	=taŋ	attention
CAT	u, waj	cataphoric pronoun
CAUS	=tak	causative
CERT	=waj	certainty
CL1	=dʑjo~dʑjaŋ, =ota~ta (<N.)	numeral classifier for things/person
CL2	=tɕak	numeral classifier on pronoun following =nis 'two'
CMPR	=sʌj, =bʰʌnda (<N.)	comparative
COM	=kusi	comitative
COND	=ja	conditional
COND.PST	=dik	past conditional
COP	mu~mʌ, kʰe, le~=l, na, ma, hʌ	copula
COP.MIR	rʌjsʌ (<N.)	mirative copula
DAT	=kaj	dative
DEN	-mʌj	denonymic

DES	=*gaɾ*	desiderative
DIR/TR	=*n~nʌ*	direct/transitive used in past tense
DIS	=*paj~pa,* =*kʰe,* =*le,* =*tɕʌhe~tɕʌj~tɕʌhene* (<N.)	phrase level discourse clitic
DIST	*o*	3rd person distal pronoun or demonstrative
DU	=*nis*	3rd person dual on pronoun and nouns
ECHO	diverse	echo word
EPIS	=*a*	epistemic 'for sure, definitely'
ERG	=*i*	ergative
EXCL	=*jal~al*	exclusive
GEN	=*ko*	genitive
IMP.HON	=*mʌj*	honorific imperative
IMPF.NEG	=*dʰaŋ*	imperfect negation
INC	=*kʰʌj*	inchoative
INST	=*i*	instrumental
INT	*ga-~gʌ-~gu-*	interrogative pronoun
INV	=*ta~tʌ~tʰa~tʰʌ*	inverse
IRR	=*tɕja~tɕa*	irrealis
LN	=*ʌ*	loan verbal nativizer
LOC1	=*haŋ~aŋ*	locational/allative
LOC2	=*kʰa*	locational/allative
NEG	=*lʌ~lu~li*	negation
NMZ:ADV1	=*tʌ*	nominalizer: adverbialization
NMZ:ADV2	=*to*	nominalizer: adverbialization
NMZ:H	=*mʌj*	nominalizer: human referent
NMZ:LOC	=*kʌ*	nominalizer: locational
NMZ:REL	=*o*	nominalizer: relativization
NMZ1	=*sa*	nominalizer 1: complementation
NMZ2	=*a~ja*	nominalizer 2: event
NPST	=*na~n*	non-past
ONO	diverse	onomatopoeia
OPT	=*pa*	optative
PART	*hʌj, dʌj~da, kja, ma, ba, ane, asam, kʰwej* (<N.), *lʌw* (<N.), *ni* (<N.), *nitʌ* (<N.), *ɾʌ~ɾa* (<N.)	clause level discourse particle

PERF	*=o*	perfect
PL	*=i*	plural
PL	*=lʌm*	plural
PL.H	*-mi~-mʌj*	plural human on 3rd person pronouns
PRO:LOC	*uh*	locational anaphoric pronoun
PROG	*=dʌj~dʰʌj* (<N)	progressive
PROX	*i*	3rd person proximal pronoun or determiner
PST	*=a~ak~aka~ka*	past
PUR	*=laŋ*	purposive
QTY	*=tɕuk~tɕjuk*	quantity
R/M	*=kaj*	reciprocal/reflexive/middle
REFL	*=sʌ~si*	reflexive
REM	*u*	remote pronoun
REM.PST	*=to*	remote past
REM.PST	*=jak*	remote past with negation
REP	*=tʌ*	reported speech/thought marker
SEQ1	*=ti*	sequential 1
SEQ2	*=jakbʌti~aktiko ~bʌtiko~bʌteko ~bʌti~bʌte~bʌt*	sequential 2
SIM	*=dʰana*	simultaneous
SLF.INTS	*lʌj*	self intensifier
SML	*=lʌm*	similative plural

List 2. Morphemes (and allomorphs) to gloss abbreviations and functions

morpheme~allomorph	gloss	function
-mi~-mʌj	PL.H	plural human on 3rd person pronouns
-mʌj	DEN	denonymic
=a	EPIS	epistemic 'for sure, definitely'
=a~ak~aka~ka	PST	past
=a~ja	NMZ2	nominalizer: event
=aka~ka	2/3.PST	2nd or 3rd person past intransitive non-singular or (di)transitive
=alaŋ~akaŋ~kaŋ	1.PST	1st person past
=dʰana	SIM	simultaneous
=dʰaŋ	IMPF.NEG	imperfect negation
=dik	COND.PST	past conditional
=dʌj~dʰʌj (<N.)	PROG	progressive
=dʑjo~dʑjaŋ, *=ota~ta* (<N.)	CL1	numeral classifier for things/person
=dʑu	2DU.IMP.TR	2nd person dual imperative (di)transitive
=dʑʌ	2DU.IMP.INTR	2nd person dual imperative intransitive
=dʑʌ~dʑ	2DU	2nd person dual
=gaɾ	DES	desiderative
=haŋ~aŋ	LOC1	locational/allative
=i	1PL.IMP.INTR	1st person plural imperative intransitive
=i	3>3SG	3rd person singular acted upon by 3rd person in inverse construction
=i	ERG	ergative
=i	INST	instrumental
=i	PL	plural
=ja	COND	conditional
=jak	ANT	anteriority with imperative
=jak	REM.PST	remote past with negation
=jakbʌti~aktiko *~bʌtiko~bʌteko* *~bʌti~bʌte~bʌt*	SEQ2	sequential 2
=jal~al	EXCL	exclusive
=kaj	DAT	dative

642

=kaj	RM	reciprocal/reflexive/middle
=kʰa	LOC2	locational/allative
=kʰʌj	INC	inchoative
=ko	GEN	genitive
=kusi	COM	comitative
=kʌ	NMZ:LOC	nominalizer: locational
=laŋ	PUR	purposive
=lʌ~lu~li	NEG	negation
=lʌm	ASS	associative plural
=lʌm	PL	plural
=lʌm	SML	similative plural
=ma	ADD	additive
=mʌj	IMP.HON	honorific imperative
=mʌj	NMZ:H	nominalizer: human referent
=n~nʌ	DIR/TR	direct/transitive used in past tense
=na~n	NPST	non-past < *cop *na*
=ne~tɕe	1>2	1st person acting upon 2nd person
=ni	1PL.IMP.TR	1st person plural imperative (di)transitive
=nis	DU	3rd person dual on pronoun and nouns
=nu	2PL.IMP.TR	2nd person dual imperative (di)transitive
=nʌ	2PL.IMP.INTR	2nd person dual imperative intransitive
=ŋ~ŋʌ	1	1st person
=o	NMZ:REL	nominalizer: relativization
=o	PERF	perfect
=pa	OPT	optative
=paj~pa, =kʰe, =le,	DIS	phrase level discourse clitic
=tɕʌhe~tɕʌj~tɕʌhene (<N.)		
=sa	NMZ	nominalizer: complementation
=se~si~sje~sja	ANTIEXP	emotional anti-experiencer
=su	1PL.EXCL	1st person plural exclusive
=sʌ	3>3PL	3rd person plural acted upon by 3rd person in inverse construction
=sʌ~si	REFL	reflexive
=sʌj	ABL	ablative
=sʌj,	CMPR	comparative
=bʰʌnda (<N.)		
=ta~tʌ~tʰa~tʰʌ	INV	inverse

=*tak*	CAUS	causative
=*taŋ*	ATT	attention
=*tɕak*	CL2	numeral classifier on pronoun following =*nis* 'two'
=*tɕja~tɕa*	IRR	irrealis
=*tɕu*	1DU.IMP.TR	1st person dual imperative transitive
=*tɕuk~tɕjuk*	QTY	quantity
=*tɕʌ*	1DU.IMP.INTR	1st person dual imperative intransitive
=*tɕʌ*	3>3DU	3rd person dual acted upon by 3rd person in inverse construction
=*tɕʌ~tɕ*	1/3DU	1st person or 3rd person dual
=*te*	2	2nd person
=*ti*	SEQ1	sequential 1
=*to*	NMZ:ADV2	nominalizer: adverbialization 2
=*to*	REM.PST	remote past
=*tʌ*	NMZ:ADV1	nominalizer: adverbialization 1
=*tʌ*	REP	reported speech/thought marker
=*tʌh*	1PL.INCL	1st person plural inclusive
=*tʌhj*	1DU.INCL	1st person dual inclusive
=*tʌhj*	3>1DU.INCL/INV	3rd person acting on 1st person dual inclusive or inverse
=*u*	2SG.IMP.TR	2nd person singular imperative (di)transitive
=*u*	3O/DIR	3rd person object in direct construction
=*ʌ*	2SG.IMP.INTR	2nd person singular imperative intransitive
=*ʌ*	LN	verbal loan nativizer
=*waj*	CERT	certainty
diverse	ECHO	echo word
diverse	ONO	onomatopoeia
ga-~gʌ-~gu-	INT	interrogative pronoun
hʌj, *dʌj~da*, *kja*, *ma*, *ba*, *ane*, *asam*	PART	clause level discourse particle
kʰwej (<N.), *lʌw* (<N.), *ni* (<N.), *nitʌ* (<N.), *rʌ~ra* (<N.)		
i	PROX	3rd person proximal pronoun or determiner
lʌj	SLF.INTS	self intensifier
mu~mʌ, *kʰe-*, *le~=l*,	COP	copula

na, ma, hʌ		
naŋ	2SG	2nd person singular
ŋa	1SG	1st person singular
ŋi ~ ni	1PL	1st person plural
o	DIST	3rd person distal pronoun and determiner
rʌjsʌ (<N.)	COP.MIR	mirative copula
u	ANA	anaphoric pronoun
u	REM	remote pronoun
u	REM	anaphoric cataphoric deictic pronoun
u, waj	CAT	cataphoric pronoun
uh	PRO:LOC	locational anaphoric pronoun

APPENDIX IV

OTHER ABBREVIATIONS AND SYMBOLS

I describe other types of abbreviations and symbols used in corpus texts (List 3) and in descriptive analyses (List 4).

List 3. Abbreviations and symbols found in corpus text

Transcriptions:

(…)   beginning of sentence preceding or end of sentence following

{…}   a morpheme is missing

//…//   clarification of context

Translations:

(…)   identity of non-overtly expressed participants

alt.   alternative translation

lit.   literal meaning

List 4. Abbreviations and symbols found in descriptive analyses

Languages, language families and Proto-languages:

N.   Nepali

T.   Tamang

E.   English

<   from

arch.   archaic

PTH   Proto-Trans-Himalayan

TH   Trans-Himalayan

PC   Proto-Chepang

PCB   Proto-Chepang-Bhujel

PKM   Proto-Kham-Magar

PH   Proto-Himalayan

IA   Indo-Aryan

Abbreviations used in morphophonological change description

>   becomes

/   in the environment of

__   environment

+   morpheme boundaries

Other abbreviations and symbols:

[…]   phonetic transcription

/…/   phonological transcription

<…>   transliteration of Devanāgarī

*   reconstructed proto-form

reconstructed proto-form whose exact phonological shape needs

       investigation

<   form derived from

>   form developed into

-      affixal boundary

=   clitic boundary

.   syllable boundary

DEVANAGARI TRANSCRIPTION CONVENTIONS

In descriptive text specifically, I use the International Alphabet of Sanskrit Transliteration (IAST), developed by the Tenth International Congress of Orientalists (1894), to transliterate words originally written in Devanāgarī and which specifically come from Nepali or Sanskrit.

The Devanāgarī vowels and basic consonants transliterated in IAST are presented in Table 9 and Table 10, respectively. Note that Devanāgarī consonants include an inherent mid central vowel sound /ʌ/, alternatively described as a schwa /ə/ in the literature, transliterated <a>, as with क /kʌ~kə/ <ka>. Only the other vowels and diphthongs appear overtly attached as diacritic to the consonants.

Additional Devanāgarī diacritics are presented in Table 11. They mark nasalized vowels (*anusvāra* and *anunāsika*), a voiceless fricative glottal (*visarga*), and a final consonant (*halanta*), i.e., a consonant that is not combined with vowel.

For the transliteration of more complex consonants and a comparison of the different transliteration systems that exist, see Shashir (2021).

Table 166. IAST transliteration of Devanāgarī vowels

	Devanāgarī	IAST
	अ	a
	आ, ा	ā
	इ, ि	i
vowels	ई, ी	ī
	उ, ु	u
	ऊ, ू	ū
	ए, े	e

	Devanāgarī	IAST
	ऐ, ै	ai
	ओ, ो	o
	औ, ौ	au

Table 167. IAST transliteration of Devanāgarī consonants

	Devanāgarī	IAST
consonants	क	ka
	ख	kha
	ग	ga
	घ	gha
	ङ	ṅa
	च	ca
	छ	cha
	ज	ja
	झ	jha
	ञ	ña
	ट	ṭa
	ठ	ṭha
	ड	ḍa
	ढ	ḍha
	ण	ṇa
	त	ta
	थ	tha
	द	da
	ध	dha
	न	na

Devanāgarī	IAST
प	pa
फ	pha
ब	ba
भ	bha
म	ma
य	ya
र	ra
ल	la
व	va
श	śa
ष	ṣa
स	sa
ह	ha

Table 168. IAST transliteration of Devanāgarī diacritics

		Devanāgarī	IAST
	anusvāra	◌ं	ṃ
diacritics	anunāsika	◌ँ	m̐
	visarga	◌ः	ḥ
	halanta	◌्	

Note that, in the prose of the dissertation and corpus text, I do not use the IAST to transliterate nouns for languages, persons, and locations; they are transcribed in the Roman alphabet, as if they were words written in English. are not transcribed in International Phonetic Alphabet (IPA) nor in IAST but in Roman alphabet, as if they were English words.

These choices result from the fact that there are already well-established ways of transliterating languages, person names and locations in a Roman alphabet based on English that may be found in the literature, newspaper, official documents, boards, and signs of Nepal. These transcriptions are often recognized by the Chepang community (even when they do not read, speak, or write English) since they are exposed to them in their daily life.

For instance, the proper name Chepang is written in Devanāgarī as चेपाङ् and appears on people's identity cards as <Chepang>. If transliterated in IAST, it becomes <cepāṅ>, a written form that is either not recognized or considered odd or unsuitable by most. The same may apply to other nouns like districts' names such as Chitwan, transcribed in Devanāgarī as चितवन्, and in IAST as <citawān>, while being rather known under the form <Chitwan>. In keeping these words in a separate frozen written form so to speak, we avoid too much incongruity and confusion with what people are already used to seeing.

REFERENCES

Abbi, Anvita. 1994. *Semantic Universals in Indian Languages*. Shimla: Indian Institue of Advenced Study.

Abraham, P. T. 1985. *Apatani grammar*.

Acharya, Jayaraj. 1991. *A descriptive grammar of Nepali and an analyzed corpus*. Georgetown University Press.

Adhikari, Dhruba. 2016. चेपाङ संक्षिप्त शब्दकोश आधारभूत व्याकरण सहित *[A Concise Dictionary of Chepang Language: with Chepang Grammar]*. Lalitpur, Nepal: Nepal Little Flower Society: Nepal Chepang Association // नेपाल क्षलटल फ्लावर सोसाइटी नेपाल चेपाङ संघ.

Adhikari, Krishna. 2010. *The importance of being a pande: A performative reading of the shamanistic practices in Chepang community*. Kirtipur, Nepal: Central Department of English, Tribhuvan University M. Phil Thesis. http://107.170.122.150:8080/xmlui/handle/123456789/124. (6 October, 2020).

Adhikari, Rishiram. 2006. *A sociolinguistic survey of the Chepang language of Chitwan district*. Kirtipur, Nepal: Central Department of English, Tribhuvan University Master's Thesis. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=39822.

Adhikari, Rishiram. 2017. Some Commonly Used Words of Chepang. *The Saptagandaki Journal* 8. 85–92. https://doi.org/10.3126/sj.v8i0.18466.

Aikhenvald, Alexandra Y. 2000. *Classifiers: A typology of noun categorization devices*. Oxford studies in typology and linguistic theory. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. 2003. Typological parameters for the study of clitics, with special reference to Tariana. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *Word: A cross-linguistic typology*, 42–78. 1st edn. Cambridge University Press. https://doi.org/10.1017/CBO9780511486241.003.

Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. In *Language Typology and Syntactic Description*, vol. III: Grammatical Categories and the Lexicon, 32. Timothy Shopen.

Aissen, Judith. 1999. Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory* 17. 673–711.

Aissen, Judith. 2003. Differential object marking: Iconicity versus economy. *Natural Language & Linguistic Theory* 21(3). 435–483.

Allen, Charles. 2016. *The Prisoner of Kathmandu: Brian Hodgson in Nepal, 1820–43*. New Delhi: Speaking Tiger Publishing PVT. LTD.

Allen, Nicholas J. 1975. *Sketch of Thulung grammar* (Cornell University East Asia Papers 6). Ithaca, N.Y.: Cornell University.

Anthony, Laurence. 2019. AntConc. Waseda University, Japan: Faculty of Science and Engineering.

Arkadiev, Peter. 2020. Non-canonical inverse in Circassian languages. *STUF - Language Typology and Universals* 73(1). 81–111. https://doi.org/10.1515/stuf-2019-0028.

Armbrecht Forbes, Ann. 1996. The discourse and practice of Kipat. *Kailash* 18(1/2). 39–80.

Arnold, David. 2004. Hodgson, Hooker, and the Himalayan frontier. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 189–205. London: Routledge Curzon.

Bandhu, Chuda Mani. 1989. The role of the Nepalese language in establishing the national unity and identity of Nepal. *Kailash*.

Baral Magar, Keshar Jung. 2012. पाल्पा तनहुँ र स्याँङजाको मगरहरुको संस्कृति *Palpa,Tanahu ra Syangjako Magarharuko Sanskriti [The culture of the Magar of Palpa, Tanahu and Syangja]*. Second. Kathmandu: Asmita Books.

Bárány, András, Oliver Bond & Irina Nikolaeva. 2019. Towards a typology of prominent internal possessors. In András Bárány, Oliver Bond & Irina Nikolaeva (eds.), *Prominent Internal Possessors*. Oxford University Press.

Barðdal, Jóhanna & Shobhana Chelliah (eds.). 2009. *The role of semantic, pragmatic, and discourse factors in the development of case*. Amsterdam and Philadelphia: Benjamins.

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

Bauman, James. 1975. *Pronouns and pronominal morphology in Tibeto-Burman*. Berkeley, California: University of California at Berkeley Doctoral Dissertation.

Beames, John. 1868. *Outlines of Indian Philology, with a Map Shewing the Distribution of Indian Languages*. 2nd edn. London, England: Trübner & Company. https://books.google.fr/books?id=vbgOAAAAQAAJ.

Benedict, Paul K. 1941. Tibetan and Chinese kinship terms. *Harvard Journal of Asiatic Studies* 6. 313–337.

Benedict, Paul K. 1972. *Sino-Tibetan: a conspectus* (Princeton-Cambridge Studies in Chinese Linguistics 2). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511753541.

Bickel, Balthasar. 1994. *Mapping operations in spatial deixis and the typology of reference frames*. Working Papers of the Cognitive Anthropology Research Group. Nijmegen: Max Planck Institute for Psycholinguistics.

Bickel, Balthasar. 1995. In the vestibule of meaning: Transitivity inversion as a morphological phenomenon. *Studies in Language* 19. 73–127.

Bickel, Balthasar. 1997. Spatial operations in deixis, cognition, and culture: Where to orient oneself in Belhare. In Jan Nuyts & Eric Pederson (eds.), *Language and conceptualization*, 46–83. Cambridge: Cambridge University Press.

Bickel, Balthasar. 2003. Belhare. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 546–70. London; New York: Routledge.

Bickel, Balthasar. 2008. On the scope of the referential hierarchy in the typology of grammatical relations. In Greville G. Corbett & Michael Noonan (eds.), *Case and grammatical relations. Papers in honor of Bernard Comrie.*, 191–210. Amsterdam, Philadelphia: John Benjamins.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, 169–240. Cambridge: Cambridge University Press (Revised second edition).

Bista, Dor Bahadur. 1967. *People of Nepal*. 1st edn. Kathmandu: Department of Publicity, ministry of information and broadcasting, his Majesty's Govt. of Nepal.

Bista, Dor Bahadur. 1987. *People of Nepal*. 5th edn. Kathmandu: Ratna Pustak Bhandar.

Bista, Dor Bahadur. 1991. *Fatalism and development: Nepal's struggle for modernization*. Orient Blackswan.

Bista, Dor Bahadur & Navin Kumar Rai. 1975. *Chepang Jatiko Samajik Arthik Sammunati Sarbekchan (Pariyojana Pratibedan)*. Kathmandu: CEDA, Tribhuvan University.

Bodt, Timotheus A. & Ismael Lieberherr. 2015. First notes on the phonology and classification of the Bangru language of India. *Linguistics of the Tibeto-Burman Area* 38(1). 66–123. https://doi.org/10.1075/ltba.38.1.03bod.

Boersma, Paul & David Weenink. 2021. Praat: doing phonetics by computer [Computer program]. http://www.praat.org/.

Bohora, Bhugol. 2002. *Socio-economic condition of Chepang of Chitwan District*. Kirtipur, Nepal: Tribhuvan University Doctoral Dissertation. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=35862.

Bornkessel-Schlesewsky, Ina & Matthias Schlesewsky. 2009. The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistics Compass* 3(1). 19–58.

Bossong, Georg. 1985. *Differenzielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, Georg. 1998. Le marquage différentiel de l'objet dans les langues d'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues de l'Europe*, 193–258. Berlin & New York: Mouton de Gruyter.

Bowerman, Melissa & Eric Pederson. 1992. Topological relations picture series. In Stephen C. Levinson (ed.), *Space stimuli kit*, vol. 1:2. Nijmegen: Max Planck Institute for Psycholinguistics.

Bradley, David. 1997. Tibeto-Burman languages and classification. In David Bradley (ed.), *Tibeto-Burman languages of the Himalayas, Papers in Southeast Asian linguistics*, vol. 14, 1–72. Canberra: Pacific Linguistics.

Bradley, David. 2002. The Subgrouping of Tibeto-Burman. In Christopher Beckwith (ed.), *Medieval Tibeto-Burman Languages*, 73–112. Leiden: Brill. https://doi.org/10.1163/9789047401308_010.

Bradley, David. 2018. Subgrouping of the Sino-Tibetan languages. *Evolang* 10. 11.

Breugel, Seino van. 2020. New perspectives on Atong kinship terms. In Q. Marak & Sarit K. Chaudhuri (eds.),. Oxford, New-York: Routledge.

Bunsen, Christian Charles Josias. 1854. *Outlines of the Philosophy of Universal History, applied to Language and Religion* (Christianity and Mankind, Their Beginnings and Prospects 3 of 7). Vol. 1. London, England: Longman, Brown, Green, and Longmans.

Burghart, Richard. 1984. The formation of the concept of Nation-State in Nepal. *The Journal of Asian Studies* 44(1). 101–125. https://doi.org/10.2307/2056748.

Burghart, Richard. 1996. The conditions of listening: The everyday experiences of politics in Nepal. In C. J. Fuller & Jonathan Spencer (eds.), *The conditions of listening: essays on religion, history and politics in South Asia*, 300–318. New Delhi: Oxford University Press.

Button, Christopher T. J. 2009. *A reconstruction of Proto-Northern-Chin in Old Burmese and Old Chinese perspective*. London, England: School of Oriental and African Studies Doctoral Dissertation.

Catford, John Cunnison. 1988. *A practical introduction to phonetics*. 1 online resource
(xiii, 239 pages) : illustrations vols. Oxford [England]: Clarendon Press ;
http://books.google.com/books?id=bEViAAAAMAAJ.

Caughley, Ross Charles. 1969a. *Chepang phonemic summary* (Tibeto-Burman Phonemic
Summaries 4). Kirtipur, Kathmandu, Nepal: Summer Institute of Linguistics and
Tribhuvan University. https://reap.sil.org/handle/9284745/36909. (7 November,
2020).

Caughley, Ross Charles. 1969b. Report of initial survey of Chepang language group.
Manuscript. Dallas, Texas, ms. 50623. SIL International Language & Culture
Archive. https://reap.sil.org/handle/9284745/50623. (8 October, 2020).

Caughley, Ross Charles. 1970a. Chepang segmental synopsis. In Austin Hale & Kenneth
L. Pike (eds.), *Tone Systems of the Tibeto-Burman Languages of Nepal: Studies
on Tone and Phonological Segments* (Occasional Papers of the Wolfenden
Society on Tibeto-Burman Linguistics 3), vol. 1, 279–299. Urbana, Illinois: The
University of Illinois.

Caughley, Ross Charles. 1970b. Chepang texts. In Austin Hale & Kenneth L. Pike (eds.),
Chepang texts, 279–299.

Caughley, Ross Charles. 1970c. Pitch, intensity, and higher levels in Chepang. In Austin
Hale & Kenneth L. Pike (eds.), *Tone Systems of the Tibeto-Burman Languages of
Nepal: Studies on Tone and Phonolgical Segments* (Occasional Papers of the
Wolfenden Society on Tibeto-Burman Linguistics 3), vol. 1, 143–157. Urbana,
Illinois: The University of Illinois.
https://hdl.handle.net/2027/mdp.39015019962417?urlappend=%3Bseq=159. (8
October, 2020).

Caughley, Ross Charles. 1971. Chepang as a pronominalised language: technical note.
Kathmandu, Nepal, ms. 49660: SAG-CDM-017. SIL International Language &
Culture Archive. https://www.reap.insitehome.org/handle/9284745/37045. (6
October, 2020).

Caughley, Ross Charles. 1972. *A vocabulary of the Chepang language* (Comparative Vocabularies of Languages of Nepal). Kirtipur, Nepal: Summer Institute of Linguistics and Institute of Nepal Studies, Tribhuvan University. https://www.sil.org/resources/archives/36902.

Caughley, Ross Charles. 1978. Participant rank and verbal cross reference in Chepang. In Joseph E. Grimes (ed.), *Papers on discourse* (Summer Institute of Linguistics Publications in Linguistics and Related Fields 51), 163–178. Dallas, Texas: Summer Institute of Linguistics. https://www.reap.insitehome.org/handle/9284745/8588. (13 October, 2020).

Caughley, Ross Charles. 1980. *The syntax and morphology of the verb in Chepang*. Canberra, Australia: Australian National University Doctoral Dissertation.

Caughley, Ross Charles. 1982. *The syntax and morphology of the verb in Chepang* (Pacific Linguistics 84). Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University. https://doi.org/10.15144/pl-b84.

Caughley, Ross Charles. 1998. Bujheli and Chepang: relationship and differences. In. Pune, India, December 1998. https://reap.sil.org/handle/9284745/86378. (7 November, 2020).

Caughley, Ross Charles. 2000. *Dictionary of Chepang: a Tibeto-Burman language of Nepal* (Pacific Linguistics 502). Canberra, Australia: Research School of Pacific and Asian Studies, Australian National University. https://doi.org/10.15144/PL-502.cover.

Caughley, Ross Charles. 2015. The syntax and morphology of the verb in Chepang. Canberra: Department of Linguistics, Research School of Pacific Studies, Australian National University.

Caughley, Ross Charles. 2016. Chepang Dictionary. SIL International.

Caughley, Ross Charles, Chuda Mani Bandhu & Ballabh Mani Dahal. 1970. Chepang segmental phonemes. *Journal of the Tribhuvan University* 5(1). 65–81. (12 October, 2020).

Caughley, Ross Charles, Churamani Bandhu & Ballabh Mani Dahal. 1969. Notes on Chepang culture. *Journal of the Tribhuvan University*. https://reap.sil.org/handle/9284745/36881. (15 October, 2020).

Caughley, Ross Charles, Ballabh Mani Dahal & Churamani Bandhu. 1971. Notes on Chepang culture. *Journal of the Tribhuvan University* 6(1). 77–89. https://reap.sil.org/handle/9284745/36881. (15 October, 2020).

Central Bureau of Statistics, National Planning Commission Secretariat. 2012. *Nepal Population and Housing Census 2011*. Kathmandu, Nepal.

Charencey, Hyacinthe de. 1862. Les dialectes du Népaul. *Revue Orientale et Américaine* 7. 57–58. https://gallica.bnf.fr/ark:/12148/bpt6k9680422f/f69.image.r=chepang.

Chelliah, Shobhana L. 2009. Semantic role to new information in Meithei. In Jóhanna Barðdal & Shobhana L. Chelliah (eds.), *The role of semantic, pragmatic, and discourse factors in the development of case*, 377–400.

Cho, Taehong, James M. McQueen & Ethan A. Cox. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2). 210–243. https://doi.org/10.1016/j.wocn.2006.03.003.

Chou, Fa-kao. 1972. Archaic Chinese and Sino-Tibetan. *Journal of the Institute of Chinese, Studies of the Chinese University of Hong Kong* 5(1). 159–237.

Comrie, Bernard. 1978. Ergativity. In Winfred P. Lehmann (ed.), *Syntactic typology: Studies in the phenomenology of language.*, 329–394. Austin: University of Texas Press.

Comrie, Bernard. 1980. Inverse verb forms in Siberia: Evidence from Chukchee, Koryak and Kamchadal. *Folia Linguistica Historica* 1(1). 61–74.

Comrie, Bernard. 1981. *Language universals and linguistic typology syntax and morphology*.

Comrie, Bernard. 1989. *Language universals and linguistic typology*. *Language universals and linguistic typology: Syntax and morphology*. Chicago: University of Chicago Press.

Corbett, Greville G. 2000. *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. & Marianne Mithun. 1996. Associative forms in a typology of number systems: evidence from Yup'ik. *Journal of Linguistics* 32. 1–17.

Creissels, Denis. 2009. Spatial cases. In Andrej L. Malchukov & A. Spencer (eds.), *The Oxford Handbook of Case*, 609–625. Oxford: Oxford University Press.

Cristofaro, Sonia. 2013. The referential hierarchy: Reviewing the evidence in diachronic perspective. In Dik Bakker & Martin Haspelmath (eds.), *Languages across boundaries*, 69–94. Berlin, New York: Mouton de Gruyter.

Croft, William. 1990. *Typology and universals*. Cambridge: Cambridge University Press.

Cust, Robert Needham. 1878. *A sketch of the modern languages of the East Endies*. London, England: Trübner & Company. https://books.google.fr/books?id=Eh4gAAAAMAAJ.

Dahal, Peshal & Ghanashyam Bhattarai. 2019. *History of Nepal customs*. Department of Customs Nepal.

Daniel, Michael & Edith A. Moravcsik. 2005. Associative plurals. In Martin Haspelmath, David Gil & Bernard Comrie (eds.), 150–153.

Daniel, Michael & Edith A. Moravcsik. 2013. The associative plural. In Martin Haspelmath & Matthew S. Dryer (eds.), *The world atlas of language structures online.* Leipzig: Max Planck: Max Planck Institute for Evolutionary Anthropology.

Das Gupta, K. 1971. *An Introduction to the Nocte language*. Shillong: North-East Frontier Agency.

Datta, Ann. 2004. Brian Hodgson and the mammals and other animals of Nepal. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 154–171. London: Routledge Curzon.

Davids, Irene M. H. & George van Driem. 1985. Limbu kinship terminology, a description. *Kailash* 115–156.

DeLancey, Scott. 1981. The category of direction in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 6(1). 83–101.

DeLancey, Scott. 1985. The Analys-Synthesis-Lexis Cycle in Tibeto-Burman: a case study in motivated change. *Typological Studies in Language* 6. 367–389.

DeLancey, Scott. 2010. Towards a history of verb agreement in Tibeto-Burman. *Himalayan Linguistics* 9(1). 1–39. https://doi.org/10.5070/H99123042.

DeLancey, Scott. 2011. "Optional" "ergativity" in Tibeto-Burman languages. *Linguistics of the Tibeto-Burman Area* 34(2).

DeLancey, Scott. 2013. The origins of Sinitic. In Zhuo Jing-Schmidt (ed.), *Increased empiricism: Recent advances in Chinese linguistics*, 73–100. Amsterdam, Philadelphia: John Benjamins. https://doi.org/10.1075/scld.2.04del.

DeLancey, Scott. 2015a. Morphological Evidence for a Central Branch of Trans-Himalayan (Sino-Tibetan). *Cahiers de Linguistique d'Asie Orientale* 44(2). 122–149. https://doi.org/10.1163/19606028-00442p02.

DeLancey, Scott. 2015b. The historical dynamics of morphological complexity in Trans-Himalayan. *Linguistics Discovery* 13. 60–79. https://doi.org/10.1349/PS1.1537-0852.A.463.

DeLancey, Scott. 2018. Deictic and sociopragmatic effects in Tibeto-Burman SAP indexation. In *Typological hierarchies in synchrony and diachrony*, 345–375. Amsterdam: John Benjamins.

DeLancey, Scott. 2019. Non-singular pronouns in Tibeto-Burman (Trans-Himalayan). *Linguistics of the Tibeto-Burman Area* 42(1). 110–136. https://doi.org/10.1075/ltba.18006.del.

Dhakal, Dubi Nanda. 2014. *A Grammar of Majhi*. Munich: Lincom Europa.

Dhungel, Ram Mani. 1994. चेपाङ समुदाय र संस्कृति *<Cepāṅg samudāya ra saṃskṛti>* *[Social life and customs of Chepang, Nepalese people]* (Saṃskaraṇ 1). Kāṭhamāḍauṃ: Sājhā Prakāśan.

Diessel, Holger. 1999. *Demonstratives: Form, function and grammaticalization*. John Benjamins. https://www.jbe-platform.com/content/books/9789027298577.

Diffloth, Gérard. 1976. Expressives in Semai. *Oceanic Linguistics Special Publications*. University of Hawai'i Press (13). 249–264. http://www.jstor.org/stable/20019159. (24 July, 2022).

Dirksmeyer, Tyko. 2008. *Spatial Deixis in Chintang. Aspects of a Grammar of Space*. Leipzig, Germany: Institut für linguistik der Philologischen Fakultät der Universität Leipzig Master's Thesis.

Dixon, R. M. W. 1979. Ergativity. *Language* 55. 59–138.

Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Doherty, Victor. 1974. The organizing principles of Gurung kinship. *Kailash* 2. 273–301.

Doornenbal, Marius. 2009. *A grammar of Bantawa: grammar, paradigm tables, glossary, and texts of a Rai language of Eastern Nepal*. Leiden, The Netherlands: Universiteit Leiden Doctoral Dissertation.

Driem, George van. 1987. *A grammar of Limbu* (Mouton Grammar Library 4). Berlin ; New York: Mouton de Gruyter. https://doi.org/10.1515/9783110846812.

Driem, George van. 1990. An exploration of Proto-Kiranti verbal morphology. *Acta Linguistica Hafniensia* 22. 27–28.

Driem, George van. 1991a. Bahing and the Proto-Kiranti verb. *Bulletin of the School of Oriental and African Studies* LIV(2). 336–356.

Driem, George van. 1991b. Tangut verbal agreement and the patient category in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* LIV(3). 520–534. http://www.himalayanlanguages.org/files/driem/pdfs/tangut_low.pdf.

Driem, George van. 1992. Le Proto-Kiranti revisité, morphologie du lohorung. *Acta Linguistica Hafniensia* 24. 33–75.

Driem, George van. 1993a. *A grammar of Dumi*. Berlin: Mouton de Gruyter.

Driem, George van. 1993b. The Proto-Tibeto-Burman verbal agreement system. *Bulletin of the School of Oriental and African Studies, University of London* 56(2). 292–334.

Driem, George van. 1997. Sino-Bodic. *Bulletin of the School of Oriental and African Studies* 60. 455–488.

Driem, George van. 2001. *Languages of the Himalayas: an ethnolinguistic handbook of the greater Himalayan region; containing an introduction to the symbiotic theory of language. (Handbuch der Orientalistik 2, Abteilung: Indien, 10.)*. Leiden: Brill. https://doi.org/10.1163/9789004492530.

Driem, George van. 2007. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1(2). 211–270. https://doi.org/10.1163/2405478X-90000023.

Driem, George van. 2011. Tibeto-Burman subgroups and historical grammar. *Himalayan Linguistics* 10(1). 31–39.

Driem, George van. 2014. Trans-Himalayan. In Thomas Owen-Smith & Nathan W. Hill (eds.), *Trans-Himalayan Linguistics: Historical and Descriptive Linguistics of the Himalayan Area*, 11–40. Berlin, Boston: De Gruyter Mouton.

Dryer, Matthew S. 1986. Primary Objects, Secondary Objects and Antidative. *Language* 62. 808–845.

Dryer, Matthew S. 1997. Are grammatical relations universal? In Joan L. Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type: Dedicated to T. Givon*, 115–143. Amsterdam: John Benjamins.

Eagle, Sonia. 1999. The language situation in Nepal. *Journal of Multilingual and Multicultural Development* 20(4–5). 272–327.

Eberhard, David M., Gary F. Simons & Charles D. Fennig. 2020. Chepang. *Ethnologue: Languages of the World*. Dallas, Texas: SIL International. https://www.ethnologue.com/language/CDM. (7 October, 2020).

Ebert, Karen H. 1991. Inverse and pseudo-inverse prefixes in Kiranti languages: Evidence from Belhare, Athpare and Dungmali. *Linguistics of the Tibeto-Burman Area* 14(1). 73–92.

Ebert, Karen H. 1997a. *A grammar of Athpare*. München: Lincom Europa.

Ebert, Karen H. 1997b. *Camling (Chamling)* (Languages of the World Materials 103). München: Lincom Europa.

Ebert, Karen H. 1999. The up-down dimension in Rai grammar and mythology. In Balthasar Bickel & Martin Gaenszle (eds.), 107–134.

Fauconnier, Stephanie. 2011. Differential agent marking and animacy. *Lingua* 121(3). 533–547.

Filimonova, Elena. 2005. The noun phrase hierarchy and relational marking: Problems and counterevidence. *Linguistic Typology* 9. 77–113.

Forbes, Charles James. 1877. Affinities of the dialects of the Chepang and Kusundah Tribes of Nipál with those of the hill tribes of Arracan. *The Journal of the Royal Asiatic Society of Great Britain and Ireland New Series* 9(2). 421–424.

Forbes, Charles James. 1878. On Tibeto-Burman languages. *Journal of the Royal Asiatic Society of Great Britain and Ireland* 10(2). 210–227. https://www.jstor.org/stable/25196796. (16 October, 2020).

Forbes, Charles James. 1881. *Comparative Grammar of the Languages of Further India: A Fragment and other essays*. London, England: W.H. Allen & Company. https://hdl.handle.net/2027/coo1.ark:/13960/t8gf1d12k. (5 October, 2020).

Gaenszle, Martin. 2004. Brian Hodgson as ethnographer and ethnologist. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 204–226. London: Routledge Curzon.

Gaige, Frederick H. 2009. *Regionalism and national unity in Nepal*. 2nd edition. Kathmandu: Himal Books.

Gast, Volker, Lisa Deringer, Florian Haas & Olga Rudolf. 2015. Impersonal uses of the second person singular: A pragmatic analysis of generalization and empathy effects. *Journal of Pragmatics* 88. 148–162. https://doi.org/10.1016/j.pragma.2014.12.009.

Gautam, Rajesh & Asoke K. Thapa-Magar. 1994. *Tribal ethnography of Nepal*. Vol. I. Book Faith India.

Gellner, David N., Joanna Pfaff-Czarnecka & John Whelpton (eds.). 1997. *Nationalism and Ethnicity in a Hindu Kingdom: The Politics of Culture in Contemporary Nepal*. Amsterdam: Harwood Academic Publishers.

Genetti, Carol. 1990. Dolakhali (Newari) wordlist (unpublished ms. contributed to STEDT database). *Sino-Tibetan Etymological Dictionary and Thesaurus - STEDT*. http://stedt.berkeley.edu/search/.

Genetti, Carol. 2003. Dolakhā Newar. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 355–370. London; New York: Routledge.

Genetti, Carol. 2007. *A grammar of Dolakha Newar* (Mouton Grammar Library 40). Berlin; New York: Mouton de Gruyter. https://doi.org/10.1515/9783110198812.

Genetti, Carol, Kristine Hildebrandt, Nathaniel A. Sims & Alexia Z. Fawcett. 2021. Direction and associated motion in Tibeto-Burman. *Linguistic Typology* 25(2). 345–388. https://doi.org/10.1515/lingty-2020-2064.

Gildea, Spike & Fernando Zúñiga. 2016. Referential hierarchies: A new look at some historical and typological patterns. *Linguistics* 54(3). https://doi.org/10.1515/ling-2016-0007.

Givón, Talmy. 1983. Topic continuity in discourse: A quantitative cross-language study. In *Typological Studies in Language*, vol. 3. John Benjamins. Amsterdam.

Givón, Talmy. 1994. The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Talmy Givón (ed.), *Voice and inversion*, 3–44. Amsterdam & Philadelphia: John Benjamins.

Glover, Warren W. 1972. *A vocabulary of the Gurung language* (Comparative Vocabularies of Languages of Nepal). Kirtipur: Summer Institute of Linguistics and Institute of Nepal Studies, Tribhuvan University.

Glover, Warren W., Jessie R. Glover & Deu Bahadur Gurung. 1977. *Gurung-Nepali-English dictionary, with English-Gurung and Nepali-Gurung Indexes* (Pacific Linguistics Series C, No 51). Australian National University. Camberra.

Gong, Xun. 2014. The personal agreement system of Zbu Rgyalrong (Ngyaltsu variety). *Transactions of the Philological Society* 112(1). 44–60. https://doi.org/10.1111/1467-968X.12007.

Grierson, George Abraham. 1909. *Tibeto-Burman families. Linguistic survey of India*. Vol. 3. Calcutta: Superintendent of Government Printing.

Grunow-Hårsta, Karen. 2008. *A descriptive grammar of two Magar dialects of Nepal: Tanahu and Syangja Magar*. Milwaukee, Wisconsin: University of Wisconsin-Milwaukee Doctoral Dissertation.

Gurung, Ganesh Man. 1970. Economic modernization in a Chepang village in Nepal. *Occasional Papers in Sociology and Anthropology* 2. 32–39. https://doi.org/10.3126/opsa.v2i0.1067.

Gurung, Ganesh Man. 1987. A note on the religious beliefs and practices among the Chepang of Nepal. *Contributions to Nepalese Studies* 14(3). 239–246. http://www.digitalhimalaya.com/collections/journals/contributions/index.php?selection=14_3. (7 October, 2020).

Gurung, Ganesh Man. 1989. *The Chepangs: A study in continuity and change*. Kathmandu: Centre for Nepal and Asian Studies, Tribhuvan University.

Gurung, Ganesh Man. 1994. A note on the religious beliefs and practices among the Chepangs of Nepal. In *Indigenous peoples: mobilization and change*, 69–84. Kathmandu, Nepal: S. Gurung (self-published).

Hargreaves, David. 2003. Kathmandu Newār (Nepāl Bhāṣā). In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 371–384. London; New York: Routledge.

Haspelmath, Martin. 2019. Differential place marking and differential object marking. *STUF - Language Typology and Universals* 72(3). 313–334. https://doi.org/10.1515/stuf-2019-0013.

Haspelmath, Martin. 2021. Role-reference associations and the explanation of argument coding splits. *Linguistics* 59(1). 123–174. https://doi.org/10.1515/ling-2020-0252.

Haspelmath, Martin & Andrea D. Sims. 2010. *Understanding morphology* (Understanding Language Series). 2nd ed. London: Hodder Education.

Haudricourt, André-Georges. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* (242). 69–82.

Heath, Jeffrey. 1991. Pragmatic disguise in pronominal-affix paradigms. In Frans Planck (ed.), *Paradigms: The Economy of Inflection*, 75–89. Berlin: Mouton de Gruyter.

Heine, Bernd & Kyung-An Song. 2011. On the grammaticalization of personal pronouns. *Journal of Linguistics* 47(3). 587–630. https://doi.org/10.1017/S0022226711000016.

Herbert, Captain J. D. 1842. Report upon the mineralogical survey of the Himalayan mountains. *Journal of the Asiatic Society of Bengal* 11. 126.

Hildebrandt, Kristine. 2007. Tone in Bodish languages: Typological and sociolinguistic contributions. In Matti Miestamo & Bernhard Wälchli (eds.), *Trends in Linguistics. Studies and Monographs [TiLSM]*. Berlin, New York: Mouton de Gruyter. https://doi.org/10.1515/9783110198904.1.77.

Hill, Nathan W. 2014. Cognates of Old Chinese *-n, *-r, and *-j in Tibetan and Burmese. *Cahiers de Linguistique Asie Orientale* 43(2). 91–109. https://doi.org/10.1163/19606028-00432p02.

Hitan Magar, Jaya Bahadur. 1992. <Mangol janajātimā hindu dharmako prabhava> [The influence of Hindu religion on the Mangol nationalities]. *Lapha* 1(2). 5–7.

Hitchcock, John T. 1966. *The Magars of Banyan Hill*. New-York: Holt, Rinehart and Winston.

Hodgson, Brian Houghton. 1848. On the Chépáng and Kúsúnda tribes of Népál. *Journal of the Asiatic Society of Bengal* 17(2). 655–658.

Hodgson, Brian Houghton. 1857a. Comparative vocabulary of the languages of the broken tribes of Nepal. *Journal of the Asiatic Society of Bengal* 26(5). 317–349.

Hodgson, Brian Houghton. 1857b. On the Chépáng and Kúsúnda tribes of Népál. *Journal of the Asiatic Society of Bengal* 46–54. https://973c63ce-3086-49ea-bb9b-6f566a8e84d2.filesusr.com/ugd/e3da62_fff263457ffb48358322f38cddbe292f.pdf. (7 October, 2020).

Hodgson, Brian Houghton. 1874a. *Essays on the languages, literature and religion of Nepál and Tibet*. London: Trübner & Company.

Hodgson, Brian Houghton. 1874b. On the Chépáng and Kúsúnda tribes of Népál. In *Essays on the Languages, Literature, and Religion of Nepál and Tibet, Part 2*, 45–54. London: Trübner & Company.

Hodgson, Brian Houghton. 1880. *Miscellaneous essays relating to Indian subjects Vol. II*. Vol. 2. London: Trübner & Company.

Hodgson, Brian Houghton. 1991. On the Chépáng and Kúsúnda tribes of Népál. In *Essays on the Languages, Literature, and Religion of Nepál and Tibet*, 45–54. New Delhi, India: Asian Educational Services.

Höfer, András. 1979. *The Caste Hierarchy and the State of Nepal: A study of the Muluki Ain of 1854*. Innsbruck: Universitätsverlag Wagner.

Holmberg, David H. 2005. *Order in paradox: Myth, ritual, and exchange among Nepal's Tamang*. Delhi: Motilal Banarsidass.

Hooker, Joseph Dalton. 1824. *Mammal & bird notes made*. Zoological Society of London Library.

Hooker, Joseph Dalton. 1848. *Himalayan journals*. Royal Botanic Gardens Kew Library.

Huang, Bufan & Qingxia Dai (eds.). 1992. 藏緬語族語言詞匯 *<Zangmianyuzu yuyan cihui> [A Tibeto-Burman Lexicon]*. Beijing: Central Institute of Minorities.

Hunter, William Wilson. 1868. *A comparative dictionary of the non-Aryan languages of India and HIgh Asia*. London: Trübner & Company.

Inskipp, Carol & Ann Datta. 2004. Zoology... amuses me much. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 134–153. London: Routledge Curzon.

Jacques, Guillaume. 2010. The inverse in Japhug Rgyalrong. *Language and Linguistics* 11(1). 127–157.

Jacques, Guillaume. 2015a. On the cluster *sr– in Sino-Tibetan. *Journal of Chinese Linguistics* 43(1A). 215–223.

Jacques, Guillaume. 2015b. The origin of the causative prefix in Rgyalrong languages and its implication for proto-Sino-Tibetan reconstruction. *Folia Linguistica* 36(1). https://doi.org/10.1515/flih-2015-0002.

Jacques, Guillaume. 2016. Tonogenesis and tonal alternations in Khaling. In *Tone and Inflection*, 41–66. Berlin, Boston: De Gruyter Mouton. https://doi.org/10.1515/9783110452754-003.

Jacques, Guillaume. 2017. The genetic position of Chinese. In *Encyclopedia of Chinese Language and Linguistics*, vol. 2, 297–306. Brill.

Jacques, Guillaume. 2018. Fossil nominalization prefixes in Tibetan and Chinese. LFK Society Young Scholars Symposium, Taipei.

Jacques, Guillaume. 2019. The labial causative in Trans-Himalayan. *Journal of the Southeast Asian Linguistics Society* 12(1). 1–11.

Jacques Guillaume. 2019. 李方桂先生調查西南少數民族語言的足跡. *Bulletin of Chinese Linguistics* 12(1). 3–12. https://doi.org/10.1163/2405478X-01201003.

Jacques, Guillaume. 2021. *A grammar of Japhug*. https://openresearchlibrary.org/content/dfa742f3-6e13-4025-a826-bb1656ff126a. (9 January, 2022).

Jacques, Guillaume. Draft. Chapter 15: Origins of Sino-Tibetan, bringing in evidence from history, biology and archaeology. 63.

Jacques, Guillaume. Draft. Sino-Tibetan/Trans-Himalayan. In Peter Arkadiev & Franz Rainer (eds.), *The Oxford handbook of historical morphology*.

Jacques, Guillaume & Anton Antonov. 2014. Direct/Inverse Systems. *Language and Linguistics Compass* 8(7). 301–318.

Jacques, Guillaume, Aimée Lahaussois, Boyd Michailovsky & Dhan Bahadur Rai. 2012. An overview of Khaling verbal morphology. *Language and Linguistics* 13(6). 1095–170.

Jest, Corneille. 1965a. *Conversation :30-02*. Audio/x-wav [171.8MB]. Village de Kandrang Ghari (West n°2). https://archives.crem-cnrs.fr/archives/items/CNRSMH_I_1967_006_030_02/. (7 October, 2020).

Jest, Corneille. 1965b. *Essais de chants :31-02*. Audio/x-wav [10.2MB]. Village de Kandrang Ghari (West n°2). https://archives.crem-cnrs.fr/archives/items/CNRSMH_I_1967_006_031_02/. (7 October, 2020).

Jest, Corneille. 1965c. *Questions et enquête :31-01*. Audio/x-wav [223.8MB]. Village de Kandrang Ghari (West n°2). https://archives.crem-cnrs.fr/archives/items/CNRSMH_I_1967_006_031_01/. (7 October, 2020).

Jest, Corneille. 1966. Les Chepang: ethnie népalaise de langue Tibéto-Birmane. *Objets et Mondes* 6(2). 169–184.

Jest, Corneille. n.d. Les Chepang: Note concernant une ethnie de langue tibéto-birmane au Népal. Paris, France, ms. 25315. The Centre for Himalayan Studies (CEH) — Fonds documentaire CEH, Paris France. https://campus-condorcet.primo.exlibrisgroup.com/permalink/33CCP_INST/iv4mie/alma991007019309705786.

Kansakar, Tej Ratna. 2010. Baram wordlist (unpublished ms. contributed to STEDT).
   Sino-Tibetan Etymological Dictionary and Thesaurus - STEDT.
   http://stedt.berkeley.edu/search/.

Kansakar, Tej Ratna, Yogendra Prasad Yadava, Krishna Prasad Chalise, Balaram
   Prasain, Dubi Nanda Dhakal & Krishna Paudel. 2014. A sociolinguistic study of
   the Baram language. *Himalayan Linguistics* 10(1). 187–225.
   https://doi.org/10.5070/H910123574.

Keenan, Edward L. 1984. Semantic correlates of the ergative/absolutive distinction.
   Linguistics 22. 197–223.

Khanal, Baikuntha. 2014. *Social exclusion of the Chepangs in the era of post-political
   revolution in Nepal: A case study of Dhading district*. Bergen, Norway: The
   University of Bergen Master's Thesis. http://hdl.handle.net/1956/8507. (7
   October, 2020).

Khatri, Anjana. 2017. *Demographic, socio-economic and educational status of Chepang
   community*. Kirtipur, Nepal: Tribhuvan University Master's Thesis.

Kim, Sahyang & Taehong Cho. 2013. Prosodic boundary information modulates phonetic
   categorization. *The Journal of the Acoustical Society of America* 134(1). EL19–
   EL25. https://doi.org/10.1121/1.4807431.

Kim, Sahyang, Holger Mitterer & Taehong Cho. 2018. A time course of prosodic
   modulation in phonological inferencing: The case of Korean post-obstruent
   tensing. (Ed.) Jiwon Hwang. *PLOS ONE* 13(8). e0202912.
   https://doi.org/10.1371/journal.pone.0202912.

Klaiman, M. H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.

Klaproth, Julius. 1823. *Asia Polyglotta*. Paris: B.A. Shubart.

Klausenburger, Jurgen. 2000. *Grammaticalization: Studies in Latin and Romance
   morphosyntax*. John Benjamins. https://www.jbe-
   platform.com/content/books/9789027299628.

Konnerth, Linda. 2021. On the nature of inverse systems: The rise of inverse marking via antipassive constructions. *Diachronica* 38(1). 25–63. https://doi.org/10.1075/dia.18055.kon.

Konnerth, Linda Anna. 2014. *A grammar of Karbi*. Eugene, Oregon: University of Oregon Doctoral Dissertation.

Konnerth, Linda & Andrea Sansò. 2021. Towards a diachronic typology of individual person markers. *Folia Linguistica* 55(s42–s1). 1–24. https://doi.org/10.1515/flin-2021-2012.

Krishan, Shree. 2001. A sketch of Raji grammar. In Yasuhiko Nagano & Randy J. LaPolla (eds.), *New research on Zhangzhung and related Himalayan languages*, 449–497. Osaka: National Museum of Ethnology.

Krishan, Shree. 2003. Darma, Chaudangsi, and Raji. In Randy J. LaPolla (ed.), *Tibeto-Burman Languages of Uttar Pradesh*, 139–272.

Ladefoged, Peter. & Keith Johnson. 2011. *A course in phonetics*. 6th ed., International ed. Boston, MA: Wadsworth/Cengage Learning. https://library.dctabudhabi.ae/sirsi/detail/1213706.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages* (Phonological Theory). Oxford, OX, UK ; Cambridge, Mass., USA: Blackwell Publishers.

Lahaussois, Aimée. 1998. Aspects of the grammar of Thulung Rai: An endangered Himalayan language.

LaPolla, Randy J. 1992. On the dating and nature of verb agreement in Tibeto-Burman. *Bulletin of the School of Oriental and African Studies* 55(2). 298–315. https://doi.org/10.1017/S0041977X00004638.

LaPolla, Randy J. 1994. Parallel grammaticalization in Tibeto-Burman Languages Evidence of Sapir's "Drift."

LaPolla, Randy J. 1995. "Ergative" Marking in Tibeto-Burman. In Yoshio Nishi, James A Matisoff & Yasuhiko Nagano (eds.), *New horizons in Tibeto-Burman morphosyntax* (Senri Ethnological Studies 41), 189–228. Osaka, Japan: National Museum of Ethnology.

LaPolla, Randy J. 2007. Hierarchical person marking in the Rawang language. In, 9. Harbin China.

Latham, Robert Gordon. 1860. On the varieties of the human species. In *The Circle of the Sciences: A series of Treatise on the principles of science, with their application to practical pursuits* (Organic Nature 1), 305–376. New Edition. London, England: Richard Griffan & Company; University of Glasgow.

Latham, Robert Gordon. 1862. *Elements of Comparative Philology*. London, England: Walton and Maberly. https://archive.org/details/elementsofcompar00lathrich.

Lawoti, Mahendra. 2005. *Towards a democratic Nepal: Inclusive political institutions for a multicultural society*. New Delhi: Sage Publications.

Lawoti, Mahendra & Susan Hangen (eds.). 2012. *Nationalism and Ethnic Conflict in Nepal: Identities and Mobilization after 1990*. Routledge. https://doi.org/10.4324/9780203118269.

Lazard, Gilbert. 2001. Le marquage différentiel de l'objet. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and language universals: An international handbook*, vol. 2, 873–885. Berlin & New York: Walter de Gruyter.

Lecomte-Tilouine, Marie. 1993. The proof of the bone: Lineage and devālī in Central Nepal. *Contributions to Indian Sociology* 27(1). 1–23. https://doi.org/10.1177/006996693027001001.

Lecomte-Tilouine, Marie. 2010. To be more natural than others: Indigenous self-determination and Hinduism in the Himalayas. In Marie Lecomte-Tilouine (ed.), *Nature, Culture, and Religion at the Crossroads of Asia*, 118–157. New Delhi: Social Science Press.

Levine, Nancy E. 1987. Caste, state, and ethnic boundaries in Nepal. *Journal of Asian Studies* 46(1). 71–88.

Levinson, Stephen C. 2003. *Space in language and cognition*. Cambridge: Cambridge University Press.

Lieberherr, Ismael. 2017. *A grammar of Bulu Puroik*. Bern, Switzerland: Universität Bern Doctoral Dissertation.

Lockwood, Hunter T. & Monica Macaulay. 2012. Prominence hierarchies. *Language and Linguistics Compass* 6(7). 431–446.

Logan, J. R. 1856. Ethnology of the Indo-Pacific islands. *The Journal of the Indian Archipelago and Eastern Asia* (New Series) New Series(11). §15. https://books.google.fr/books?id=eXMEAAAAQAAJ.

Lot-Falck, É. 1961. A propos d'un tambour de Chamane Tounguse. *L'homme* 1(2).

Malchukov, Andrej L. 2008. Animacy and asymmetries in differential case marking. *Lingua* 118. 203–221.

Malchukov, Andrej L. & Peter De Swart. 2008. Differential case marking and actancy variations. In Andrej L. Malchukov & Andrew Spencer (eds.), *The Oxford book of case*, 339–355. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199206476.013.0023.

Manandhar, Badri Prasad. 2000. *Ethnobiology of the Chepang, a case study of Makwanpur District*. Kirtipur, Nepal: Tribhuvan University Master's Thesis. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=25876.

Manandhar, Narayan P. 1989. Medicinal plants used by Chepang tribes of Makawanpur District, Nepal. *Fitoterapia* 60(1). 61–68.

Masica, Colin. 1991. *The Indo-Aryan languages* (Cambridge Language Surveys). Cambridge University Press. https://doi.org/10.1017/S0022226700015383.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia.

Matisoff, James A. 1991. Sino-Tibetan linguistics: Present state and future prospects. *Annual Review of Anthropology*. Annual Reviews 20. 469–504. http://www.jstor.org/stable/2155809. (3 July, 2022).

Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: system and philosophy of Sino-Tibetan reconstruction*. Berkeley: University of California Press. https://books.google.com/books?hl=en&lr=&id=fp8lDQAAQBAJ&oi=fnd&pg=PR1&dq=tibeto-burman&ots=R5S1YMeA6R&sig=AJg0JVbsGr5jsh88Ot6w5UuTCEY.

Matisoff, James A. 2015. Sino-Tibetan Etymological Dictionary and Thesaurus. 1185.

Mauri, Caterina & Andrea Sansò. 2018. Linguistic strategies for ad hoc categorization: theoretical assessment and cross-linguistic variation. *Folia Linguistica* 52(s39–s1). 1–35. https://doi.org/doi:10.1515/flih-2018-0001.

Mauri, Caterina & Andrea Sansò. 2019. Nouns & co. converging evidence in the analysis of associative plurals. *STUF - Language Typology and Universals* 72(4). 603–626. https://doi.org/doi:10.1515/stuf-2019-0023.

Mauri, Caterina & Andrea Sansò. 2021. Heterogeneous sets: a diachronic typology of associative and similative plurals. *Linguistic Typology* 000010151520212072. https://doi.org/doi:10.1515/lingty-2021-2072.

Mazaudon, Martine. 1978. Consonantal mutation and tonal split in Tamang Sub-Family of Tibeto-Burman. *Kailash* (6.3). 157–79.

Mazaudon, Martine. 2003. Tamang. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 291–314. London; New York: Routledge.

Mazaudon, Martine & Alexis Michaud. 2008. Tonal contrasts and initial consonants: A case study of Tamang, a 'missing link' in tonogenesis. *Phonetica* 65(4). 231–256. https://doi.org/10.1159/000192794.

McGregor, William. 2010. Optional ergative case marking systems in a typoogical-semiotic perspective. *Lingua* 120(7). 1610–1636.

McGregor, William & Jean-Christophe Verstaete. 2010. Optional ergative marking and its implications for linguistic theory. *Lingua* 120(7). 1607–1609.

Michailovsky, Boyd. 1988. *La langue Hayu*. Paris: Centre National de la Recherche Scientifique.

Michailovsky, Boyd. 1989a. Bahing (unpublished ms. contributed to STEDT). *Sino-Tibetan Etymological Dictionary and Thesaurus - STEDT.*

Michailovsky, Boyd. 1989b. Limbu (unpublished ms. contributed to STEDT). http://stedt.berkeley.edu/search/.

Michailovsky, Boyd. 1994. Manner versus place of articulation in the Kiranti initial stops. In Kitamura Hajime, Tatsuo Nishida & Yasuhiko Nagano (eds.), *Current Issues in Sino-Tibetan Linguistics.*, 766–781. Osaka, Japan: The Organizing Committee of the 26th International Conference on Sino-Tibetan Languages and Linguistics.

Michailovsky, Boyd. 2003. Hayu. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 518–532. London; New York: Routledge.

Michaud, Alexis & Bonny Sands. 2020. Tonogenesis. In *Oxford Research Encyclopedia of Linguistics*. Oxford University Press. https://doi.org/10.1093/acrefore/9780199384655.013.748.

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60. 847–94.

Moravcsik, Edith A. 1978. On the case marking of objects. In Joseph H. Greenberg (ed.), *Universals of human language* (Syntax), vol. 4, 249–289. Stanford, CA: Standford University Press.

Moravcsik, Edith A. 2004. A semantic analysis of associative plurals. *Studies in Language* (27). 469–503.

Moronval, Frédéric. 2017. *Vitalités linguistique et religieuse chez les Néwar bouddhistes de la vallée de Kathmandu*. Rouen, France: Université de Rouen Normandie Doctoral Dissertation. https://tel.archives-ouvertes.fr/tel-01697607. (30 October, 2020).

Moseley, Christopher (ed.). 2010. *Atlas of the World's Languages in Danger*. 3rd edn. Paris: UNESCO Publishing.

Müller, Friedrich Max. 1854. *Letter to Chevalier Bunsen on the Classification of the Turanian Languages*. London, England: A. & G.A. Spottiswoode. https://books.google.fr/books?id=d5leAAAAcAAJ.

Nebesky-Wojkowitz, René von. 1959. Kusunda and Chepang: Notes on two little-known tribes of Nepal. *Bulletin of the International Committee on Urgent Anthropological and Ethnological Research* 2. 77–84. https://973c63ce-3086-49ea-bb9b-6f566a8e84d2.filesusr.com/ugd/e3da62_df8165d35d7043fb9df3dc12ee7f6421.pdf. (7 October, 2020).

Nepal National Educational Planning Commission. 1956. *Education in Nepal: Report of the Nepal education planning commission.* Kathmandu, Nepal: His Majesty's Government.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Noonan, Michael. 2003. Chantyal. In Graham Thurgood & Randy J. LaPolla (eds.), *The Sino-Tibetan languages*, 315–335. London; New York: Routledge.

Opgenort, Jean-Robert M. L. 2004. *A grammar of Wambule: Grammar, lexicon, texts and cultural survey of a Kiranti tribe of eastern Nepal*. Leiden: Brill.

Oppitz, Michael. 1982. Death and kin amongst the Northern Magar. *Kailash* 377–421.

Oppitz, Michael. 1988. *Die Dreierallimaz bet den Magar West-Nepal. [Woman against service: The tripartite alliance system among the Magars of Western Nepal]*. Frankfurt am Main: Suhrkamp.

Pandey, Sudarshan. 2001. *Altitudinal impacts on vegetation distribution and forest utilization along with socio-economic pattern of Chepang community: a case study of Brashbang hill, Dhading District Central Nepal*. Kirtipur, Nepal: Tribhuvan University Doctoral Dissertation. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=25150.

Patna, Dinesh Raj. 1968. वि. सं. १८३२ देखि वि. सं. १८५८ सम्मको नेपालको इतिहासमा नयाँ प्रकाश पार्ने केही पत्रहरु <b.s. 1832 dekhi b.s. 1858 sammako Nepālko itihāsmā nayā prakāsh parne kehi patraharu> [Some documents to shed new light on the history of Nepal from 1775 to 1801 CE]. *Purnima* 5(1). 27–51.

Paudyal, Netra Prasad. 2003. *A sketch grammar of Darai*. Kirtipur, Nepal: Tribhuvan University Master's Thesis.

Paudyal, Netra Prasad. 2015. Aspects of Chintang syntax. University of Zurich. https://doi.org/10.5167/UZH-113367.

Pignède, Bernard. 1966. *Les Gurungs [The Gurung]*. The Hague: Mouton.

Pittman, Richard S. 1970. Gurung, Tamang, Thakali, Sherpa, and Chepang prosodies. In Austin Hale & Kenneth L. Pike (eds.), *Tone systems of the Tibeto-Burman languages of Nepal 2: Lexical lists and comparative studies* (Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics 3), vol. 2, 1–8. Urbana, Illinois: The University of Illinois. https://www.sil.org/resources/archives/36972.

Pokharel, Kalpana. 2002. *Socio-economic condition of Chepang in Shaktikhor VDC, Chitwan*. Kirtipur, Nepal: Tribhuvan University Doctoral Dissertation. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=30047.

Pons, Marie-Caroline. 2017. Morphological evidence for "Raji-Raute" and its genetic position within Trans-Himalayan. Berner Zirkel für Sprachwissen-schaft, University of Bern.

Pons, Marie-Caroline. 2021. On the origin of 2nd person prefix #tV- in Trans-Himalayan languages. *Linguistics of the Tibeto-Burman Area* 44(2). 215–254. https://doi.org/10.1075/ltba.20004.pon.

Popham, Barbara. 1978. Testing the Chepang primer. *READ: The Adult Literacy & Literature Magazine* 13(1). 26–28. https://www.sil.org/resources/archives/23322. (2 October, 2020).

Post, Mark W. 2019. Topographical deixis in Trans-Himalayan (Sino-Tibetan) languages. *Transactions of the Philological Society* 117(2). 234–255. https://doi.org/10.1111/1467-968X.12155.

Post, Mark W. 2020. The distribution, reconstruction and varied fates of topographical deixis in Trans-Himalayan (Sino-Tibetan): Implications for the reconstruction of an early Trans-Himalayan environment. *Diachronica* 37(3). 368–409. https://doi.org/10.1075/dia.19018.pos.

Pradhan, Kumar. 1991. *The Gorkha conquests: The process and consequences of the unification of Nepal, with particular reference to eastern Nepal*. Calcutta; New York: Oxford University Press.

R Core Team. 2021. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org.

Rai, Bal Krishna & Klaus Peter Kuegler. 1975. *A vocabulary of the Danuwar Rai language with Nepali and English*. Kirtipur: Summer Institute of Linguistics, Institute of Nepal and Asian Studies.

Rai, Krishna Prasad, Anna Holzhausen & Andreas Holzhausen. 1975. *Kulung body part index from: Kulung-Nepali-English glossary*. Kathmandu: Summer Institute of Linguistics and Institute of Nepal and Asian Studies, Tribhuvan University.

Rai, Navin Kumar. 1985. *People of the Stones: The Chepang of Central Nepal*. Kathmandu, Nepal: Centre for Nepal & Asian studies (CNAS), Tribhuvan University.

Rai, Novel Kishore. 1985. *A descriptive study of Bantawa*. Poona: Deccan College Post-Graduate and Research Institute.

Rastogi, Kavita. 2012. *A descriptive grammar of Raji*. Delhi: Aviram Prakashan.

Reetz, Henning. 2021. Pitch Praat Script 6.1.2. GitHub. https://github.com/HenningReetz/Praat-scripts/blob/main/Pitch/Pitch_6_1_2.praat.

Regmi, Dan Raj. 2007. *The Bhujel language*. Kirtipur, Nepal: Tribhuvan University Doctoral Dissertation. (5 October, 2020).

Regmi, Dan Raj. 2009. Participant reference in Bhujel, Chepang and Banakariya. *Nepalese Linguistics* 24. 299–310.

Regmi, Dan Raj. 2012. *A grammar of Bhujel*. München: Lincom Europa.

Regmi, Dilli Raman. 1961. *Modern Nepal: Rise and growth in the eighteenth century*. Calcutta: Firma K. L. Mukhopadyaya.

Regmi, Mahesh C. (ed.). 1970. *Regmi Research Series*. Vol. 2. Kathmandu: Regmi Research (Private) Ltd.

Regmi, Mahesh C. 1974. Notes on the Kipat system V. *Regmi Research Series* 6(9).

Regmi, Mahesh C. 1976. *Landownership in Nepal*. Berkeley: University of California Press.

Regmi, Mahesh C. 1978. *Land tenure and taxation in Nepal*. Kathmandu: Ratna Pustak Bhandar.

Regmi, Pratigya. 2017. Thangmi surnames. *Nepalese Linguistics* 32. 102–109.

Riboli, Diana. 1993. Shamanic paraphernalia and dances among the Chepang shamans (South Nepal). In Mihály Hoppál & Pál Páricsy (eds.), *Shamanism and Performing Arts*, 122–127. Budapest, Hungary: Ethnographic Institute, Hungarian Academy of Sciences.

Riboli, Diana. 1994a. Shamanic practices in the Chepang of Central and Southern Nepal. Presented at the 13th International Conference Anthropological and Ethnological Sciences, Mexico City, Mexico.

Riboli, Diana. 1994b. Shamanic rites of the Terai Chepangs. *East and West*. Istituto Italiano per l'Africa e l'Oriente (IsIAO) 44(2/4). 327–352. https://www.jstor.org/stable/29757157. (7 October, 2020).

Riboli, Diana. 2000. *Tunsuriban. Shamanism in the Chepang of Southern and Central Nepal*. (Trans.) Philippa Currie. Kantipath, Kathmandu, Nepal: Mandala Book Point.

Riboli, Diana. 2004. Extrasensory trances and trances of movement: Reflections on different types of trances. In Art Leete & R. Paul Firnhaber (eds.), *Shamanism in the Interdisciplinary Context: Papers from the 6th Conference of the International Society for Shamanistic research, Viljandi, Estonia, August 2001*, 120–130. Boca Raton, Florida: Brown Walker Press.

Riboli, Diana. 2011. 'We play in the black jungle and in the white jungle'. The Forest as a Representation of the Shamanic Cosmos in the Chants of the Semang-Negrito (Malaysia) and the Chepang (Nepal). *Shaman* 19. 153–168.

Riboli, Diana. 2014. Dreamed violence and shamanic transformation in Nepal and Malaysia. In Ryan Hurd & Kelly Bulkeley (eds.), *Lucid Dreaming: New Perspectives on Consciousness in Sleep — Science, Psycology, and Education* (Practical and Applied Psychology), vol. 1, 73–91. Santa Barbara, California: Praeger. https://books.google.fr/books?id=uRGJBAAAQBAJ&pg=RA1-PA73.

Riboli, Diana. 2020. Shamanism and gender (in)equality in South and South-east Asia: The Chepang of Nepal and the Semang-Negrito of Peninsular Malaysia. In Torri Davide & Sophie Roche (eds.), *The Shamaness in Asia: Gender, Religion and the State* (Vitality of Indigenous Religions). London, England: Routledge. https://doi.org/10.4324/9780429319808.

Riboli, Diana. Forthcoming. The threatening invisibility of the Christian God and its consequences among the Chepang (Nepal). In Anne De Sales & Marie Lecomte Tilouine (eds.), *Encounters with the Invisible in the Himalaya*. Delhi, India: Oxford University Press.

Rose, Françoise. 2015. When "You" and "I" mess around with the hierarchy: a comparative study of Tupi-Guarani hierarchical indexing systems [Quando "Tu" e "Eu" se mexem na hierarquia: um estudo comparativo de sistemas de indexação hierárquica Tupí-Guaraní]. *Boletim do Museu Paraense Emílio Goeldi* 10(2). 347–362.

Rutgers, Roland. 1993. The Chepang verb: A synchronic analysis with some diachronic observations. *Acta Linguistica Hafniensia* 26(1). 107–137. https://doi.org/10.1080/03740463.1993.10415456.

Sagart, Laurent, Guillaume Jacques, Yunfan Lai, Robin J. Ryder, Valentin Thouzeau, Simon J. Greenhill & Johann-Mattis List. 2019. Dated language phylogenies shed light on the ancestry of Sino-Tibetan. *Proceedings of the National Academy of Sciences* 116(21). 10317–10322. https://doi.org/10.1073/pnas.1817972116.

Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* 13. 250–82.

Sapkota, Suren & Ram K. Uranw. 2013. *A sociolinguistic survey of Chepang: A report submitted to the Linguistic Survey of Nepal (LinSuN)*. Kathmandu, Nepal: Tribhuvan University. https://web.archive.org/web/20180831100942/http://cdltu.edu.np/site/images/linsun/chepang.pdf.

Schackow, Diana. 2015. *A grammar of Yakkha*. Language Science Press. https://doi.org/10.26530/OAPEN_603340.

Schlossberg, Jonathan. 2017. "Local nouns" and the differential marking of place. Presented at the 12th Conference of the Association of Linguistic Typology, Canberra, Australia.

Schmidt, Ruth Laila. 1994. *A practical dictionary of modern Nepali*. Ratna Sagar.

Schorer, Nicolas. 2016. *The Dura language: grammar and phylogeny* (Brill's Tibetan Studies Library volume 5/17). Leiden; Boston: Brill. https://doi.org/10.1163/9789004326408.

Selkirk, Elisabeth. 1986. On derived domains in sentence phonology. *Phonology* 3. 371–405.

Seržant, Ilja A. & Alena Witzlack-Makarevich. 2018. *Diachrony Of Differential Argument Marking*. Zenodo. https://doi.org/10.5281/ZENODO.1219168.

Shafer, Robert. 1955. Classification of the Sino-Tibetan languages. *Word* 11(1). 94–111. https://doi.org/10.1080/00437956.1955.11659552.

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

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

Shashir, Reddy. 2021. Modern Transcription of Sanskrit. *Shashir's Notes*. http://shashir.autodidactus.org/shashir_umich/sanskrit_tra nscription.html.

Shibatani, Masayoshi. 2006. On the conceptual framework for voice phenomena. *Linguistics* 44(2). 217–269. https://doi.org/10.1515/LING.2006.009.

Shibatani, Masayoshi. 2019. What is nominalization? Towards the theoretical foundations of nominalization. In Roberto Zariquey & David Fleck (eds.), *Nominalization in languages of the Americas* (Typological Studies in Language), 15–167. Amsterdam: John Benjamins Publishing Company. https://doi.org/10.1075/tsl.124.02shi.

Shneiderman, Sara & Mark Turin. 2006. Revisiting ethnography, recognizing a forgotten people: The Thangmi of Nepal and India. *Studies in Nepali History and Society* 11(1). 97–181.

Shrestha, Bal Gopal. 2007. Ethnic nationalism in Nepal and the Newars. In Mahendra Lawoti (ed.), *Contentious politics and democratization in Nepal*, 199–225. Los Angeles, London, New Delhi, Singapore: Sage Publications.

Shrestha, Basan. 1997. *Forest management practices and socio-economic determinants in Commercial utilization of chiuri in Chepang Community*. Kirtipur, Nepal: Tribhuvan University Master's Thesis. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=12045.

Shrestha, Nisha. 2004. *Socio-economic condition of Chepang; a case study of Hadikhola VDC of Makawanpur District*. Kirtipur, Nepal: Tribhuvan University Doctoral Dissertation. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=36661.

Shuya, Zhang, Guillaume Jacques & Lai Yunfan. 2019. A study of cognates between Gyalrong languages and Old Chinese. *Journal of Language Relationship* 17(1–2). 73–92. https://doi.org/10.31826/jlr-2019-171-210.

Siewierska, Anna. 1998. On nominal and verbal person marking. *Linguistic Typology* 2. 1–56.

Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.

Smith, Brianne J. 2022. *A sociolinguistic study of Pahari: A language of Nepal*. SIL International.

Song, Jae J. 2001. *Linguistic typology: morphology and syntax*. Harlow, Essex: Longman.

Stirling, Lesley & Lenore Manderson. 2011. About you: Empathy, objectivity and authority. *Journal of Pragmatics* 43(6). 1581–1602. https://doi.org/10.1016/j.pragma.2010.12.002.

Stolz, Thomas, Sander Lestrade & Christel Stolz. 2014. *The crosslinguistics of Zero-Marking of spatial relations*. De Gruyter (A). https://doi.org/doi:10.1524/9783050065304.

Sun, Hongkai & et al. 1991. 藏缅语语音和词汇 *<Zangmianyu yuyin he cihui> [Tibeto-Burman phonology and lexicon]*. Beijing: Chinese Social Sciences Press.

Sun, Jackson T. S. & Shidanluo. 2002. Caodeng Jiarongyu yu rentong dengdi xiangguan de yufa xianxiang [empathy hierarchy in Caodeng rGyalrong grammar]. *Language and Linguistics* 3(1). 79–99.

Sun, Tianshin Jackson. 1993. Tani synonym sets unpublished ms. contributed to STEDT). *Sino-Tibetan Etymological Dictionary and Thesaurus - STEDT*. http://stedt.berkeley.edu/search/.

Sun, Tianshin Jackson. 1994. Linguistic characteristics of the Tani (Mirish) branch of Tibeto- Burman. *Bulletin of the Institute of History and Philology, Academica Sinica* 65(1). 175–220.

Tachibana, Kenichi. 2009. 〈*Tasha / Jiko*〉*Hyoushou no Minzokushi: Neparu Senjumin Tyepan no Mikurosonzairon [An ethnography of other / self Representations: Micro-ontology of the Chepang; An indigenous people of Nepal]*. Tokyo: Fukyosha.

Talmy, Leonard. 1972. *Semantic structures in English and Atsugewi*. Berkeley, California: Department of Linguistics, University of California Doctoral Dissertation.

Talmy, Leonard. 1983. Spatial orientation: Theory, research, and application. In Herbert L. Pick & Linda P. Acredolo (eds.), *How language structures space*. New York and London: Plenum Press. https://doi.org/10.1007/978-1-4615-9325-6_11.

Tautscher, Gabriele. 2007. *Himalayan mountain cults: Sailung, Kalingchok, Gosainkund territorial rituals and Tamang histories*. Kathmandu: Vajra Publications.

Thapa, B. B. 1974. Chepang Haru Sanga Kehi Din (A Few Days with the Chepangs). *Madhuparka* 7. 78–68.

Thapaliya, Bhojraj. 1987. *Chepang jati ek parichaya*. Kathmandu : gopal thapaliya. http://107.170.225.98/cgi-bin/koha/opac-detail.pl?biblionumber=7267.

Thompson, Chad L. 1990. On the Treatment of Topical Objects in Chepang: Passive or Inverse? *Studies in Language* 14(2). 405–427. https://doi.org/10.1075/sl.14.2.06tho.

Thurgood, Graham. 1985. Pronouns, verb agreement systems, the sub-grouping of Tibeto-Burman.

Thurgood, Graham. 2007. Tonogenesis revisited: Revising the model and the analysis. In J. G. Harris, Burusphat Somsonge & J. E. Harris (eds.), *Studies in Tai and Southeast Asian Linguistics*, 263–291. Bangkok: Ek Phim Thai Ltd.

Toba, Sueyoshi. 1989. The pronominal affix system in Khaling. In David Bradley (ed.), *Prosodic analysis and Asian linguistics: to honour R. K. Sprigg*, 201–204. Camberra: Pacific Linguistics.

Toba, Sueyoshi, Ingrid Toba & Novel Kishore Rai. 2005. Diversity and endangerment of languages in Nepal. *United Nations Educational, Scientific and Cultural Organization, Kathmandu Office KAT-CLT-2005/02* 7(7). 64–64. http://unesdoc.unesco.org/images/0013/001399/139952m.pdf.

Toffin, Gérard. 1981. L'organisation sociale des Pahari (ou Pahi), population du centre Népal. *L'Homme* 21. 39–68. https://www.persee.fr/doc/hom_0439-4216_1981_num_21_2_368184.

Tolsma, Gerard Jacobus. 1999. *A grammar of Kulung*. Leiden, The Netherlands: Leiden University Doctoral Dissertation.

Turin, Mark. 2004a. Newar-Thangmi lexical correspondences and the linguistic classification of Thangmi. *Journal of Asian and African Studies* 68. 97–120.

Turin, Mark. 2004b. Thangmi kinship terminology in comparative perspective. *Himalayan Languages: Past and Present* 101–139. http://www.digitalhimalaya.com/projectteam/turin/downloads/kinship_mouton.pdf.

Turner, Ralph Lilley. 1931. *A comparative and etymological dictionary of the Nepali Language*. London: K. Paul, Trench, Trubner.

Upreti, D. R. 1966. <Cepāng bhanne nepāli jāti [Chepang, a Nepali tthnic group]. In *Hamro Sanskriti*, 29–33.

Varya, T. V. 1971. The Chepang. In *Rising Nepal*.

Viel, Adrien. 2020. Note préliminaire sur les Chepang du Népal et leur territoire. *Arts et Cultures*.

Vinding, Michael. 1979. A preliminary report on kinship terminologies of the Bodish section of Sino-Tibetan speaking peoples. *Kailash* VII(3–4). 191–225.

Waterhouse, David M. 2004. Brian Hodgson - A biographical sketch. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 1–24. London: Routledge Curzon.

Watters, David E. 2002. *A grammar of Kham*. Cambridge University Press. https://doi.org/10.1017/CBO9780511486883.

Watters, David E. 2003. Some preliminary observations on the relationship between Kham, Magar, (and Chepang). Unpublished.

Watters, David E. 2004. *A dictionary of Kham: Taka dialect (a Tibeto-Burman language of Nepal)*. Kirtipur, Kathmandu, Nepal: Central Department of Linguistics, Tribhuvan University. https://reap.sil.org/handle/9284745/36863. (6 October, 2020).

Watters, David E. & Dan Raj Regmi. 2008. Bhujel "Direct-Inverse." *Nepalese Linguistics* 23. 429–436. http://himalaya.socanth.cam.ac.uk/collections/journals/nepling/pdf/Nep_Ling_23.pdf. (6 October, 2020).

Weidert, Alfons. 1987. *Tibeto-Burman tonology: A comparative account* (Amsterdam Studies in the Theory and History of Linguistic Science). Vol. 54. Amsterdam; Philadelphia: John Benjamins Pub. Co.

Weinberg, Miranda. 2013. Revisiting history in language policy: The case of medium of instruction in Nepal. *Working Papers in Educational Linguistics* 28(1). 61–80.

Whelpton, John. 2004. The political role of Brian Hodgson. In David M. Waterhouse (ed.), *The origins of Himalayan studies: Brian Houghton Hodgson in Nepal and Darjeeling 1820–1858*, 25–38. London: Routledge Curzon.

Wickham, Hadley. 2016. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Romain François, Lionel Henry & Kirill Müller. 2021. *dplyr: A grammar of data manipulation. R package version 1.0.7*. https://CRAN.R-project.org/package=dplyr.

Widmer, Manuel. 2014. *A descriptive grammar of Bunan*. Bern, Switzerland: Universität Bern.

Wilde, Christopher P. 2017. A phonological comparison of Gamale, Sheram and Ghusbang – three Kham varieties. *Journal of the Southeast Asian Linguistics Society* 10.1. 67–90. http://hdl.handle.net/10524/52398 University of Hawai'i Press.

Witzlack-Makarevich, Alena, Taras Zakharko, Lennart Bierkandt, Fernando Zúñiga & Balthasar Bickel. 2016. Decomposing hierarchical alignment: Co-arguments as conditions on alignment and the limits of referential hierarchies as explanations in verb agreement. *Linguistics* 54(3). 531–561. https://doi.org/10.1515/ling-2016-0011.

Wood, Hugh B. 1965. The development of education in Nepal. *Studies in Comparative Education* 5.

Yadav, Nandlal Ray. 1997. *The potential role of forest for socio-economic development of Chepang Community; a case study of Korak VDC, Chitwan District*. Kirtipur, Nepal: Tribhuvan University Master's Thesis. http://202.70.82.238/cgi-bin/koha/opac-detail.pl?biblionumber=13658.

Yadava, Yogendra Prasad. 2007. *Linguistic and ethnographic documentation of Baram*. Endangered Languages Archive.

Yonjan Tamang, Amrit. 2012. *तमङ साहित्यको इतिहास <Tamang sāhityako itihās> [History of Tamang literature]*. Kathmandu: Nepal Academy.

Yu, Kristine M. 2010. Laryngealization and features for Chinese tonal recognition. In *INTERSPEECH*, 1529–1532. Makuhari, Chiba, Japan.

Zhang, Hanzhi, Ting Ji, Mark Pagel & Ruth Mace. 2020. Dated phylogeny suggests early Neolithic origin of Sino-Tibetan languages. *Scientific Reports* 10(1). 20792. https://doi.org/10.1038/s41598-020-77404-4.

Zhang, Menghan, Shi Yan, Wuyun Pan & Li Jin. 2019. Phylogenetic evidence for Sino-Tibetan origin in northern China in the Late Neolithic. *Nature*. Springer Science and Business Media LLC 569(7754). 112–115. https://doi.org/10.1038/s41586-019-1153-z.

Zhu, Dexi. 1982. *Yǔ fǎ jiǎngyì [Lecture notes on grammar]*. Beijing: Shangwu Yinshuguan.

Zúñiga, Fernando. 2006. *Deixis and Alignment - Inverse systems in indigenous languages of the Americas.* Amsterdam: Benjamins.

Zúñiga, Fernando. 2014. Inversion, obviation, and animacy in native languages of the Americas: Elements for a cross-linguistic survey. *Anthropological Linguistics* 56(3/4). 1–22.

1894. *Xme Congrès International des Orientalistes, Session de Genève*. Rapport de la Commission de Transcription.