

Topics in the grammar of Kalkoti

David Hultman

Department of Linguistics
Bachelor's Thesis 15 ECTS credits
Linguistics – Bachelor's Course, LIN612
Spring semester 2023
Supervisor: Henrik Liljegren
Reviewer: Ljuba Veselinova



Topics in the grammar of Kalkoti

David Hultman

Abstract

This thesis presents a study of the phonological and morphosyntactic characteristics of Kalkoti, an understudied Indo-Aryan language of northern Pakistan. Kalkoti belongs to the Shina group of Indo-Aryan languages, but shows heavy influence from the Kohistani language Gawri, including the development of a complex tone system. The data used for the study were mostly collected in Pakistan in 2006 and 2015 as part of other projects, but a small amount of new data was collected as part of this study. Some notable results concern the behavior of the tone system as well as the heavily reduced nominal and verbal morphology. The ergative and genitive cases of nouns have become formally identical, and perfective verbs have by and large lost agreement, both of which set Kalkoti apart from its Shina relatives as well as from Gawri.

Keywords

Kalkoti, Shina languages, Indo-Aryan languages, language contact

Sammanfattning

Denna uppsats presenterar en studie av fonologiska och morfosyntaktiska drag i kalkoti, ett understuderat indoariskt språk från norra Pakistan. Kalkoti tillhör shina-gruppen inom de indoariska språken, men uppvisar stor påverkan från kohistanispråket gawri. Bland annat har kalkoti utvecklat ett komplext tonsystem likt gawris. Det mesta av datan som använts för studien samlades in i Pakistan år 2006 och 2015 som del av andra projekt, men en liten mängd ny data har samlats in som del av denna studie. Några nämnvärda resultat behandlar tonsystemets beteende, och den kraftigt reducerade substantiv- och verbmorfologin. Ergativ och genitiv-kasus har sammanfallit för substantiv, och de flesta perfektiva verb har förlorat kongruensböjning. Båda dessa fenomen särskiljer kalkoti såväl från närbesläktade shinaspråk som från gawri.

Nyckelord

Kalkoti, shinaspråk, indoariska språk, språkkontakt

Contents

Abbreviations	1
1 Introduction	2
2 Background	3
2.1 Location, speakers, and name	3
2.2 Relations to other languages	4
2.3 Previous studies	5
3 Method	6
3.1 Data	6
3.2 Analysis	7
4 Phonology	8
4.1 Consonants	8
4.1.1 The status of aspiration	9
4.1.2 Flapping of /d/	9
4.1.3 Utterance-final devoicing	10
4.1.4 Loan-exclusive consonants	10
4.1.5 Loss of /h/	11
4.1.6 Plosive-/r/ clusters	11
4.1.7 Nasal-plosive clusters	12
4.2 Vowels	13
4.2.1 Vowel qualities	13
4.2.2 Vowel length	13
4.2.3 Nasal vowels	14
4.3 Tone	14
4.3.1 Transcription of tone	15
4.3.2 Toneless words and high tones	15
4.3.3 Low tones	17
5 Nominals and noun phrases	21
5.1 Noun inflection	21
5.1.1 Number	22
5.1.2 Oblique case	25
5.2 Pronouns	26
5.2.1 Personal pronouns	26
5.2.2 Other pronouns	27
5.3 Noun phrase elements	27
5.3.1 Adjectives	27
5.3.2 Numerals	28
5.3.3 Other modifiers	29
5.4 Case and role marking	30
5.4.1 Ergative case	30
5.4.2 Genitive case	31
5.4.3 Dative case	32
5.4.4 Other role markers	33

5.5	Gender	34
6	Verbs	36
6.1	Finite forms	36
6.1.1	Imperfective	36
6.1.2	Perfective	38
6.1.3	Past tense marking	40
6.1.4	Irrealis	41
6.1.5	Imperative	42
6.2	Nonfinite forms	42
6.2.1	Converb	42
6.2.2	Infinitives and verbal nouns	43
6.3	Valency modification	44
6.3.1	Passive	44
6.3.2	Causative	45
7	Clauses	46
7.1	Clause types	46
7.1.1	Intransitive	46
7.1.2	Transitive	47
7.1.3	Ditransitive	48
7.1.4	Copula clauses	49
7.1.5	Quotative clauses	51
7.2	Negation and questions	51
7.2.1	Negation	52
7.2.2	Polar questions	53
7.2.3	Wh-questions	53
7.3	Multiclausal constructions	54
7.3.1	Subordination	54
7.3.2	Clause chaining	55
8	Discussion	56
8.1	Methodological limitations	56
8.2	Discussion of synchronic results	56
8.3	Historical perspective and relation to other languages	56
8.4	Future work	57
9	Conclusion	58
	References	59

Abbreviations

Languages are identified with ISO 639-3 codes shown in square brackets. The following grammatical abbreviations are used in glossing:

ACC	accusative
CAUS	causative
COP	copula
CVB	converb
DAT	dative
DIM	diminutive
ERG	ergative
F	feminine
GEN	genitive
IMP	imperative
INDF	indefinite
INF	infinitive
IPFV	imperfective
IRR	irrealis
M	masculine
NEG	negation
NOM	nominative
OBL	oblique
PASS	passive
PFV	perfective
PL	plural
PRS	present
PST	past
Q	polar question
QUOT	quotative
REFL	reflexive
SG	singular
VN	verbal noun

The following general abbreviations are also used:

C	consonant
V	vowel
HKIA	Hindu Kush Indo-Aryan
IA	Indo-Aryan
IPA	International Phonetic Alphabet
lit.	literally
JK	a speaker of Kalkoti
MI	a speaker of Kalkoti
MJ	a speaker of Kalkoti
U	a speaker of Kalkoti

1 Introduction

This thesis is a study of Kalkoti [xka], an Indo-Aryan language spoken in the village Kalkot in northern Pakistan. It is a relatively understudied language, and no detailed description of Kalkoti has ever been published. It is noteworthy as many of its recent historical developments can be attributed to influence from the neighboring Gawri language [gwc], including the development of a complex tone system. Studying Kalkoti in more depth might therefore give insights into the processes of language change and contact.

The previous work done on Kalkoti includes the publication of a word list and a sociolinguistic description by Rensch (1992), as well as two articles by Henrik Liljegren: Liljegren (2009) and Liljegren (2013). The work by Liljegren identified the language as belonging to the Shina group of IA languages, instead of the Kohistani group as was previously believed (Rensch 1992; Bashir 2003). Liljegren's work showed that the apparent Kohistani characteristics of Kalkoti are likely due to heavy influence from the neighboring Kohistani language Gawri. In addition to the historical and comparative work, Liljegren (2013) also provides a sketch of the phonology and the morphology of Kalkoti.

This study has been helped by the fact that multiple languages related to Kalkoti have been studied in much more detail. Among Shina languages, the closest genealogical relative of Kalkoti is the language Palula [phl], which has been extensively studied and described by Henrik Liljegren (e.g. Liljegren 2016). Gawri, while not as closely related in a genealogical sense, is still very important to the study of Kalkoti due to the history of language contact. Gawri has been studied by Joan Baart (e.g. Baart 1999), whose work has also been helpful in understanding the structure of Kalkoti.

This thesis mainly expands upon the previous analysis of Kalkoti made by Liljegren (2013). I have been given access to the data underlying Liljegren's earlier work, which had only been partially analyzed before. I have also gathered a small amount of new data as part of this project.

The primary aim of this thesis is to describe the major phonological and morphosyntactic characteristics of Kalkoti. It also aims to put these characteristics into a broader perspective, comparing them to Palula and Gawri. This thesis does not attempt to give a comprehensive treatment of any one aspect of Kalkoti grammar, as that is not within the scope of the thesis project, and is not possible with the limited available data.

The structure of the thesis is as follows. Chapter 2 gives a background to the study of Kalkoti. Chapter 3 describes the data and method of investigation used. Thereafter, four chapters present the results of the thesis, organized by grammatical topic: chapter 4 is concerned with phonology, chapter 5 is concerned broadly with the the components of noun phrases, chapter 6 is concerned with verbs, and chapter 7 is concerned with the structure of clauses. These four chapters present selected portions of the analyzed data to illustrate the grammatical structures identified, as well as discussion about the chosen analysis and, where relevant, the relation of the Kalkoti structures to those of other languages. Finally, chapter 8 discusses the thesis project as a whole, and chapter 9 concludes the thesis.

2 Background

This chapter presents an overview of the geographical and social setting of the Kalkoti language (section 2.1), its relation to other languages (section 2.2), and previous studies that are relevant to Kalkoti (section 2.3).

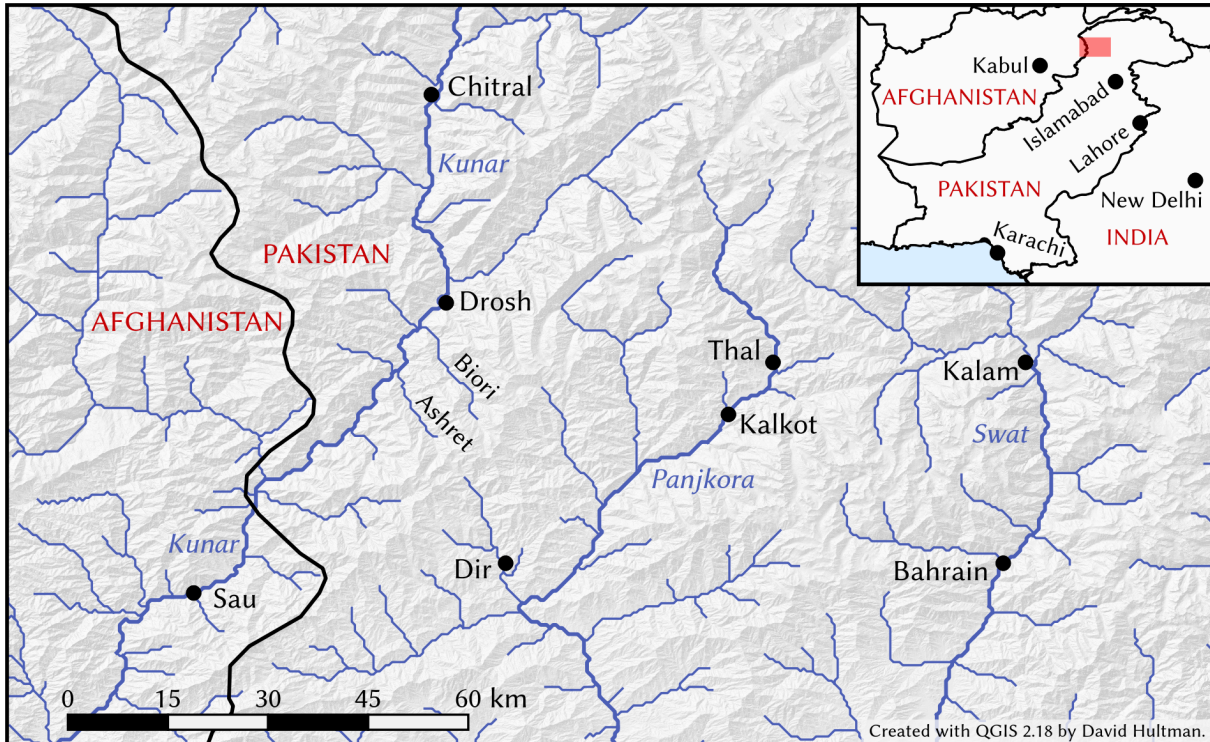


Figure 1: Location of Kalkot in northern Pakistan.

2.1 Location, speakers, and name

Figure 1 shows the location of the village Kalkot, where Kalkoti is spoken, as well as the surrounding area in the Upper Dir district of northern Pakistan. Kalkot is located on the Panjkora river in the Hindu Kush mountains at an altitude of around 1800 meters above sea level. Liljegren (2013) notes the possibility that language varieties closely related to Kalkoti may be spoken in other locations, as no fully comprehensive language survey has been carried out in the region. An overview of the other languages spoken in the region is given in section 2.2.

The most recent demographic information on Kalkoti that I know of stems from an unpublished linguistic survey performed in 2002 by Muhammad Zaman and Shamshi Khan. This survey places the population of Kalkot around 9000 people, and estimates that 70% of the inhabitants of Kalkot speak Kalkoti, whereas the other 30% speak a variety of the Gawri language. This would place the number of Kalkoti speakers around 6300 people in 2002.

Multilingualism is common among Kalkoti speakers. Rensch (1992) reports that Kalkoti men understand and use both Gawri and Pashto, and that Kalkoti women generally also know Pashto, but that only some women speak Gawri. Urdu likely also has an increasingly important role, as the national language of Pakistan. However, no information is available about the current extent of Urdu knowledge and usage in Kalkot.

The language is generally known as ‘Kalkoti’, after the name of the village. As noted by Liljegren (2013), the local pronunciation of the toponym ‘Kalkot’ begins with an aspirated con-

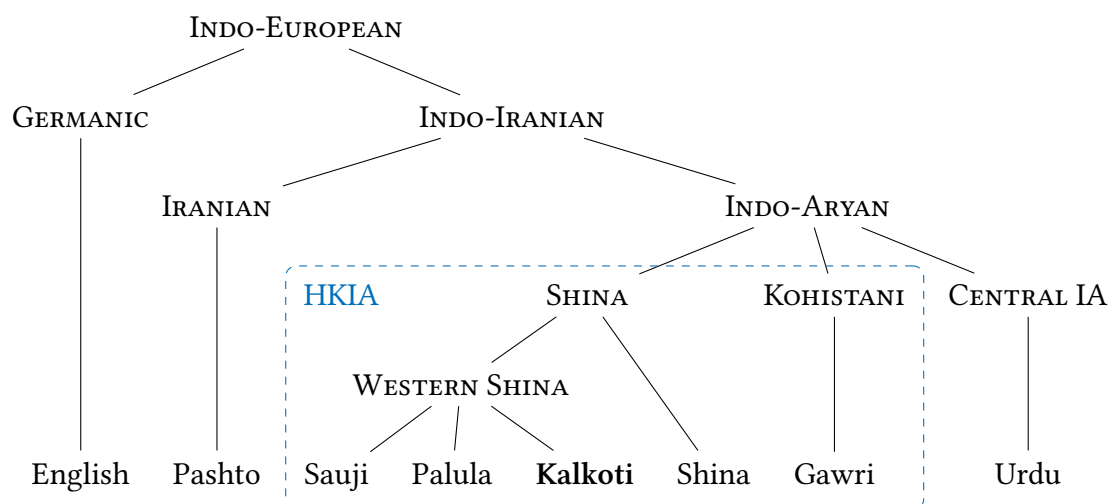


Figure 2: Partial Indo-European family tree, including the Hindu Kush Indo-Aryan areal group circled in blue. All groups have additional members not shown here.

sonant that is not reflected in the conventional spelling. The pronunciation in Kalkoti appears to be [k^hælkotː], written *khālkootː* in the transcription scheme used in this thesis. According to the survey by Zaman and Khan, Kalkoti is also known locally as ‘Goedijaa’, in contrast to the variety of Gawri that is also spoken in Kalkot, which is known as ‘Daraagi’.

2.2 Relations to other languages

Figure 2 illustrates the classification of Kalkoti and other languages relevant to this thesis within the Indo-European language family. Indo-Aryan is well-established as a constituent of the Indo-Iranian branch of the Indo-European family, but internal classification within Indo-Aryan is a complex and controversial topic. Genealogical classification typically aims to produce a tree, such as the one in figure 2, where languages are grouped based on common ancestry. Many such schemes have been proposed for Indo-Aryan, but Masica (1991, p. 446) suggests that that kind of classification may not be very useful, or even possible, for Indo-Aryan languages. The reason for this, according to Masica, is prolonged historical contact and language mixing throughout the area where Indo-Aryan languages are spoken, which has given Indo-Aryan some characteristics of a dialect continuum. Cardona and Jain (2003) disagree, and consider it to be “beyond doubt” that a classification of Indo-Aryan languages into “affiliated subgroups” is both possible and useful, though they admit that the details of the classification are debatable. No attempt is made to tackle this issue in this thesis project, so the placement of the Shina, Kohistani, and Central Indo-Aryan language groups directly below Indo-Aryan in figure 2 should not be interpreted as representing any fundamental division within Indo-Aryan: the vast majority of Indo-Aryan languages do not belong to any of these three groups, and the figure omits intermediate levels of classification.

Shina and Kohistani are among the Indo-Aryan language groups traditionally classified as ‘Dardic’ (which also includes several other languages). This term refers to the northwesternmost Indo-Aryan languages, spoken primarily in the mountains of northern Pakistan and neighboring regions of India and Afghanistan. Modern scholars do not posit a united ‘Dardic branch’ within Indo-Aryan (Masica 1991, p. 462; Bashir 2003), instead considering it to be an areal group that, owing to its relative isolation in the mountains, has on the whole undergone qualitatively different developments than other Indo-Aryan languages. To emphasize the areal

nature of the group, and to avoid unwanted connotations, the geographically descriptive term Hindu Kush Indo-Aryan is used here instead, following Liljegren (2016).

Kalkoti belongs to the Shina language group, which is named after Shina [scl], the largest language of the group. It is thought by Bashir (2003) that a westward migration of Shina speakers took place a few hundred years ago, corresponding to modern-day Shina varieties west of the main area of Shina speakers. This western group of Shina varieties includes Palula [phl] and Sauji [sdg] as reported by Bashir (2003). Liljegren (2009) presents evidence that Kalkoti, too, belongs to this group. Consequently, Palula and Sauji are the closest genealogical relatives to Kalkoti. Palula is spoken mainly in the Ashret and Biori side valleys (see figure 1) connecting to the Kunar river (Liljegren 2016) and Sauji is spoken in the village Sau further downstream on the Kunar river (Knobloch 2020).

The Gawri [gwc] language belongs to the Kohistani group of languages (Bashir 2003). It has mainly been studied by Joan Baart, who calls it ‘Kalam Kohistani’ after the village Kalam (location shown in figure 1), where most of its speakers are located (including Baart’s main language consultants). However, Gawri is spoken in a larger area, including Thal and a minority of the population of Kalkot (Baart 1999; Rensch 1992), and the name ‘Gawri’ will be used as a neutral term for this language, following Liljegren (2013). Kalkoti has undergone heavy linguistic influence by Gawri as a result of recent contact, making Gawri important to the study of Kalkoti.

Pashto, which belongs to the Iranian branch of Indo-European, is also relevant, due to its status as a lingua franca in the area where Kalkoti is spoken. This is mainly evident in a large number of loanwords in Kalkoti which can be traced to Pashto.

2.3 Previous studies

To my knowledge, the earliest published work identifying Kalkoti as a distinct language variety is Rensch (1992), which includes some demographic and sociolinguistic information about Kalkoti (and other languages), as well as a comparative word list of 210 items.

Kalkoti has been further studied in two articles by Henrik Liljegren. Liljegren (2009) is mainly concerned with the historical development of the western Shina group, but includes a brief sketch of the phonology and verb morphology of Kalkoti. Liljegren (2013) presents a more developed sketch of Kalkoti grammar, including an initial analysis of its tone system.

Kalkoti’s close genealogical relative Palula has been studied in much more detail, and Liljegren (2016) is a comprehensive grammar of Palula. Many grammatical structures are shared between Palula and Kalkoti, which has aided the analysis of the Kalkoti data in this study. Liljegren and Haider (2011) present an extensive Palula vocabulary, which has been used for lexical comparisons between Palula and Kalkoti.

A number of publications about Gawri by Joan Baart are relevant. Baart (1997) is an in-depth description of the phonology of Gawri, including a detailed analysis of its tone system. Baart (1999) deals with the morphosyntax of the language. Finally, Baart and Sagar (2004) presents annotated Gawri texts and an extensive word list.

Masica (1991) presents an overview of Indo-Aryan languages as a whole. While this book does not specifically mention Kalkoti, it is relevant to this study in that it helps place the grammatical features of Kalkoti in a broader context within the Indo-Aryan family.

3 Method

This chapter presents the data used in this thesis project, and describes the analysis performed.

3.1 Data

The data for this thesis comes from several sources. Table 1 gives a summary of each of the datasets used, ordered chronologically by collection date. The abbreviations MI, MJ, JK, and U refer to different speakers of Kalkoti. Each dataset is given a reference consisting of the consultant abbreviation, two-digit year, and a meaningless letter for disambiguation. These references are used throughout the thesis to indicate the origin of example sentences.

Table 1: Kalkoti data used.

	Consultant	Date/Loc	Analysis performed	Reference
Word list (254 elicited words)	MI	May 2006, Kalkot	Transcribed (NH)	MI06a
Sentence questionnaire (222 elicited sentences)	MJ MI	May 2006, Kalkot	Transcribed (NH, HL, DH) Glossed (HL, DH)	MJ06a MI06b
Verb survey (123 elicited sentences)	MI	May 2006, Kalkot	Transcribed (NH, HL)	MI06c
Pronoun survey (62 elicited sentences)	MJ	May 2006, Kalkot	Transcribed (NH, HL)	MJ06b
Stories (Narration of two short stories)	MJ MI	May 2006, Kalkot	Transcribed (NH, HL) Glossed (NH, HL)	MJ06c MI06d
Numerals (60 elicited numerals)	JK	Oct 2015, Islamabad	Transcribed (NL, DH)	JK15c
Valency questionnaire (87 elicited sentences)	JK	Oct 2015, Islamabad	Transcribed (NL, DH) Glossed (DH)	JK15d
North Wind and the Sun (Written translation)	JK	Oct 2015, Islamabad	Transcribed (NL, DH) Glossed (DH)	JK15e
Pear story (2 minutes free narration)	JK	Oct 2015, Islamabad	Transcribed (NL, DH) Glossed (DH)	JK15g
Online elicitation (33 elicited words/sentences)	U	Apr 2023, online	Transcribed (DH) Glossed (DH)	U23a

NH = Naseem Haider, HL = Henrik Liljegren, NL = Noa Lange, DH = David Hultman

The data can be divided into three main groups. The data from 2006 was gathered in Kalkot by Muhammad Zaman and Naseem Haider, and formed the basis of Liljegren’s earlier articles on Kalkoti. The data from 2015 was gathered as part of the project *Language contact and relatedness in the Hindukush region*. All datasets have accompanying audio recordings and English translations, and some were fully or partially transcribed or glossed by other researchers as notated in the table. I was given access to this data by Henrik Liljegren.

Lastly, some data was gathered during the final phase of this project, using WhatsApp to video chat with a native speaker of Kalkoti. I was put into contact with this speaker through my supervisor. Unfortunately, scheduling a time proved difficult and we only managed to have one initial meeting. Nevertheless, the small amount of data gathered was useful in answering a few grammatical questions that would otherwise have gone unanswered.

The Kalkoti consultants MI, MJ, JK, and U are all male, and highly multilingual. The speech of JK differs from that of the other three speakers in a few ways (see sections 4.1.1, 5.4.1, and 6.1.1). It is not known if these differences are entirely individual, or indicative of systematic variation among Kalkoti speakers.

3.2 Analysis

All the data was transcribed using the audio files (sometimes retranscribed when previous researchers had used a different transcription scheme). When relevant, the data was also glossed, based on the translations to English that accompany the Kalkoti data. This is to some degree a cyclical process: I have updated the transcription and glossing several times as I understood the language better.

As my hypotheses could not be checked by a native speaker of Kalkoti, I have only considered my glossing to be reliable under certain criteria. If a Kalkoti word seems to be used in a similar sense multiple times in the data, I have considered that to be reliable. Similarly, if a Kalkoti word appears only once in the data, but has an identifiable cognate in another language, matching the likely meaning indicated by the English translation, I have also considered it reliable. In other cases, the glossing has been considered unknown or uncertain, which is notated throughout the thesis with a question mark.

Throughout the thesis, example sentences are presented with interlinear glosses conforming to the Leipzig Glossing Rules.¹ The abbreviations used are listed at the beginning of the thesis.

The processed data was used to investigate several phonological and morphosyntactic questions, concerning phonemes and allophony, tone, noun morphology, the case system, nominal gender, verbal grammatical categories, verb morphology, syntax of noun phrases, and some aspects of the syntax of clauses.

The audio files constitute the phonetic data underlying the analysis of phonology. The segmental transcriptions were mostly done by ear. However, the computer program Praat 6.0.37 (Boersma and Weenink 2018) was used to plot the pitch contour of utterances as an aid in the transcription and analysis of tone (section 4.3).

¹<https://www.eva.mpg.de/lingua/resources/glossing-rules.php>

4 Phonology

This chapter aims to give an overview of the most prominent phonological characteristics of Kalkoti. A comprehensive study of Kalkoti phonology is not possible with the present data, so some questions will remain unanswered, particularly regarding the tone system.

4.1 Consonants

Table 2 shows the consonant phonemes of Kalkoti present in the data, in one possible analysis. There are a number of uncertainties relating to the analysis, which will be explained below. The phonemes are given in IPA transcription. Three phonemes are shown in parentheses, as their phonemic status is marginal or questionable. The notation used for Kalkoti in the rest of this thesis is also shown, in angle brackets, when it differs from the IPA.²

Table 2: Kalkoti consonant phonemes.

	Labial	Dental	Retroflex	Palatal	Velar	Uvular
Nasal	m	n	(ŋ <ṅ>)		(ŋ <ng>)	
Plosive	p	t	ʈ <ṭ>		k	q
	p ^h <ph>	t ^h <th>			k ^h <kh>	
	b	d	ɖ <ḍ>		g	
Affricate		ts		tʃ <č>		
			tʂ ^h <č̣>	tʃ ^h <č̣h>		
				dʒ <ǰ>		
Fricative	(f)	s	ʂ <ṣ>	ç <ṣ̌>	x	
		z			ɣ <ḡ>	
Flap		r <r>				
Approximant	v <w>	l		j <y>		

Comparing this to the inventory of consonant phonemes in Palula (Liljegren 2016, p. 53) and Gawri (Baart 1997, p. 19), it can be seen that Kalkoti and Gawri are alike in lacking voiced aspirated stops, which are present in Palula. Both Gawri and Kalkoti have de-aspirated voiced aspirated stops as part of a process of tonogenesis (see section 4.3). On the other hand, Kalkoti and Palula are alike in lacking the lateral fricative /ɬ/ which is found in Gawri (as well as in Sauji; Knobloch 2020, p. 12). However, Kalkoti has [ɬ] as an allophone of /l/, see section 4.1.3. Kalkoti sets itself apart from both Palula and Gawri in having lost the phoneme /h/. Kalkoti also has some apparent gaps in the aspiration contrast, though those may be due to incomplete Kalkoti data.

Most Kalkoti consonants are consistently pronounced close to the IPA symbols shown in the consonant table, but a few show allophonic variation. Some variation is context-dependent, such that multiple allophones produced by the same speaker can be ascribed to a single phoneme. Other variation is between speakers.

²This notation scheme largely follows Liljegren (2013), with *g* *ḡ* used instead of *g* *ɣ* for convenience, and without *r*, as it does not seem to contrast with *ɖ*. It is used for convenience as well as for ease of comparison to related languages, which are customarily transcribed using similar systems.

4.1.1 The status of aspiration

The contrast between plain voiceless /p t t̤ k/, aspirated /p^h t^h t̤^h k^h/, and voiced /b d d̤ g/ seems to be robust. Table 3 illustrates this, though the data set is too limited to find perfect minimal pairs for most contrasts.

Table 3: Illustration of the three-way stop contrast.

	Labial	Dental	Palatal	Velar
Unaspirated	/pɪɾɪ/ 'uncle'	/tɛ/ 'he'	/t̤ɪka:r/ 'bird'	/ka:r/ 'work'
Aspirated	/p ^h ɪt/ 'fly'	/t ^h ɛ/ 'to'	/t̤ ^h mɪl/ 'chopped'	/k ^h ɔ:l/ 'ate'
Voiced	/bɪl/ 'became'	/dɛ/ 'from'	/d̤ɪɾɪ/ 'tongue'	/ga:ŋ/ 'why'

The phoneme /t/ is almost always pronounced without aspiration, but unlike plosives at other places of articulation, there does not appear to be a contrasting aspirated retroflex plosive. The sound [t^h] appears to be entirely absent from the Kalkoti phonetic data, with one notable exception, namely the dative suffix /-t̤ɛ/. This suffix is only used for the first- and second person singular dative pronouns /mɔ̤t̤ɛ/ 'to me' and /t̤ɔ̤t̤ɛ/ 'to you', which seem to vary freely in pronunciation between unaspirated [mɔ̤t̤ɛ t̤ɔ̤t̤ɛ] and aspirated [mɔ̤t̤^hɛ t̤ɔ̤t̤^hɛ].

There also appears to be no consistent aspiration contrast in retroflex affricates, but the details are not the same as for plosives. The phones [t̤^h] and [t̤] appear to be in complementary distribution, with [t̤^h] only appearing before vowels, and [t̤] elsewhere. These will be considered allophones of the same phoneme /t̤^h/: [ɪt̤^hɪ] 'eye', [t̤^hɔ̤:r] 'walnut', [ɪt̤] 'bear', [dr̤:t̤] 'grape'. The hypothesis of an underlyingly aspirated /t̤^h/ is preferable over a hypothetical underlying /t̤/, as de-aspiration in the syllable coda is a natural rule. Indeed, aspirated sounds do not seem to be present in Kalkoti syllable codas in general. However, the phoneme /t̤^h/ is rare, and it could be the case that the apparent complementary distribution is due to a low sample size. Furthermore, some speakers seem to have replaced the [t̤^h] allophone with [s̤], pronouncing 'eye' as [ɪs̤ɪ]. This conforms to a general tendency of aspirated affricates being susceptible to be transformed into the corresponding fricative (Masica 1991, p. 102).

There exists a historical tendency for aspirated stops to become fricatives over time (Masica 1991, p. 103, citing Greek and Iranian as examples), though Masica notes that this phenomenon is rare in Indo-Aryan. However, at least some speakers of Kalkoti pronounce /t^h/ as a dental fricative [θ], while retaining other aspirated stops as stops. In the available data this is most noticeable in the speech of JK, who consistently pronounces e.g. /t^hæ:l/ 'did' as [θæ:l̥] with a clearly audible fricative.

4.1.2 Flapping of /d/

The phonetic Kalkoti data contains both voiced retroflex plosives [d̤] and retroflex flaps [ɾ], but these do not appear to contrast with each other and can be considered to be realizations of the same phoneme /d/. As illustrated in example 1, [d̤] occurs consistently in word-initial position. In word-final position there is variation between [d̤] and [ɾ]. Word-internally after a vowel (but before either a consonant or vowel), [ɾ] is used consistently.

- (1) Realizations of /d/
- a. Only [d]: [d̪ɑːleː] ‘gift’, [d̪ʊb] ‘sunken’
 - b. Variable: [bæːd, bæːɽ] ‘much, very’
 - c. Only [ɽ]: [l̪əɽk̪oːr] ‘boys’, [gəɽɪl] ‘removed’

The above presentation does not include any instances of /d/ after another consonant. The only possible examples of that in the data are in the cluster /ɲd/, which is examined separately in section 4.1.7.

4.1.3 Utterance-final devoicing

Kalkoti has a phonological process causing the sonorants /m n l r/ to be sometimes pronounced as voiceless [m̥ n̥ l̥ r̥] at the end of an utterance. Some speakers pronounce the devoiced allophone of /l/ as a fricative [ɻ]. The elicited word lists contain many examples of the contextual variation between voiced and voiceless allophones within the same word, as all words were elicited both in isolation and in the frame *rās=thä __ mānaan* ‘for this (we) say __’, as shown for /r/ in example 2a. In full sentences it is most commonly observed for /l/ as in example 2b, due to the verb-final syntax and prevalence of /l/-final verb forms (e.g. /šɑːl/ ‘put on’).

- (2) a. /ěːr/ [eːʔ̥] ‘sheep’ [MI06a-238]
/rɛs=tʰɛ̃ ɛːr mɛnɑːn/ [ræstʰæ̃ ɛːr mɛnɑː] ‘for this (we) say sheep’
- b. /mɪ bɪjǎːl tɛːp šɑːl/ [mɪ bɪjǎːl tɛːp šɑːʔ̥] ‘I put my hat on yesterday’ [MI06c-47]

Different speakers apply this process to different extents. For example, it seems to occur more often with the ending *-aál* than with *-il* (two common perfective suffixes, see section 6.1.2). The devoicing process is probably connected to the glottalization that appears on (underlyingly) rising-tone vowels in an utterance-final syllable (see section 4.3.2), but at least some speakers seem to apply it more generally.

This devoicing process is very likely due to Gawri influence, as a very similar process is present in Gawri (Baart 1997, p. 23).

4.1.4 Loan-exclusive consonants

The consonants /q ts f x z ɣ/ are only found in loanwords (see table 4). Despite this, at least all but /f/ are used consistently in the available Kalkoti data, so they must be considered to be part of the synchronic phonological system of Kalkoti. The labiodental fricative /f/ seems to vary with /p^h/: ‘photo’ occurs both as [fʊɽʊ] and [p^hʊɽʊ] in the data, so it is not quite as fully integrated into the sound system of Kalkoti as the other loan-phonemes. This pattern of [f~p^h] variation in loanwords is known to occur to some degree throughout Indo-Aryan (Masica 1991, p. 108).

Table 4: Examples of loan-exclusive phonemes.

Phoneme	Kalkoti word	Comparison	Meaning
/q/	<i>qisä</i>	Pashto قصه (<i>qisa</i>)	‘story’
	<i>muqa</i>	Pashto موقع (<i>mauqa</i>)	‘occasion’
	<i>sabaq</i>	Pashto سبق (<i>sabaq</i>)	‘lesson’
/ts/	<i>tsaaq</i>	Gawri <i>tsak</i>	‘thief’
	<i>mätsin</i>	Gawri <i>mätsin</i>	‘fish’
/f/	<i>fuṭú</i>	English <i>photo</i>	‘photo’
	<i>sirif</i>	Pashto صرف (<i>sirf</i>)	‘only, merely’
/x/	<i>xuṣaal</i>	Pashto خوشال (<i>xuṣaal</i>)	‘happy, content’
	<i>wax</i>	Pashto وخت (<i>waxt</i>)	‘time’
/z/	<i>zímidaar</i>	Gawri <i>zimidaar</i>	‘farmer’
	<i>bázaar</i>	Pashto بازار (<i>baazaar</i>)	‘market’
	<i>meez</i>	Pashto مېز (<i>mez</i>)	‘table’
/ɣ/	<i>ǰalti</i>	Pashto غلطي (<i>ǰalati</i>)	‘mistake’
	<i>narug</i>	Pashto ناروغ (<i>naarog</i>)	‘sick’

4.1.5 Loss of /h/

A notably absent phoneme in Kalkoti is /h/. It has historically been lost through regular sound change (see Kalkoti /ɪn/ and /ɛsil/ in table 5). It does not appear to have been regained in loanwords (see Kalkoti /im/ and /ɔkɔmɔt/ in table 5), though it is also plausible that these loanwords were introduced before the sound change removing /h/ took place. Only one word in the data seems to contradict this, namely the determiner /xɛr/ ‘every’: its cognates in Palula and Gawri have /h/, but this was not lost in Kalkoti. In the Kalkoti data, ‘every’ is variably pronounced as [hɛr] or [xɛr]. As there is no evidence for a contrast between [h] and [x], this word will be considered to have the phoneme /x/.

Table 5: Loss of /h/ in Kalkoti.

Kalkoti	Comparison	Meaning
<i>in</i>	Palula <i>hin</i>	‘is’
<i>àsil</i>	Palula <i>hansíl-</i>	‘laugh.PFV’
<i>im</i>	Gawri <i>him</i>	‘snow’
<i>ukumat</i>	Urdu <i>hukūmat</i>	‘government’
<i>xär ?</i>	Palula <i>har</i>	‘every’

It is likely that the loss of /h/ occurred in conjunction with the loss of voiced aspiration which gave rise to low tones (section 4.3.3). Indeed, the lost /h/ also corresponds to low tones at least some of the time.

4.1.6 Plosive-/r/ clusters

Kalkoti does not have many consonant clusters. The only clusters allowed within the syllable onset are /tr-/ and /dr-/. Table 6 shows some Kalkoti words beginning with /tr-/ or /dr-/, as well as the cognates in Palula and Gawri. As noted by Liljegren (2009), historically labial and velar plosives have assimilated to the dental place of articulation of /r/ in these clusters in Kalkoti,

but not in Palula, while Gawri has had an entirely different development. The assimilation to /tr/ and /dr/ thus seems to be a development unique to Kalkoti.

Table 6: Kalkoti words with plosive-/r/ clusters.

English	Kalkoti	Palula	Gawri
‘brother’	<i>dra</i>	<i>bhróo</i>	<i>jää</i>
‘see.PFV’	<i>driṣ</i>	<i>dhríṣṭ-</i>	<i>liç</i>
‘village’	<i>draám</i>	<i>ghróom</i>	<i>laam</i>
‘guest’	<i>traač</i>	<i>práaču</i>	<i>čač</i>
‘three’	<i>traa</i>	<i>tróo</i>	<i>ṭaa</i>
‘work’	<i>traam</i>	<i>kráam</i>	<i>ṭam</i>

Liljegren (2013) also reports that some words may be pronounced with a word-final cluster [-tʃ] when uttered carefully in isolation, by some speakers. Elsewhere, those words are pronounced with a final [-t]. No examples of this were observed in the data analyzed for this thesis.

4.1.7 Nasal-plosive clusters

The presence of nasal-plosive clusters in Kalkoti is at best marginal. However, there is some evidence that Kalkoti allows the clusters /nd nɖ ndʒ ng/ as complex codas. Liljegren (2013) finds some words to have free variation between the clusters [nd nɖ ng] and the plain nasals [n ŋ] in the syllable coda, e.g. [pɑ:nd~pɑ:n] ‘path’. In fact, Liljegren proposed that the nasals [n] and [ŋ] may not be phonemic at all, always deriving from underlying /nd/ and /ng/. In support of this view is the fact that [n] and [ŋ] never appear word-initially. The pronunciations with nasal-plosive clusters were mainly observed in word list elicitation.

When words are uttered within a sentence instead of in isolation, the picture is more consistent than the variation observed in word lists. The cluster [ng] is present, but only intervocalially. The cluster [nɖ] seems entirely absent both intervocalically and word-finally.

Table 7: Retroflex and velar nasals.

	Intervocalic	Word-final
Retroflex	<i>gäñil</i> [gəŋɻl] ‘tied’	<i>phun</i> [pʰʊŋ] ‘flower’
Velar	<i>ṭangur</i> [ṭəŋgʊɾ] ‘pear’	<i>naang</i> [nɑ:ŋ] ‘snake’

It thus seems that the only traces of [nɖ] appear when a word is articulated clearly and in isolation. This suggests that the retroflex nasal, if not quite phonemic yet, is undergoing a process of phonemicization.

The cluster /ndʒ/ is also present, but rare, appearing in the word /pɑ:ndʒ/ ‘five’ and maybe a few others. Unlike the other coda clusters, this one seems to consistently remain distinct from a plain nasal, though the affricate is usually reduced to something more like palatalization of the nasal: [pɑ:nʲ]. In the related form /pændʒɛ:ç/ ‘fifteen’, the affricate is clearly audible: [pændʒɛ:ç].

4.2 Vowels

Kalkoti has a system of 10 contrastive vowels, of which 4 are short and 6 are long, shown in table 8. Again, IPA symbols are given, as well as the spelling used by Liljegren (2013) which I have followed exactly. The IPA symbols chosen to represent the vowels are to some degree arbitrary choices among multiple allophones, especially for the short vowels, which are highly variable.

Table 8: Vowels of Kalkoti.

	Front	Back
High	i: <ii> ɪ <i>	u: <uu> ʊ <u>
Mid	e: <ee> ɛ <ä>	o: <oo> ɒ <a>
Low	æ: <ää>	ɑ: <aa>

It may be possible to have an analysis where vowel quality is primary, and length is always secondary/derived, as favored by Liljegren (2009). However, I believe this would run into a few problems. The vowel quality range of /ɛ/ appears to overlap with that of /æ:/, and similarly for the other short vowels, while duration is generally much more consistent, notwithstanding a small number of words that show variable length (see section 4.2.2). Furthermore, the long vowels show five distinct tone patterns (see section 4.3), which is explained through autosegmental tones attaching in different ways to two vocalic moras, but short vowels show a much more restricted set of tone patterns consistent with being composed of only a single mora.

4.2.1 Vowel qualities

The low short vowels /ɛ/ and /ɒ/ (written *ä* and *a*) have large phonetic variation: /ɛ/ has allophones in the range of [ɛ~æ~ɐ~ə] and /ɒ/ has allophones in the range of [a~ɑ~ɒ~ɔ]. The majority of this variation follows no pattern that I can identify. The variation is not attributable to further phonemic contrasts, as the same word often displays many of the different allophones when said by different speakers, or even by the same speaker at different times. A more thorough acoustic analysis of these vowels would be needed to learn more.

There is one apparently consistent rule, which is that short front vowels are centralized before retroflex consonants, affecting /ɪ/ and /ɛ/. The high front vowel /ɪ/ is centralized to [ɨ] before retroflex consonants, as in /driʃ/ [driʃ] ‘saw’. The low front vowel /ɛ/ vowel is pronounced as a schwa [ə] before retroflex consonants, as in /gɛɲɪl/ [gəɲɪl] ‘tied’, or /lɛɲkóɾ/ [ləɲkóɾ] ‘boys’.

4.2.2 Vowel length

The length contrast between short /ɪ ʊ ɛ ɒ/ and long /i: u: e: o: æ: ɑ:/ is consistent in most words. However, some words show variation between short and long vowels. The words /mi/ ‘I, my’ and /lɛɲkóɾ/ ‘boys’ generally have short vowels, but have been observed with a lengthened vowel before a pause, sounding more as if they were /mi:/ and /lɛɲko:ɾ/. These words will be assumed to underlyingly have short vowels. Throughout the data that I have observed, the lengthened versions of /ɪ ʊ/ in words with variable length appear to be similar in vowel quality to /i: o:/ rather than to /e:/ or /u:/. This is different from the (historical) shortening process described below, where /e:/ and /u:/ do participate.

Long Kalkoti vowels have a restricted distribution: they are almost exclusively found in the final syllable of words, or at least the final syllable of roots before affixation. Only a few words contradict this tendency: /dʒæ:le:/ ‘gift’ (a Pashto loanword) and /tri:de:/ ‘the day after tomorrow’ are the clearest examples of long vowels in a non-root-final syllable. Some examples of long vowels before a suffix are /míçɑ:l-óm/ ‘men-OBL.PL’ (see section 5.1.2) and /trɑ:r-il/ ‘sell-PFV’ (see section 6.1.2).

Some irregular or derivational suffixed forms appear to have fused into roots that obey the restriction on long vowels, with shortening of an originally long vowel as a result. Example 3 shows this with two common irregular plural forms and two diminutives. This is likely a fossilized process that is no longer productive in the language.

- (3) a. Shortening in plurals with /-ɑ:l/:
 /mě:ç/ ‘man’ → /míçɑ:l/ ‘men’
 /tre:r/ ‘woman’ → /triçɑ:l/ ‘women’
- b. Shortening in diminutives with /-tɔ:ç, -tɛ:ç/:
 /pu:/ ‘boy’ → /pɔtɔ:ç/ ‘little boy’
 /pe:/ ‘girl’ → /pitɛ:ç/ ‘little girl’

4.2.3 Nasal vowels

A sequence of a long vowel followed by /n/ is sometimes realized as a long nasal vowel. This primarily applies to the imperfective verbal suffixes /-u:n, -ɑ:n, -i:n/ (see section 6.1.1), which vary between [ũ: ã: ï:] and [u:n ɑ:n i:n]. Not all speakers use the nasal vowel variant equally often. In particular, the suffixes almost always had a consonantal [n] in the speech of JK. All speakers, including JK, use the nasal vowels before the past tense suffix /-s/, e.g. /b-u:n-s/ [bũ:s] ‘(he) was going’. No word was recorded with a [ns] cluster. In addition to those verb suffixes, there is at least one loanword which has a nasal vowel phonetically, /pæ:ns/ [pã:s] ‘money’, borrowed from Gawri *pää-s* ‘money’ (Baart and Sagar 2004, p. 280).

Given that the pronunciation with a full consonant [n] exists at least utterance-finally, and there is no contrast between [Ṽ:] and [V:n], it does not seem necessary to posit phonemic nasal vowels. Instead, they will be considered allophones of /V:n/ and written accordingly (*buuns*, *pääns*).

4.3 Tone

Kalkoti has a system of contrastive lexical tone, meaning that semantically unrelated words may differ only in the pitch or pitch contour that they are pronounced with (Yip 2002, p. 2). Tone systems are most typically known from East Asian, African, and American languages, but are also present in smaller numbers elsewhere. Multiple Indo-Aryan languages have developed tone systems, typically resulting from the loss of voiced aspirated stops (Masica 1991, p. 119). The Kalkoti tone system has arisen through a combination of the loss of voiced aspiration, a historical process of apocope, a pitch-accent system that already existed before the recent process of tonogenesis started, as well as influence from Gawri, which also has a complex tone system.

The analysis of the Kalkoti tone system presented here is an extension of the one by Liljegren (2013). Liljegren identified five different tone melodies that a Kalkoti word could have: mid-high level pitch, low level pitch, low rising pitch, high rising pitch, and high falling pitch.

Those categories are basically maintained in this analysis (though with slightly different theoretical treatment), though the distinction between low rising and high rising pitch seems to be very subtle or possibly neutralized in some environments.

Tone is generally considered to reside on an ‘autosegmental’ tier separate from vowels and consonants (Yip 2002, p. 65). This is suitable for the description of Kalkoti tone. Following Liljegren (2013), the analysis presented here takes vocalic moras to be the tone-bearing units of Kalkoti, and the analysis makes use of marked high (H) and low (L) tones that can be ‘linked’ to a vocalic mora of a word, though most moras remain unmarked for tone.

The discussion of Kalkoti tones is divided into several parts: first, section 4.3.1 discusses how tone is notated in this thesis. Section 4.3.2 discusses tone melodies which do not involve low tones. Low tones show behavior that is not entirely symmetric to that of high tones, and are discussed in section 4.3.3.

At the end of this section, on pages 19 and 20, the pitch contours of eight sentences from the JK15d data are shown in figures produced using Praat. These figures show examples of the phonetic realization of most tonal phenomena discussed in this section. They are kept separate instead of integrated with the text for formatting reasons.

4.3.1 Transcription of tone

The analysis presented here assumes the presence of underlying H and L tones. In IPA transcription high tone is marked by an acute accent /´/, and the low tone is marked by a grave accent /`/. Long vowels contain two moras to which a tone can attach, so H attaching to the first mora gives falling tone, marked as /â:/, and H attaching to the second mora gives rising tone, marked as /ã:/. In this section on tone, however, the non-IPA transcription used throughout the rest of the thesis will be used instead. This is more convenient, since long vowels are written doubled (*aa*), so the position of the H tone in a long vowel can simply be indicated by the position of the acute accent (rising *a´a*, falling *áa*).

As tone is a contrastive part of Kalkoti phonology, I have attempted to write tone throughout the thesis. However, my transcription of tone is often uncertain, for several reasons. Sentence prosody often makes it difficult to know if a word has high pitch due to lexical tone or prosodic prominence. This is especially the case with the data from 2006: I have found tones easier to transcribe in the data from 2015. This may be due to better audio quality of the recordings, or due to the consultant (JK) speaking slowly and hyperarticulating the tonal distinctions in an elicitation setting. To be more certain of the transcriptions would require finding clear contrasts, which is not always possible with the available data. In the cases I am unsure, I have defaulted to not assuming the presence of a marked tone. The marked H and L tones throughout the thesis are fairly certain, but the lack of marking should not be interpreted as a confident statement that there is no marked tone there.

4.3.2 Toneless words and high tones

In a monosyllabic word with a long vowel, there are three distinct possibilities for the presence or absence of a high tone: toneless /VV/, falling /V´V/, and rising /V´V´/. These possibilities are exemplified by the words *šii* ‘house’, *púu* ‘boy’ and *yeé* ‘mother’ in figures 3 and 4.

There does not seem to be a need to assume two underlying high tones */V´V´/ for any word. This conforms to a common pattern known as the ‘obligatory contour principle’, stating that two adjacent (underlying) tones cannot be identical (Yip 2002, p. 99). The falling pattern of /V´V/ is sometimes realized as a more-or-less flat high tone, which can be seen in *šáak* in figure 3.

This is not taken to be a distinct tone pattern, as / $\acute{V}V$ / words appear to vary allophonically in the extent to which the drop in pitch is noticeable.

High tone can also appear on short vowels and in polysyllabic words, as in *míšaálúm* ‘men (obl.)’ in figure 3, which has two high tones. Another example of high tone in a polysyllabic word is given by *minaák* ‘frog’ in figure 4. In words of the shape /CVCVV(C)/, the observed tone patterns are /C \acute{V} CVVC/ (e.g. *míšaál* ‘men’), /CVCV \acute{V} C/ (e.g. *minaák* ‘frog’), and toneless /CVCVVC/ (e.g. *gänäär* ‘old’; see figure 5). No hypothetical */CVC \acute{V} VC/ pattern is observed.

The status of high tone on short monosyllabic words is questionable. The vast majority, such as *šiṣ* ‘head’ and *tam* ‘tree’, do not appear to have a marked high tone. However, a few words such as *múr* ‘died’ and *dríg* ‘long’ do at least sometimes show a higher pitch, and participate in glottalization (see below), suggesting that they have a high tone.

The figures also show the words =*eel* ‘out of’, =*daa* ‘from’, and *tani* (reflexive possessive pronoun), here notated as toneless. All three are function words (two role marking clitics and one pronoun). Function words are often prosodically unemphasized, so it is not surprising that most Kalkoti function words appear to be toneless. The words =*daa* and *tani* in figure 4 have a slight drop in pitch, but this is more likely assimilation to the preceding high tone than an underlyingly marked high tone present in those words.

The behavior of Kalkoti high tones is similar to the pitch-accent system in Palula (Liljegren 2016, p. 72), but with some significant differences. Kalkoti allows words without a high tone, giving a three-way distinction /VV $\acute{V}V$ V \acute{V} / in long monosyllables, but Palula requires every word to have an accent. Kalkoti also allows multiple high tones in a word (e.g. *míšaálúm*), whereas Palula only allows one accented mora per word.

Table 9: Three Kalkoti tone classes compared to Palula.

Kalkoti	Palula	English
<i>daan</i>	<i>dáand</i>	‘tooth’
<i>raat</i>	<i>ráat</i>	‘blood’
<i>čoor</i>	<i>čúur</i>	‘four’
<i>bii</i>	<i>bíi</i>	‘seed’
<i>šiṣ</i>	<i>šiṣ</i>	‘head’
<i>šáak</i>	<i>šaák</i>	‘wood’
<i>púu</i>	<i>poó</i>	‘boy’
<i>róo</i>	<i>rhoó</i>	‘song, music’
<i>taár</i>	<i>tóoru</i>	‘star’
<i>deéd</i>	<i>déedi</i>	‘father’s mother’
<i>yeé</i>	<i>yéei</i>	‘mother’
<i>dríg</i>	<i>dhrígu</i>	‘long’
<i>meéš</i>	<i>míiš</i>	‘man’

Historical development As noted by Liljegren (2013), the rising tone in Kalkoti words like *taár* ‘star’ and *deéd* ‘father’s mother’ likely stems from a historical process of apocope that caused short word-final vowels to be lost in Kalkoti under some circumstances, leaving behind a high tone. This vowel was preserved in Palula, however. Table 9 shows some monosyllabic Kalkoti words along with their Palula cognates, which show some examples of rising tone (or high tone on monosyllables) arising from apocope. It also appears that toneless Kalkoti

words tend to correspond to Palula falling tones, and falling-tone Kalkoti words surprisingly correspond to Palula rising tones.

The correspondences are not without exception, however. The Kalkoti word *meěš* ‘man’ has the Palula cognate *míuš* without a final vowel, and the source of the rising tone of *meěš* (as seen in figure 9) is unknown to me. Tone assignment in Kalkoti has likely been influenced by loanwords from Gawri, which itself has a similarly complex tone system to Kalkoti, though I have not investigated the details of how Kalkoti assigns tones to loanwords from Gawri.

Glottalization Kalkoti has a process of glottalization, whereby a high tone on an utterance-final vocalic mora is instead realized as glottalization of the vowel it attaches to. This applies to rising tones on long vowels (e.g. *yaál* ‘(he) came’), as well as marked high tones on short vowels (e.g. *múr* ‘died’). The glottalization is transcribed using a small glottal stop character [ʔ], but the phonetic realization of glottalization is somewhat variable. Creaky voice throughout the affected vowel and the presence of a full glottal stop have both been observed. When affecting high tones, the glottalization process completely cancels the rising pitch that the vowel would otherwise have (see e.g. *traál* in figure 3). However, with short vowels, there is some evidence suggesting that a word like *múr* can still receive high pitch in addition to glottalization when utterance-final.

This glottalization process is connected to a process of consonant devoicing which occurs in a similar environment (see section 4.1.3). Like with the devoicing process, there is a similar process found in Gawri (Baart 1997, p. 44). The variation between high or rising pitch and glottalization is visible in word list elicitation (see example 2 on page 10). In speech, however, most sentences end with a verb, and only a few verb forms show glottalization. This is most common with the perfective suffix *-aál*, which is often very audibly glottalized. Because of the verb-final syntax of Kalkoti, the underlying rising tone of *-aál* is not usually encountered in simple sentences. However, it surfaces when followed by the polar question marker =*ää* (section 7.2.2) or when embedded in a relative clause (section 7.3.1).

(4) a. Utterance-final glottalization:

/... trǎ:l/ ‘... made’ → [trǎ:ʔɪ]
 /... mǔʁ/ ‘... died’ → [mǔʁʔɪ] or [mǔʁʔɪ]?

b. No glottalization:

/... trǎ:l=æ:/ ‘... made?’ → [trǎ:læ:]
 /... bǐl/ ‘... became’ → [bǐl] or [bǐɪ]

4.3.3 Low tones

In addition to high tones, Kalkoti also has marked low tones, which generally stem from the loss of voiced aspiration. One instance where this has become contrastive is in the verb forms *buun* ‘(he) is going’ and *bùun* ‘(he) is becoming’. In both verbs, the stem is *b-*, but the latter used to begin with **bh-* (compare Palula *bhíu* ‘to become’), which remains as low tone on the inflectional suffix *-uun*.

Figure 6 shows two examples of low tone words: *dǎǎlee* ‘gift’ and *gìin* ‘took, got, received’. Low tone appears to be a property of the syllable as a whole: there is no contrast between /*V̇V*/ and /*V̇V̇*/. In the transcriptions, the low tone will be indicated on the first mora of long vowels. It also affects short vowels, such as *dǎrin* ‘ground’ in figure 7, and *bǎt* in figure 8.

Historical development Table 10 shows some Kalkoti words where low tone has arisen through loss of aspiration, along with Palula cognates for comparison. Most examples of this are after the voiced stops /b d ɖ dz g/. However, Palula also has clusters such as /lh/ and /jh/, which correspond to Kalkoti low tones after /l/ and /j/. Similarly, independent /h/ was also lost in Kalkoti, giving rise to low tone.

Table 10: Kalkoti low tones from lost aspiration.

Kalkoti	Palula	English
<i>bùun</i>	<i>bháanu</i>	‘(he) becomes’
<i>dàrin</i>	<i>dharaáṅ</i>	‘ground’
<i>ḍàär</i>	<i>ḍheér</i>	‘belly’
<i>jägär</i>	<i>jhangaár</i>	‘liver’
<i>güin</i>	<i>ghüinu</i>	‘(he) took, got’
<i>lüun</i>	<i>lhoón</i>	‘salt’
<i>yüun</i>	<i>yháandu</i>	‘(he) comes’
<i>àsil</i>	<i>hansílu</i>	‘(he) laughed’
<i>dùúr</i>	<i>dhúura</i>	‘far away’
<i>gòór</i>	<i>ghúuru</i>	‘horse’

However, this pattern is not without exceptions. Palula also has /mh nh rh/ clusters, but the Palula words *mhaás* ‘meat’, *nhiáara* ‘near’, and *rhoó* ‘music’ do not correspond to low tones in the Kalkoti cognates *maas* ‘meat’, *niyär* ‘near’ and *róo* ‘song’. This is likely due to some kind of influence by Gawri, which also lacks low tone in its cognates to those words. Similarly, Kalkoti does have low tone in the word *ràat* ‘night’ (which forms a minimal pair with *raat* ‘blood’), which cannot be explained through a comparison to Palula *raát* ‘night’. Gawri *raat* ‘night’ does also have low tone, however.

Low-rising tones The conditions for causing low tone and rising tone can intersect, producing a low-rising tone, in e.g. the word *gòór* ‘horse’ shown in figure 9. Liljegren (2013) reports a consistent phonetic difference between high-rising tones as in *meěš* and low-rising tones as in *gòór* when produced in isolation. However, figure 9 does not demonstrate a clear difference between the two words. This may be explained as assimilation from the preceding high tone raising the initial pitch of *gòór*.

Non-contrastive low tones Low tones are not contrastive after voiceless stops. However, voiceless aspirated stops cause an allophonic lowering of the pitch of the subsequent vowel: as reported by Liljegren (2013), *khoor* ‘left’ has the same low-level pitch as *ḍàär* ‘belly’. This effect seems to be most consistent with long vowels, but some degree of lowering can be seen in the short vowel of *phun* ‘flower(s)’ in figure 10. It seems likely that the original allophonic process that lowered tone after voiced aspiration also targeted voiceless aspiration. However, since Kalkoti has maintained voiceless aspiration, this has not become contrastive. Instead, voiceless aspiration in Kalkoti has a similar synchronic status to the ‘depressor consonants’ of some Bantu languages (Yip 2002, p. 157). One might consider the consonants themselves as being specified for low tone, which can spread onto the vowel. There is also some indication that the loaned consonants /x ɣ z/ behave similarly to voiceless aspiration, and give rise to allophonic low tone.

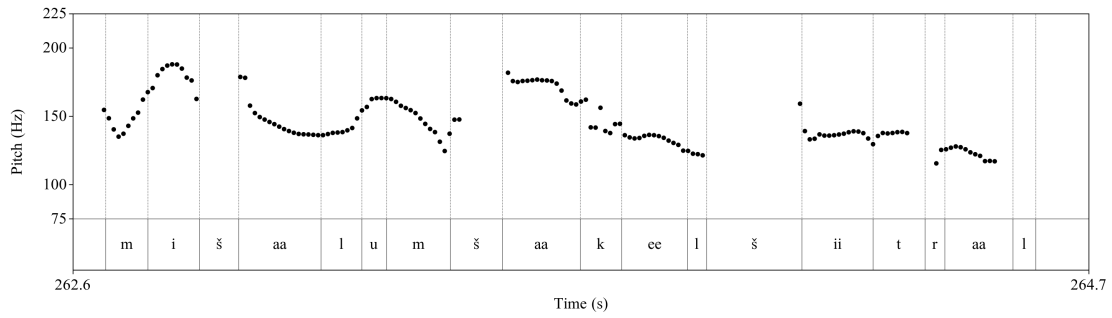


Figure 3: Pitch for JK15d-24: *míšaalum šáak=eel šii traál* ‘The men built a house out of wood.’



Figure 4: Pitch for JK15d-34: *púu minaaák tani ye=daa pálaál* ‘The boy hid the frog from his mother.’

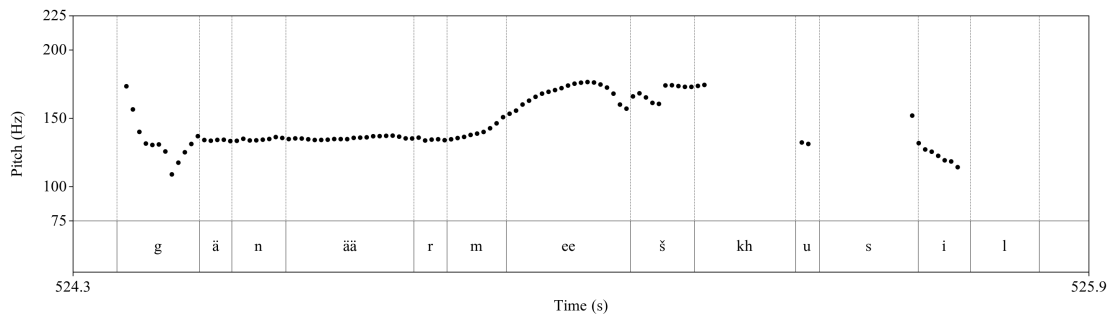


Figure 5: Pitch for JK15d-47: *gänäär meés khusil* ‘The old man coughed.’

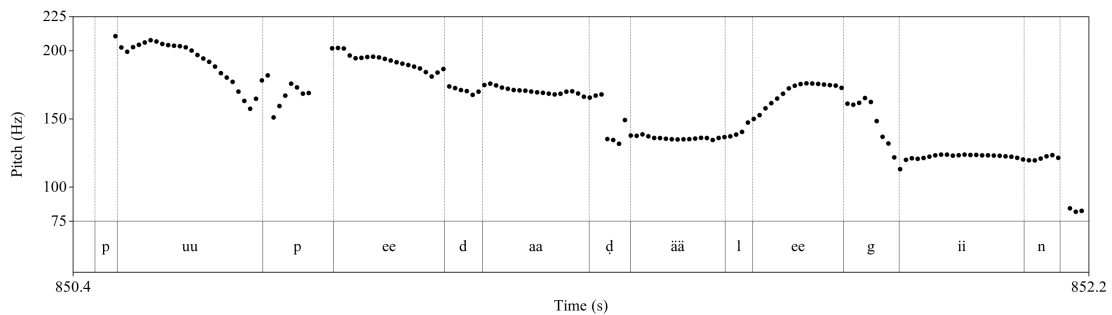


Figure 6: Pitch for JK15d-86: *púu pé=daa däälee giin* ‘The boy received a gift from the girl.’

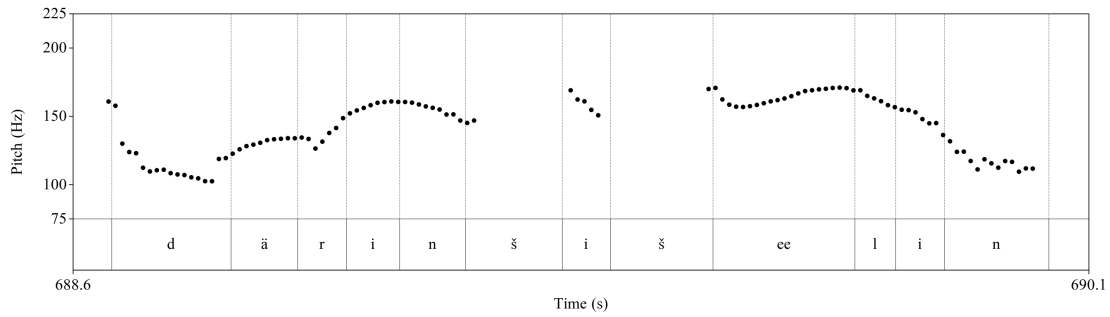


Figure 7: Pitch for JK15d-68: *därin šišeel in* 'The ground is dry.'

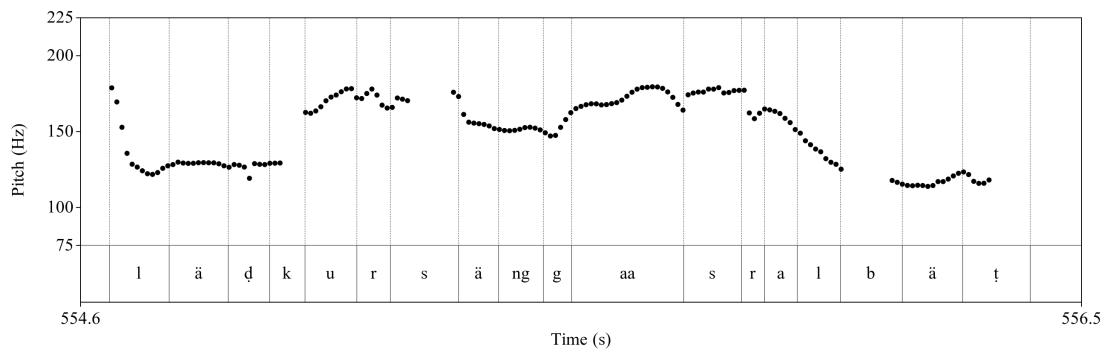


Figure 8: Pitch for JK15d-51: *lädkür sāngaás=ral bät* 'The children sat down on the bench.'

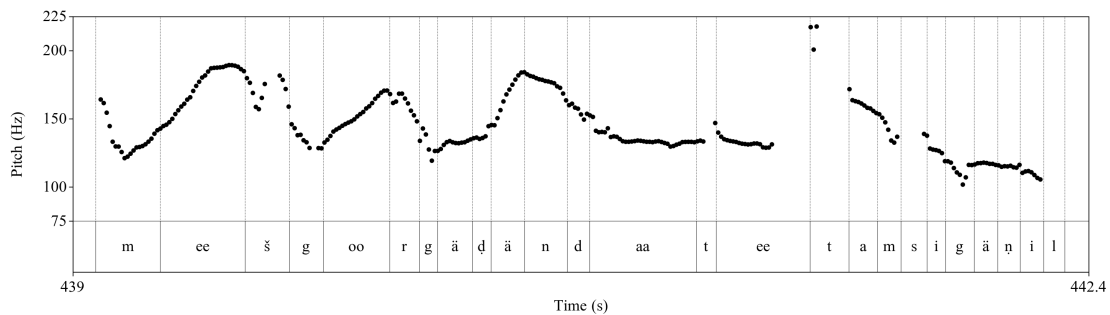


Figure 9: Pitch for JK15d-40: *meš gòor gädán=daa tee tám=si gänil* 'The man tied the horse with a rope to the tree.'

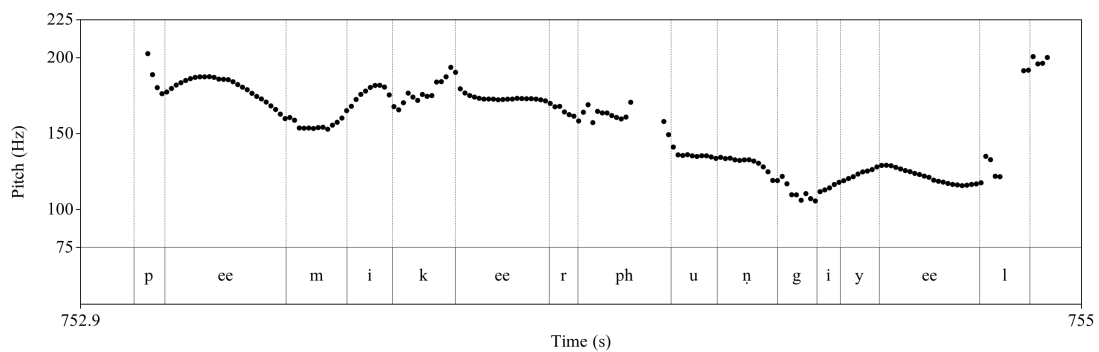


Figure 10: Pitch for JK15d-75: *pée mi keer phun gi yeel* 'The girl brought flowers to me.'

5 Nominals and noun phrases

This chapter describes the components and structure of Kalkoti noun phrases, as well as the system of case and role marking in Kalkoti.³ The term ‘nominal’ is used here as a convenient umbrella term for nouns and pronouns. The morphology of nouns is discussed in section 5.1 and pronouns are discussed in section 5.2. The structure of noun phrases, referring to prenominal modifying elements, is discussed in section 5.3. Section 5.4 describes the category of case in Kalkoti, as well as the closely related system of role marking clitics (or postpositions). Finally, section 5.5 discusses gender, which despite being a nominal feature is most consistently expressed outside the noun phrase, namely in verbal agreement.

5.1 Noun inflection

Compared to related languages, Kalkoti noun inflection is heavily reduced. The morphology is sensitive to the categories of number and case, but in a limited way. There is only one regular affix that is unambiguously part of the nominal morphology, which is the oblique plural suffix *-um*. This means that number is only consistently distinguished for oblique nouns, and the oblique case is only distinct in plural nouns. Some nouns maintain distinct singular and plural forms through stem modification or irregular suffixation, however (see section 5.1.1). In the direct and oblique cases, this amounts to three distinct forms for number-distinguishing nouns, and two forms otherwise. These possibilities are illustrated in table 11. There is also a morpheme *=ä*, which attaches to the oblique form in an entirely agglutinative manner, marking both ergative and genitive cases. It is included in table 11 for completeness and for comparison with table 14, but it is structurally equivalent to other role marking clitics discussed in section 5.4, and it is not considered part of the morphology of nouns. Table 11 is intended as a schematic overview, and usage examples illustrating these distinctions are provided in sections 5.1.1 and 5.1.2.

Table 11: Declension of *meěš* ‘man’ and *mälgir* ‘friend’.

	Singular	Plural		Singular	Plural
Direct	<i>meěš</i>	<i>míšaal</i>	Direct	<i>mälgir</i>	<i>mälgir</i>
Oblique	<i>meěš</i>	<i>míšaal-úm</i>	Oblique	<i>mälgir</i>	<i>mälgir-um</i>
Erg-gen	<i>meěš=ä</i>	<i>míšaal-úm=ä</i>	Erg-gen	<i>mälgir=ä</i>	<i>mälgir-um=ä</i>

These paradigms make a good example of the relationship between Kalkoti and Palula. Palula marks the oblique plural with a suffix ending in *m*, like in Kalkoti (Liljegren 2016, p. 45). Furthermore, the Palula oblique singular and direct plural forms are primarily marked by suffixed short vowels. Considering the historical process of apocope that has taken place in Kalkoti, it is not surprising that those forms lack overt marking in Kalkoti.

On the other hand, Gawri noun inflection looks quite different, both in the morphological means by which declension is performed, as well as the pattern of syncretism. Gawri generally inflects nouns for number and oblique case through non-concatenative morphology (changing

³The Kalkoti role markers discussed in section 5.4 are notated as clitics throughout this thesis. However, the line between a case marking clitic and a postposition is not easy to determine, or perhaps even to clearly define, so I make no definite statement concerning their exact structural status. Postpositions are not considered part of the noun phrase that they follow, but the rationale for including the discussion of these role markers in this chapter is discussed in section 5.4.

the vowels, tone, or both), and not suffixes. Furthermore, the same form is generally used for the direct plural as well as both oblique forms (Baart 1999, p. 36). That is, the typical Gawri noun has direct singular different from all other forms, while the typical Kalkoti noun has oblique plural different from all other forms.

5.1.1 Number

The category of number is marginal in Kalkoti, and not always overtly marked. As far as the available data shows, all nouns distinguish number in the oblique case. Imperfective verb agreement (see section 6.1.1) also consistently distinguishes between singular and plural masculine subjects. However, only some nouns retain distinct singular and plural forms in the direct (unmarked) case. This subset includes loanwords that have retained plural suffixes, and four very common words for humans that have irregular plurals. This discussion of number will describe the regular situation of lacking number marking in the direct case, but distinguishing it in the oblique case through the presence of *-um*. Finally, the nouns with irregular plural forms will be discussed.

Lack of plural forms in the direct case Regular nouns do not have dedicated plural forms in the direct case. This is true for most inanimate objects, e.g. *tam* ‘tree, trees’ and *čädeer* ‘basket, baskets’. It also applies to some words for humans, e.g. *traač* ‘guest, guests’, as demonstrated in example 5, where the same form can control both singular and plural agreement.

- (5) a. *äsi šii=thä lapaár traäč y-àan*
 1PL.GEN house=to tomorrow guest come-IPFV.M.PL
 ‘The guests will come to our house tomorrow’ [U23a-28]
- b. *ä traäč lapaár y-ùun*
 a guest tomorrow come-IPFV.M.SG
 ‘A guest will come tomorrow’ [U23a-31]

The Gawri cognate to *traač* is *čač*, which has a vowel change in the plural form *čäč* (Baart 1997, p. 82). This kind of alternation is common in Gawri, but Kalkoti shows no sign of it. Another method by which Gawri distinguishes number in many nouns is through tone alternations (sometimes on its own, and sometimes combined with a segmental change; Baart 1999, p. 32). Though the tone system of Kalkoti is far from thoroughly investigated, tone alternation to indicate number has also not been observed in the Kalkoti data.

Number-sensitivity of *-um* The plural oblique suffix *-um* depends on both number and case. Number is discussed here, and the distribution of oblique case is discussed in section 5.1.2. Example 6 shows three sentences where the marker *-um* is present on nouns with plural reference. For nouns that have irregular plural forms (either by stem change or suffixation; discussed further below), the *-um* suffix applies to the plural form. This is exemplified by *lädkúr-um* ‘children (obl.)’⁴ and *šägird-aan-um* ‘students (obl.)’. By contrast, example 7 shows singular nouns in the same syntactic environment (before a role marking clitic), where *-um* is

⁴As is common in Indo-European languages, the masculine plural form is used as a generic plural for groups of mixed gender.

not present. Note the contrast between plural *mälgir-um=ä* ‘friends (erg.)’ and singular *mäl-gir=ä* ‘friend (gen.)’, as well as plural *lädkúr-um=thä* ‘boys (dat.)’ and singular *púu=ä* ‘boy (erg.)’, where the *-um* suffix co-occurs with the plural stem.

- (6) a. *mi mälgir-um=ä un fläm driş*
1SG.GEN friend-OBL.PL=ERG this film see.PFV
‘My friends saw this film.’ [MJ06a-146]
- b. *is lädkúr-um=thä kitaáb-uni dit*
1PL.ERG boy.PL-OBL.PL=DAT book-PL give.PFV
‘We gave the books to the children.’ [JK15d-36]
- c. *suwaa šägird-aan-um=ä şiş şil-uun*
all student-PL-OBL.PL=GEN head hurt-IPFV.M.SG
‘All the students have headaches.’ [U23a-21]
- (7) a. *mi mälgir=ä şiş şil-uun*
1SG.GEN friend=GEN head hurt-IPFV.M.SG
‘My friend has a headache’ [U23a-20]
- b. *púu=ä pée muuk=thä nal-il*
boy=ERG girl face=DAT look-PFV
‘The boy looked at the girl.’ [JK15d-3]
- c. *dukandaár treer=thä tsádar päş-uun*
shopkeeper woman=DAT chador show-IPFV.M.SG
‘The shopkeeper is showing chadors to the woman.’ [MJ06a-20]

Irregular plural forms Some nouns (mostly, but not exclusively, denoting people) have separate forms that are used when referring to two or more entities. This concerns four very common nouns referring to humans whose plural forms have a stem change, as well as Pashto loanwords which retain a plural suffix. The common nouns with a separate plural stem are shown in table 12 and will be discussed first.

Table 12: Irregular Kalkoti plural forms formed by a stem change.

Singular	Plural	Meaning
<i>meéş</i>	<i>míşaal</i>	‘man’
<i>treer</i>	<i>triyaal</i>	‘woman’
<i>púu</i>	<i>lädkúr</i>	‘boy’
<i>pée</i>	<i>lädkír</i>	‘girl’

If number were a category with a regular overt expression in Kalkoti nouns, one might speak of the forms in table 12 as suppletive. However, as there is no regular distinction that the stem change supplants (only a lack of a distinction), this does not meet all the conventional criteria to be considered suppletion (Veselinova 2006, p. 46). This is in contrast to the suppletive verbs discussed in section 6.1.2, which show a much clearer case of suppletion.

The forms *mīšaal* and *triyaal* are seemingly related to the corresponding singulars by the addition of the suffix *-aal* (and shortening of the stem, see section 4.2.2), but this is not a productive morphological process in Kalkoti. The plurals *läḍkúr* and *läḍkír* are unrelated to the corresponding singulars. Gawri has a similar pattern for ‘man, men’, namely *miiš* ~ *mišääl* (Baart 1997, p. 105). The plural forms *läḍkúr* ‘boys’ and *läḍkír* ‘girls’ appear related to the Urdu words *larḳā* ‘boy’ and *larḳī* ‘girl’, but I do not know how those roots came to be used as plural forms in Kalkoti. As far as I am aware, this development is limited to Kalkoti.

The distinction between the forms in table 12 corresponds to the number distinction in verbs, as shown in example 8, where the singular and plural forms of ‘boy’ control singular and plural agreement. As illustrated in example 6, the plural oblique suffix *-um* is added to the plural forms of these nouns. Example 9 additionally shows singular and plural forms used with corresponding numerals.

- (8) a. ***púu nāt-uun***
 boy play-IPFV.M.SG
 ‘The boy is playing.’ [JK15d-62]
- b. ***läḍkúr šii muuk=thä trap-aan***
 boy.PL house face?=DAT run-IPFV.M.PL
 ‘The children are running home.’ [MJ06a-24]
- (9) a. ***täsi ä pee aas***
 3SG.GEN one girl COP.PST
 ‘She had one daughter.’ [MJ06a-101]
- b. ***thi čoor läḍkúr in***
 2SG.GEN four boy.PL COP.PRS
 ‘You have four children.’ [MJ06a-97]

The above examples have mostly used the words for ‘boy’ and ‘boys’, but the other words show the same distinctions (see e.g. example 11 on page 25). The corpus of Kalkoti data used for this project contains a total of over 150 tokens corresponding to the eight forms in table 12. Of these, all but two tokens seem to be consistent with the hypothesis of a number distinction. The two exceptions concern one instance of *läḍkír* seemingly being used about one girl, and one instance of *läḍkúr* being used about one boy, both in elicited sentences. The latter is given as example 56a on page 51. This may indicate that *läḍkúr* and *läḍkír* have not become fully plural-exclusive, only showing a strong tendency towards it. At the same time, those exceptions were elicited translations from Urdu (which uses the cognate words as singular nouns), so those data points may not reflect natural Kalkoti speech.

Some Pashto loanwords constitute the other group of nouns with irregular plural forms, having retained plural suffixation when loaned into Kalkoti. Three such nouns whose singular and plural forms both appear in the data are shown in table 13.

Table 13: Retained plural suffixes in Pashto loanwords.

Singular	Plural	Meaning
<i>ḍakṭar</i>	<i>ḍakṭar-aan</i>	‘doctor’
<i>ingär</i>	<i>ingär-aan</i>	‘blacksmith’
<i>kitaáb</i>	<i>kitaáb-uni</i>	‘book’

The lexical extent of this pattern is not known. It seems to commonly apply to words denoting professions such as ‘doctor’ and ‘blacksmith’ in the table. It may be present in *älugaán* ‘potatoes’, though the data does not contain the singular form for comparison. The loanword *kitaáb* ‘book’ has the plural form *kitaáb-uni*, which shows the only plural suffix other than *-aan* present in the data. The corresponding Pashto plural form is *kitaab-una*, which has a different final vowel. When these loanwords are used in the oblique case, the oblique plural suffix *-um* is added after the irregular plural suffix, as in example 6c.

5.1.2 Oblique case

The only morphological case distinction in Kalkoti nouns is between ‘direct’ and ‘oblique’ cases, and this distinction is only expressed in plural nouns. The oblique plural is marked with the suffix *-um*. This suffix does not appear to have any highly distinct allomorphs, but in at least one word (*míšaal-úm*) it carries a high tone. It was established as plural-exclusive in section 5.1.1, and this section is only concerned with the distinction between direct and oblique cases in plural nouns. The most frequent use of oblique case is before cliticized role markers (including the ergative/genitive marker =*ä*).

- (10) a. *is lädkúr-um=thä kitaáb-uni dit*
 1PL.ERG boy.PL-OBL.PL=DAT book-PL give.PFV
 ‘We gave the books to the children.’ [JK15d-36]
- b. *mi mälgir-um=ä un fläm driş*
 1SG.GEN friend-OBL.PL=ERG this film see.PFV
 ‘My friends saw this film.’ [MJ06a-146]
- c. *suwaa šägird-aan-um=ä şiş şil-uun*
 all student-PL-OBL.PL=GEN head hurt-IPFV.M.SG
 ‘All the students have headaches.’ [U23a-21]

At least one speaker (JK) generally used the *-um* suffix on its own in an ergative function, without the additional =*ä* morpheme.

- (11) a. *míšaal-úm triyaal-um=dii nal-il*
 man.PL-OBL.PL woman.PL-OBL.PL=at? look-PFV
 ‘The men searched for the women.’ [JK15d-11]

The unmarked counterpart to oblique, the ‘direct’ case, is primarily used for (non-ergative) subjects of verbs, as well as for direct objects of verbs. Example 12 shows plural nouns in the direct case in both these roles.

- (12) a. *lädkúr sāngaás=ral båt*
 boy.PL bench=onto sit.PFV
 ‘The children sat down on the bench.’ [JK15d-51]
- b. *boor mīšaal kh-uun*
 leopard man.PL eat-IPFV.M.SG
 ‘Leopards eat men.’ [lit. ‘Leopard eats men.’] [MI06b-64]

Example 12b apparently uses a singular subject to convey a general fact about leopards (the subject and object are distinguished by word order). Some more conventional sentences with plural nouns as direct objects are example 10a on page 25 and example 28c on page 34.

5.2 Pronouns

This section describes personal pronouns, as well as the various other pronouns present in the Kalkoti data.

5.2.1 Personal pronouns

Kalkoti personal pronouns distinguish person and number, as well as (at least) four cases: nominative, accusative, ergative, and genitive. They do not distinguish gender, but third-person pronouns distinguish a category which Liljegren (2013) identifies as distance (near/far). The usage parameters of the distance distinction are outside the scope of the current study, and both series will be glossed simply as third person.

Table 14: Personal pronouns in Kalkoti.

	1SG	2SG	3SG		1PL	2PL	3PL	
			near	far			near	far
Nominative	<i>ma</i>	<i>tu</i>	<i>roo</i>	<i>soo</i>	<i>bā</i>	<i>tis</i>	<i>rin</i>	<i>tin</i>
Accusative	<i>ma</i>	<i>tu</i>	<i>rās</i>	<i>tās</i>	<i>āsaan</i>	<i>tusaan</i>	<i>rāna</i>	<i>tāna</i>
Ergative	<i>mi</i>	<i>thi</i>	<i>rā</i>	<i>tā</i>	<i>is</i>	<i>tis</i>	<i>rin</i>	<i>tin</i>
Genitive	<i>mi</i>	<i>thi</i>	<i>rāsi</i>	<i>tāsi</i>	<i>āsi</i>	<i>tusi</i>	<i>rāni</i>	<i>tāni</i>

Most of these pronoun forms are attested multiple times throughout the data, but several 2PL and 3PL forms were only present in the pronouns elicitation data (MJ06b).

There are several patterns of syncretism in the pronoun forms: 1SG and 2SG pronouns do not distinguish nominative forms from accusative (*ma, tu*), or ergative forms from genitive (*mi, thi*), and 2PL and 3PL pronouns do not distinguish nominative forms from ergative (*tis, rin, tin*). Only third person singular and first person plural pronouns distinguish all four cases. The syncretism in 1SG and 2SG forms is shared with Palula (Liljegren 2016, p. 126), but the nominative-ergative syncretism in 2PL and 3PL pronouns is not.

The pronominal case system has more distinctions than that of nouns. The usage of the ergative and genitive pronoun forms is discussed in sections 5.4.1 and 5.4.2, together with the ergative and genitive marker used with nouns. The distinction between nominative and accusative pronouns does not align with any category applicable to nouns. The accusative is used before role marking clitics as well as direct objects of verbs, as shown in example 13

(overlapping with both direct and oblique noun cases). The nominative case is used otherwise, mainly for non-ergative subjects.

- (13) a. *mi täs=dä ä xat g-i yaál*
 1SG.ERG 3SG.ACC=from a letter take-CVB come.PFV.M
 ‘I brought a letter from him.’ [MJ06a-27]
- b. *mi täs driş*
 1SG.ERG 3SG.ACC see.PFV
 ‘I saw it.’ [MJ06a-182]

In addition to the four cases presented in table 14, there are also special dative forms of the first and second person pronouns: *ma=tä* ‘to me’ and *tu=tä* ‘to you’, which are formed with the clitic =*tä* attached to the accusative form. The recipients in ditransitive constructions are normally marked with the clitic =*thä* (see section 5.4.3), including third-person pronouns (e.g. *täs=thä* ‘to him’). The initial consonant of this clitic seems to have become retroflex in the environment after *ma* and *tu*.

5.2.2 Other pronouns

There is a reflexive possessive pronoun *tani* ‘one’s own’, which is used instead of genitive personal pronouns when the possessor is coreferential with the subject of the clause it is in. The same form *tani* is used regardless of person.

- (14) a. *púu minaák tani yeé=daa päl-aál*
 boy frog REFL mother=from hide-PFV
 ‘The boy hid the frog from his mother.’ [JK15d-78]
- b. *ma tani tip ş-uun*
 1SG.NOM REFL hat put_on-IPFV.M.SG
 ‘I will put on my hat.’ [MI06c-48]

There are also interrogative pronouns: *guwaá* ‘what’ and *kii* ‘who’ (see also section 7.2.3). The interrogative *guwaá* can be preceded by *xär* ‘every’ to form *xärguwaá* ‘everything’.

5.3 Noun phrase elements

The Kalkoti noun phrases present in the data minimally consist of either a noun or a pronoun. Nouns may be modified by a variety of other elements including demonstratives, genitive phrases, an indefinite article, numerals, and adjectives. All of these modifiers are optional, and precede the noun if present.

5.3.1 Adjectives

There are not many examples of adjectives in the data, but from the available data it seems likely that most adjectives are invariable, without any inflection. However, some very common adjectives have separate masculine and feminine agreement forms that employ vowel changes in the root, as demonstrated in table 15, along with the masculine and feminine Gawri forms

Table 15: Vowel alternations in Kalkoti adjectives.

Meaning	Masculine example	Feminine example	Gawri word (M, F)
‘big’	gan <i>ʃiʃ</i> ‘big head’	gän <i>ä ši</i> ‘a big house’	<i>gän, gin</i>
‘small’	lukuṭ <i>puṭooš</i> ‘little boy’	likiṭ <i>piṭeeš</i> ‘little girl’	<i>lukuṭ, likiṭ</i>
‘good’	<i>un meéš</i> raán <i>in</i> ‘this man is good’	<i>geél</i> reén <i>in</i> ‘bread is good’	<i>raan, reen</i>

for comparison. For two of these adjectives, the vowel alternations are identical to what Baart (1997) reports for Gawri, and may have been loaned, but for ‘big’, a different alternation is observed. The usage of these vowel alternations in Kalkoti may be somewhat inconsistent. The example with *reén* ‘big (f.)’ is sentence U23a-12. However, the same speaker also gave the following example where *raán* is used modifying a feminine noun:

- (15) *soo raán ä treer in*
3SG.NOM good.M? INDF woman COP.PRS

‘She is a good woman.’

[U23a-15]

5.3.2 Numerals

Kalkoti numerals from 1 to 40, as well as some higher multiples, are shown in table 16, based on the JK15c data.

Table 16: Kalkoti numerals.

1	<i>äk</i>	11	<i>äkaáš</i>	21	<i>äk-tée-bíiš</i>	31	<i>äkaáš-tée-bíiš</i>	50	<i>diš-tée-dubíiš</i>
2	<i>duu</i>	12	<i>baáš</i>	22	<i>duu-tée-bíiš</i>	32	<i>baáš-tée-bíiš</i>	60	<i>träbíiš</i>
3	<i>traa</i>	13	<i>treéš</i>	23	<i>traa-tée-bíiš</i>	33	<i>treéš-tée-bíiš</i>	70	<i>diš-tée-träbíiš</i>
4	<i>čoor</i>	14	<i>čändeéš</i>	24	<i>čoor-tée-bíiš</i>	34	<i>čändeéš-tée-bíiš</i>	80	<i>čurbíiš</i>
5	<i>paanj</i>	15	<i>pänjeéš</i>	25	<i>paanj-tée-bíiš</i>	35	<i>pänjeéš-tée-bíiš</i>	90	<i>diš-tée-čurbíiš</i>
6	<i>šu</i>	16	<i>šureéš</i>	26	<i>šu-tée-bíiš</i>	36	<i>čoor-kum-dubíiš</i>	100	<i>paanjbíiš</i>
7	<i>saat</i>	17	<i>sätaáš</i>	27	<i>saat-tée-bíiš</i>	37	<i>traa-kum-dubíiš</i>	110	<i>diš-tée-paanjbíiš</i>
8	<i>iš</i>	18	<i>ištaáš</i>	28	<i>iš-tée-bíiš</i>	38	<i>duu-kum-dubíiš</i>	120	<i>šubíiš</i>
9	<i>num</i>	19	<i>änbíiš</i>	29	<i>num-tée-bíiš</i>	39	<i>äk-kum-dubíiš</i>	200	<i>du soo</i>
10	<i>diš</i>	20	<i>bíiš</i>	30	<i>diš-tée-bíiš</i>	40	<i>dubíiš</i>	1000	<i>zir</i>

The number system is vigesimal, meaning that numbers are built around multiples of twenty: *bíiš* ‘20’, *dubíiš* ‘2×20’ = ‘40’, *träbíiš* ‘3×20’ = ‘60’, and so on. The vigesimal system is an areal feature (Bashir 2003, p. 823) shared with almost all HKIA languages, including Palula (Liljegren 2016, p. 163) and Gawri (Baart 1999, p. 58). The precise structure of the numerals is more similar to Gawri, including how a stretch of numerals before 40 ‘count backwards’: *čoor-kum-dubíiš* ‘36’ is more literally ‘4-from-40’, and so on. Some numerals, like *traa*, have a long vowel when counting, but can have a short vowel (as *trä*) when modifying a noun.

- (16) a. *un kāmra=ä sirif trä däwaal in*
 this room=GEN only three wall COP.PRS
 ‘This room has only three walls.’ [MJ06a-134]
- b. *tä du čädeer țangur pur-il*
 3SG.ERG two basket pear fill-PFV
 ‘He filled two baskets with pears.’ [JK15g]

5.3.3 Other modifiers

In addition to adjectives and numerals, there is also an indefinite article *ä*. It seems identical in form and function to the indefinite article described by Baart (1999, p. 63) for Gawri, but may be an independent development from the Kalkoti numeral *äk* ‘one’. Like in Gawri, the Kalkoti indefinite article immediately precedes the noun, coming after adjectives.

- (17) a. *ma tu=țä [raán ä kitaáb] d-am*
 1SG.NOM 2SG=DAT good INDF book give-1SG.IRR
 ‘I will give you a good book.’ [MJ06a-166]

Kalkoti has three essentially different kinds of genitive modifiers: a genitive personal pronoun, the reflexive genitive *tani*, and a noun phrase used with the genitive marker =*ä*.

- (18) a. *ma [täsi ğalti] päš-uun*
 1SG.NOM 3SG.GEN fault see-IPFV.M.SG
 ‘I see his faults.’ [MI06c-78]
- b. *púu [tani num khilunú] xuš b-il*
 boy REFL new toy happy become-PFV
 ‘The boy liked his new toy.’ [JK15d-8]
- c. *[mi mälgir=ä şiş] şil-uun*
 1SG.GEN friend=GEN head hurt-IPFV.M.SG
 ‘My friend has a headache’ [U23a-20]

There are also a few demonstratives present in the data: *un* and *atu*, which are glossed as ‘this’, as well as *aar/äär/ääđ*, which is glossed as ‘that’. The latter has a pronunciation that varies between speakers. The semantics of these demonstratives is not possible to investigate within the scope of this thesis.

Finally, example 19 shows a construction involving the word *mäni* ‘called’, seemingly a form of the verb *män-* ‘say’.

- (19) a. *äsi draám=mi [şilkin mäni ä baağ] in*
 1PL.GEN village=in Shelkin called INDF place COP.PRS
 ‘In our village, there is a place called Shelkin.’ [MJ06c-7]

5.4 Case and role marking

In Indo-Aryan in general, morphological case markers and nonmorphological role markers like postpositions behave in a similar way: both appear after a noun and serve to mark the relationship that the noun phrase headed by that noun has to the rest of the sentence. Masica (1991, p. 231) proposes a hierarchy of three ‘layers’ to the post-nominal case-like elements. Masica’s Layer I consists of ‘inherited synthetic’ forms, often characterized by allomorphy and number-sensitive case forms. Layer II consists of ‘new agglutinative’ forms, which are often constructed by adding onto a Layer I element. Layer III consists of ‘quasi-analytic elements’.

By Masica’s classification, only the plural oblique *-um* can be said to belong to Layer I in Kalkoti. All other Kalkoti role markers are number-invariant, and for plural nouns are added after *-um*. No phonological processes apply, not even to avoid vowel hiatus: the genitive form of *kāmrá* ‘room’ is *kāmrá=ä*, pronounced [kəmrá.æ], with both vowels clearly audible. The only allomorphy of any Layer II role marker concerns an irregular form of the dative clitic that is only used with two specific pronouns (see section 5.4.3).

I have chosen to consistently write these Layer II morphemes as clitics, as they are bound morphemes that do not appear anywhere except after nouns or pronouns. Butt and King (2004) have argued in more detail that similar Layer II markers in Urdu and Hindi are clitics, based on tests including coordination and the presence of a focus clitic that can intervene between an oblique noun and the case marker. The available Kalkoti data does not allow performing the same tests, but in the absence of further evidence, writing the case markers as clitics seems like the most neutral choice.

As is conventional, I will still use the term ‘case’ for the ‘grammaticalized’ categories of ergative and genitive (which also have specialized forms for pronouns, without a separate role marker), as well as dative, even though they are not morphological categories of nouns.

5.4.1 Ergative case

As is the norm among Indo-Aryan languages, Kalkoti has an ergative case used for the subjects of transitive perfective verbs. For nouns, the ergative case is marked using the clitic *=ä*. The personal pronouns have separate ergative forms, and do not use the *=ä* clitic.

- (20) a. *treer=ä púu taa=waa päläṭ-il*
 woman=ERG boy blanket=into cover-PFV
 ‘The woman covered the boy with a blanket.’ [JK15d-43]
- b. *mi mälgir-um=ä un fläm driṣ*
 1SG.GEN friend-OBL.PL=ERG this film see.PFV
 ‘My friends saw this film.’ [MJ06a-146]
- c. *is çiiir báazaar=mi traar-il*
 1PL.ERG milk market=in sell-PFV
 ‘We sold milk in the market.’ [MJ06b-24]

In the data available to Liljegren (2013) the ergative marker appeared to be obligatory. However, in the speech of one of the Kalkoti consultants (JK), the conditions governing the appearance of *=ä* are unclear. Some sentences such as 21a seem like they should have the ergative marker, but they do not, while other sentences like 21b have *=ä* as expected. This cannot simply be governed by inflection classes of the subject noun: *púu* ‘boy’ can appear as the subject

of a transitive perfective clause both with and without the ergative marker. Sentences like 21c suggest that JK uses the oblique form on its own to mark ergative case, when he does not use the clitic =*ä*.

- (21) a. *púu máwa kh-aál*
 boy fruit eat-PFV
 ‘The boy ate the fruit.’ [JK15d-1]
- b. *púu=ä péé muuk=thä nal-il*
 boy=ERG girl face=DAT look-PFV
 ‘The boy looked at the girl.’ [JK15d-3]
- c. *triyaal-um maas päč-aál*
 woman.PL-OBL.PL meat cook-PFV
 ‘The women cooked the meat.’ [JK15d-79]

The characterization of the Indo-Aryan ergative case as marking transitive perfective subjects is known to not be entirely accurate for several IA languages (Butt and Ahmed 2011). Instead, the ergative may be more accurately defined using semantic or pragmatic notions such as control over one’s actions. The apparent optionality of =*ä* points to the possibility of a similar condition governing Kalkoti ergativity. The available data, consisting mostly of elicited translations into Kalkoti, are not suited to investigating the semantics of the Kalkoti ergative, but it would be a suitable topic for future study.

5.4.2 Genitive case

Most pronouns have a distinct genitive form. However, 1SG and 2SG pronouns merge the genitive and the ergative case forms into one. The same appears to be true for all nouns: genitive nouns are marked with the clitic =*ä*, like ergative nouns. The two main functions of the genitive case appear to be as a prenominal genitive modifier (examples 22a and 22b; see also section 5.3.3) and as the subject in a predicative possession construction (examples 22c and 22d; see also section 7.1.4).

- (22) a. *äsi kučur tipä tipä múr*
 1PL.GEN dog now now die.PFV
 ‘Our dog has just died.’ [MJ06a-184]
- b. *tä zaar=ä gilaas píil*
 3SG.ERG poison=GEN glass drink.PFV
 ‘She drank the glass of poison.’ [MI06b-144]
- c. *äsi bääd gän ä šii in*
 1PL.GEN very big.F INDF house COP.PRS
 ‘We have a very big house.’ [MJ06a-139]
- d. *un šii=ä paanĵ dār in*
 this house=GEN five door COP.PRS
 ‘This house has five doors.’ [MJ06a-143]

Having a single form with both ergative and genitive function is a pattern known from several ergative languages throughout the world (Dixon 1994, p. 57). However, it is not common in Indo-Aryan, where the ergative forms are more often identical to instrumental forms (Masica 1991, p. 238).

5.4.3 Dative case

The Kalkoti dative case is marked by the clitic (or perhaps postposition) =*thä* after nouns and at least some pronouns. The primary function of the dative is to mark recipients of ditransitive verbs (see section 7.1.3).

- (23) a. *is lädkúr-um=thä kitaáb-uni dit*
 1PL.ERG boy.PL-OBL.PL=DAT book-PL give.PFV
 ‘We gave the books to the children.’ [JK15d-36]
- b. *pée púu=thä tani kučur sipat thää́l*
 girl boy=DAT REFL dog talk? do.PFV
 ‘The girl talked to the boy about her dog.’ [JK15d-18]
- c. *thi räs=thä kiteek rupee dit*
 2SG.ERG 3SG.ACC=DAT how_much money give.PFV
 ‘How much money did you give him?’ [MJ06a-60]

After the pronouns *ma* ‘I, me’ and *tu* ‘you (sg.)’, the dative marker is pronounced with a retroflex consonant (that may or may not be aspirated, see section 4.1.1).

- (24) a. *rä ma=ṭä angrizii päš-aál*
 3SG.ERG 1SG.ACC=DAT English show-PFV
 ‘He taught me English.’ [MJ06a-167]
- b. *zamaán=ä tu=ṭä guwaá män-il*
 Zaman=ERG 2SG.ACC=DAT what say-PFV
 ‘What did Zaman say to you?’ [MJ06a-68]

The dative case can also be used to encode the experiencer in certain so-called ‘dative subject’ constructions (see sections 7.1.1 and 7.1.4).

- (25) a. *puṭooš=thä šidäl y-ùun*
 boy=DAT cold come-IPFV.M.SG
 ‘The child is feeling cold.’ [MJ06a-70]
- b. *tu=ṭä un kitaáb män-il in*
 2SG.ACC=DAT this book read-VN COP.PRS
 ‘You will have to read this book.’ [MJ06a-89]

5.4.4 Other role markers

Apart from the ergative/genitive and dative case markers, there are many other role markers that appear to occupy the same syntactic position, most of which denote spatial relationships. Example 26 shows some: =*ij* marks spatial destinations, =*räl* marks motion onto something, =*mi* marks location on or inside, and =*dii* marks location at or next to something. The last of these, =*dii*, is more commonly used in a grammaticalized function marking the possessor in an alienable predicative possession construction (see section 7.1.4).

- (26) a. *ma báazaar=iĵ goo*
 1SG.NOM market=to go.PFV.M.SG
 ‘I went to the market.’ [MJ06a-23]
- b. *seekäl=räl nām-il*
 bicycle=onto put-PFV
 ‘(... and) placed (it) onto the bicycle.’ [JK15g]
- c. *púu kheér=mi nikhät*
 boy field=on appear.PFV
 ‘The boy appeared on the field.’ [JK15d-81]
- d. *čhätree dār=dii in*
 umbrella door=at COP.PRS
 ‘The umbrella is next to the door.’ [MJ06a-43]

Another of these markers, =*dä* (sometimes lengthened as =*daa*), deserves a special mention for its large polysemy. The basic meaning may be something like ‘from’, which is how it is glossed. It marks spatial sources (example 27a), sources of possession (example 13a on page 27), instruments (example 27b), means of travel (example 56c on page 51), recipients of questions (example 27c), standards of comparison (example 27d), and temporal sequence (example 35c on page 41).

- (27) a. *ma tä báazaar=dä patä waal aas*
 1SG.ACC DAT market=from back? come_down.INF COP.PST
 ‘I had to come back from the market.’ [MJ06a-192]
- b. *púu lumeeĉ=dä naang már-il*
 boy stick=from snake kill-PFV
 ‘The boy killed the snake with a stick.’ [JK15d-28]
- c. *ustad=ä pée=dä suwaal täpoos thääł*
 teacher=ERG girl=from question ask do.PFV
 ‘The teacher asked the girl a question.’ [MJ06a-56]
- d. *siir räs=dä tikir in*
 sun 3SG.ACC=from strong COP.PRS
 ‘(... that) the sun is stronger than him.’ [JK15e]

5.5 Gender

The term ‘grammatical gender’ refers to a classification of nouns into groups based on agreement patterns (Corbett 1991, p. 105), though the details of the definition are not universally agreed upon. Kalkoti has a category of grammatical gender: nouns are either masculine or feminine. This is primarily reflected in the inflection of imperfective verbs, which take a suffix whose form is decided by the gender and number of the subject. The gender distinction is also relevant to two perfective verbs (see section 6.1.2) and to some extent also in adjectives (see section 5.3.1). Example 28 shows verb agreement with masculine singular, masculine plural, and feminine human nouns.

- (28) a. *púu nāt-uun*
 boy(M) play-IPFV.M.SG
 ‘The boy is playing.’ [JK15d-62]
- b. *gānaraaṭ draám=mi biš-aan*
 old.people(M) village=in live-IPFV.M.PL
 ‘The old people live in town.’ [JK15d-56]
- c. *treer ālug-aán gād-een*
 woman(F) potato-PL? dig.up-IPFV.F
 ‘The woman is digging for potatoes.’ [JK15d-73]

In the terminology of Corbett (1991), Kalkoti has three ‘agreement classes’ (M.SG, M.PL, and F), corresponding to two genders (masculine and feminine). Only the masculine gender has a further number distinction in agreement, as the *-een* verb suffix (or the variant *-iin* used by other speakers) is used both with singular and plural feminine subjects (see section 6.1.1).

The presence of masculine and feminine genders is unsurprising. Liljegen (2019) investigated grammatical gender in 25 HKIA languages, Kalkoti among them. He found that a masculine-feminine division without further distinctions is by far the most common configuration, appearing in 18 out of 25 languages, including Palula and Gawri. Like other related languages, Kalkoti appears to generally associate masculine gender with the back vowels *u*, *uu*, *oo*, *aa* and feminine gender with the front vowels *i*, *ii*, *ee*. This can be seen in lexical items like *puu* ‘boy’ and *pee* ‘girl’, in adjective agreement (see section 5.3.1), as well as in verb agreement (see sections 6.1.1 and 6.1.2).

Kalkoti gender assignment is semantically predictable on the basis of biological sex for nouns denoting humans, but does not seem to be predictable for inanimate objects. In the Kalkoti data, the gender of inanimate nouns is rarely visible, as inanimate objects are rarely the subject of a verb, which is the position that most often controls gender agreement. Example 29 shows some of the sentences in which inanimate nouns do control gender agreement.

- (29) a. *mi daan šil-uun*
 1SG.GEN tooth(M) hurt-IPFV.M.SG
 ‘I have a toothache.’ [MJ06a-74]
- b. *phānoos dūkur-uun*
 ball(M) roll-IPFV.M.SG
 ‘The ball is rolling.’ [JK15d-65]

- c. *šii* *ǰäl-**een***
house(F) burn-IPFV.F

‘The house is burning.’

[JK15d-67]

Table 17 shows the non-human nouns whose gender has been possible to determine through Kalkoti sentences such as the above, as well as a comparison with a likely cognate or loanword source, whose gender in that language is notated in parentheses. As far as this limited data shows, the masculine-feminine distinction has been consistently retained in both inherited and loaned Kalkoti vocabulary.

Table 17: Gender of non-human nouns in Kalkoti.

	Kalkoti word	Comparison
Masculine	<i>daan</i> ‘tooth’	Palula <i>dáand</i> ‘tooth’ (m.)
	<i>šiš</i> ‘head’	Palula <i>šiš</i> ‘head’ (m.)
	<i>wää</i> ‘water’	Palula <i>wii</i> ‘water’ (m.)
	<i>phänoos</i> ‘ball’	Pashto پڼډوس (<i>pəndos</i>) ‘ball’ (m.)
	<i>boor</i> ‘leopard’	Gawri <i>boor</i> ‘leopard, lion’ (m.)
Feminine	<i>paan</i> ‘path’	Palula <i>páand</i> ‘path’ (f.)
	<i>niin</i> ‘sleep’	Palula <i>níindra</i> ‘sleep’ (f.)
	<i>šii</i> ‘house’	Gawri <i>šit</i> ‘house’ (f.)
	<i>bäkhin</i> ‘arm’	Gawri <i>bookhin</i> ‘elbow’ (f.)
	<i>päšin</i> ‘bird’	Gawri <i>päçhin</i> ‘bird’ (f.)
	<i>geél</i> ‘bread’	Gawri <i>geel</i> ‘bread’ (f.)

The above discussion has focused on nouns and adjectives. Gender is not directly relevant to Kalkoti pronouns, as they do not have an inherent gender. However, the same form, e.g. *soo* ‘he, she’, can control either masculine or feminine agreement based on the gender of its referent.

- (30) a. *soo* *ǰ-**iun***
3SG.NOM come-IPFV.M.SG

‘He is coming.’

[JK15d-8]

- b. *soo* *ǰ-**iin***
3SG.NOM come-IPFV.F

‘She is coming.’

[JK15d-9]

6 Verbs

This chapter is concerned with the morphology and usage of Kalkoti verbs. The discussion will be divided into three parts. Section 6.1 concerns finite verb forms, section 6.2 concerns nonfinite forms, and section 6.3 describes the passive and the causative, which are not like other inflection, in that they derive new verb stems to which regular inflection attaches.

It is worth mentioning here the existence of a copula with two forms (*in* ‘is’ and *aas* ‘was’). However, the copula is morphologically and syntactically distinct from other verbs, and it is instead described in section 7.1.4.

6.1 Finite forms

The finite declarative forms of Kalkoti verbs are split into three primary TAM values: imperfective (section 6.1.1), perfective (section 6.1.2), and irrealis (section 6.1.4). The first two are plentiful in the data and are better understood than irrealis, which only occurs a few times in the data. There is also an imperative form (section 6.1.5).

The three declarative categories differ when it comes to agreement. Perfective verbs generally do not show any agreement, though some marginal agreement exists with two basic motion verbs. Imperfective verbs show agreement for the gender and number of the subject, but not person. Irrealis verbs, however, show agreement for the person and number of the subject, but seemingly not gender. Imperatives seem to lack agreement entirely. Examples for each of these agreement patterns are given in the subsequent sections.

In addition to the three basic categories, there is an ‘outer’ layer of past-tense marking that can be indicated with the suffix *-s* attached after the perfective or imperfective suffixes (which is elaborated on in section 6.1.3).

The suffixes corresponding to the declarative forms are summarized in table 18, with two caveats. Firstly, the perfective form is often irregular and *-il* (*-is*) is simply the most common marker. Secondly, the lack of 1PL and 2PL irrealis forms in the table is not to be taken as a sign that they do not exist in Kalkoti, just that they do not exist in the data. Usage examples of all of these forms are provided in the subsequent sections.

Table 18: Finite declarative verb forms.

	Perfective	Imperfective			Irrealis			
		M.SG	M.PL	F	1SG	2SG	3SG	3PL
Plain	<i>-il</i> (+ others)	<i>-uun</i>	<i>-aan</i>	<i>-iin</i>	<i>-am</i>	<i>-ä</i>	<i>-ä</i>	<i>-ään</i>
+ Past	<i>-is</i> (+ others)	<i>-uuns</i>	<i>-aans</i>	<i>-iins</i>				

The semantics of the TAM categories is not a primary focus of this study. However, it should be noted that the decision to consider the core inflectional distinction between perfective and imperfective to be aspectual rather than having to do with tense rests mainly upon the existence of the past imperfective form (see section 6.1.3), which shows that tense is to some degree an independent category.

6.1.1 Imperfective

The formation of imperfective verbs appears to be completely regular, with one complication that is described at the end of this section. The imperfective stem is followed by a suffix

determined by the gender and number of the subject: *-uun* for masculine singular, *-aan* for masculine plural, and *-iin* or *-een* for feminine, either singular or plural. There is no contrast between *-iin* and *-een*, just a difference between speakers: JK always uses *-een*, while the data from other speakers only has *-iin*.

Usage examples of these suffixes are given throughout the thesis, in particular in sections 5.1.1 and 5.5. Person is not a factor in imperfective agreement: first-person and second-person subjects show agreement based only on their number as well as the gender of their referent (such as in examples 31a and 31b). The imperfective verb form is used in most present-tense contexts, and it can also be used in future contexts (such as in example 31a). It can also be combined with past tense marking (see section 6.1.3) to produce a past imperfective form. Examples 31c and 31d illustrate the lack of a number contrast with feminine agreement.

- (31) a. *lupaarä ma trap-uun*
 tomorrow 1SG.NOM run-IPFV.M.SG
 ‘Tomorrow I (male) will run.’ [MI06c-24]
- b. *bä näṭ-aan*
 1PL.NOM play-IPFV.M.PL
 ‘We (male or mixed group) are playing.’ [MJ06a-16]
- c. *soo ä šää=ral biš-iin*
 3PL.NOM INDF charpoi=on sit-IPFV.F
 ‘She is sitting on a charpoi.’ [MJ06b-4]
- d. *tin šää=ral biš-iin*
 3PL.NOM charpoi=on sit-IPFV.F
 ‘They (all female) are sitting on the charpoi.’ [MJ06b-8]

In order to support the morphological analysis, table 19 shows the verbs that appear in the data in all three imperfective forms.⁵

Table 19: Imperfective verb forms.

Masculine singular	Masculine plural	Feminine	English
<i>buun</i> ‘(he) goes’	<i>baan</i> ‘(they, m.) go’	<i>biin</i> ‘(she) goes’	‘go’
<i>duun</i>	<i>daan</i>	<i>diin</i>	‘give’
<i>bišuun</i>	<i>bišaan</i>	<i>bišiin</i>	‘sit’
<i>čhinuun</i>	<i>čhinaan</i>	<i>čhiniin</i>	‘cut’
<i>khusuun</i>	<i>khusaan</i>	<i>khusiin</i>	‘cough’
<i>bùun</i>	<i>bàan</i>	<i>bìin</i>	‘become’
<i>yùun</i>	<i>yàan</i>	<i>yìin</i>	‘come’

As each of the three imperfective form categories always ends with the same segmental material, that ending is considered a suffix. This results in the somewhat unusual situation

⁵It should be noted explicitly that the forms in table 19 are not the only data underlying the analysis of imperfective verbs. Many other verbs also appear with one or two different imperfective suffixes, and all appear to be entirely regular, agreeing with the pattern described here.

that the root morphemes of some verbs consists of just a single consonant: *buun* ‘(he) goes’ is morphologically *b-uun*, since the *-uun* ending is an entirely regular grammatical marker that cannot contribute any lexical information. This same one-consonant stem *b-* ‘go’ can also be observed in the formation of converbs (see section 6.2.1).

There is one complication, which is that some one-consonant verb stems impart a low tone on the suffix: e.g. *bùun* ‘(he) becomes’ forms a minimal pair with *buun* ‘(he) goes’. While tone is notationally marked on the vowel, it is not bound to a single segment, rather applying to moras or syllables, so the form *bùun* may be considered to be composed of a root *b-* carrying with it a floating low tone, followed by the regular suffix *-uun*.

6.1.2 Perfective

The perfective verb form seems to be the most varied in its formation, in direct contrast to the regular imperfective form. Two main regular verb types will be recognized, as well as several irregular verbs, all of which have perfective forms that are idiosyncratic to varying degrees.

Regular verbs Many verbs form the perfective by adding *-il* or *-aál* to the verb stem. The *-aál* suffix is glottalized at the end of an utterance, which is taken to be a reflex of rising tone (see section 4.3 for discussion). Only the perfective forms constructed with these suffixes will be considered to be morphologically compositional. Table 20 shows some regular perfective forms observed in the data, with imperfective forms for comparison.⁶

Table 20: Imperfective and perfective forms of regular verbs.

IPFV (M.SG)	PFV	English
<i>b-ùun</i>	<i>b-il</i>	‘become’
<i>täl-uun</i>	<i>täl-il</i>	‘pour’
<i>khus-uun</i>	<i>khus-il</i>	‘cough’
<i>traar-uun</i>	<i>traar-il</i>	‘sell’
<i>ämuş-uun</i>	<i>ämuş-il</i>	‘forget’
<i>ş-uun</i>	<i>ş-aál</i>	‘put on (clothes)’
<i>kh-uun</i>	<i>kh-aál</i>	‘eat’
<i>duw-uun</i>	<i>duw-aál</i>	‘wash, clean’
<i>päl-uun</i>	<i>päl-aál</i>	‘hide, steal’

There does not appear to be a consistent semantic or phonological division between the verbs taking the *-il* suffix and those with *-aál*. However, there appears to be a partial distinction involving transitivity. All regular intransitive verbs seem to use the short suffix *-il*. In the available data this can be seen in the forms *khusil* ‘coughed’, *àsil* ‘laughed’, *čikaril* ‘cried’, *ringil* ‘wept’, *träpil* ‘ran’, and *ičil* ‘gave up, lost?’. Transitive verbs, on the other hand, do not seem to predictably fall into one group or the other, as can be seen from the examples in the table.

Irregular verbs Outside the two regular verb classes, there exists a variety of strategies for perfective formation, illustrated in table 21. Some patterns or tendencies can be identified,

⁶The imperfective forms are shown with the masculine singular suffix *-uun*, as that is the only imperfective form that is directly attested in the data for all these verbs.

but not enough to consider any of these perfective forms morphologically compositional. Two irregular verbs use *l*-endings, but not in the regular way identified above: *thää́l* ‘did’ has a front vowel, and *yaál*, *yeél* ‘came’ shows gender agreement. Another group of verbs have perfective forms that end in *t*, though the suffix vowel varies unpredictably. Some other verbs employ stem modification: *píil* ‘drank’ has a lengthened vowel, and *múr* ‘died’ has a vowel quality change. Finally, two verbs are entirely suppletive, one of which also shows gender agreement.

Table 21: Irregular perfective forms.

IPFV (M.SG)	PFV	English
<i>th-uun</i>	<i>thää́l</i>	‘do’
<i>y-ùun</i>	<i>yaál</i> (M) <i>yeél</i> (F)	‘come’
<i>d-uun</i>	<i>dit</i>	‘give’
<i>w-uun</i>	<i>waat</i>	‘come down’
<i>nikh-uun</i>	<i>nikhät</i>	‘come out’
<i>pil-uun</i>	<i>píil</i>	‘drink’
<i>mär-uun</i>	<i>múr</i>	‘die’
<i>bış-uun</i>	<i>bät</i>	‘sit’
<i>päs-uun</i>	<i>driş</i>	‘see’
<i>b-uun</i>	<i>goo</i> (M.SG) <i>gee</i> (M.PL, F)	‘go’

It is not a surprise that the very frequent verbs ‘go’ and ‘see’ should be the ones to show suppletion: this agrees perfectly with the ‘suppletion hierarchy’ proposed by Veselinova (2006, p. 92). The verbs ‘die’ and ‘sit’ also appear in this hierarchy, a step below ‘see’. This corresponds well to the Kalkoti forms *múr* and *bät* which, while they appear historically related to their imperfective counterparts, do have a synchronically distinct perfective stem.

Agreement Agreement in Kalkoti perfective verbs is very limited, and is only present in two basic motion verbs (as illustrated in later examples). Perfective agreement is an interesting topic in Indo-Aryan languages in general, as transitive perfective verbs tend to agree with the object rather than the subject. This is the case in both Palula and Gawri (Liljegren 2016, p. 293; Baart 1999, p. 137), giving those languages a split-ergative agreement pattern mirroring that of the split-ergative case marking. In Kalkoti, however, the alignment question is not applicable to verbs, as the only two verbs that retain agreement in their perfective forms are intransitive. Before illustrating this marginal agreement, the lack of agreement in regular perfective verbs will be illustrated in example 32. This example shows the verb *khaál* ‘ate’ used with both masculine and feminine objects without a change in its form, as well as *duwaál* ‘washed’ with a feminine subject, showing that the *-aál* suffix agrees neither with the subject nor the object (the examples from JK lack overt ergative case marking, which differs from other Kalkoti speakers, but the system of verbal agreement appears to be identical between all speakers).

- (32) a. *púu máwa kh-aál*
boy fruit(M) eat-PFV

‘The boy ate the fruit.’

[JK15d-1]

- b. *zamaán=ä geél kh-aál*
 Zaman=ERG bread(F) eat-PFV
 ‘Zaman ate the food.’ [MJ06a-72]
- c. *yeé bàčee duw-aál*
 mother baby wash-PFV
 ‘The mother washed the baby.’ [JK15d-12]

The verbs that do show agreement in their perfective forms are *yaál*, *yeél* ‘came’ and *goo*, *gee* ‘went’. Interestingly, the distinction between the two forms does not seem to be the same for the two verbs: the masculine plural form is grouped with the masculine singular for ‘come’, and with the feminine form for ‘go’. The different agreement forms are illustrated in table 22 with examples from the MJ06a, MI06c, and JK15g data.

Table 22: Agreement in perfective verb forms.

	Masculine singular	Masculine plural	Feminine
‘came’	<i>soo yaál</i> ‘He came.’	<i>lädkúr yaál</i> ‘The boys came.’	<i>soo yeél</i> ‘She came.’
‘went’	<i>ma bas=dä goo</i> ‘I went by bus.’	<i>lädkúr gee</i> ‘The boys went.’	<i>treer bázaar=par gee</i> ‘The woman went to the market.’

6.1.3 Past tense marking

Imperfective and perfective verbs can take a past tense marker *-s*, which is added after the other suffixes. When applied to imperfective forms, it produces a past imperfective form used to provide information on something that was ongoing in the past. Example 33a from a narrated story shows this form used to ‘set the scene’ of a narration with an ongoing action. In example 33b from the translation of the North Wind story, the past imperfective is used perhaps to give a sense of repetition.

- (33) a. *ä deer bä, káteeyi mälgir, thäl=iĵ b-aan-s*
 one time 1PL.NOM some friend Thal=to go-IPFV.M.PL-PST
 ‘One day we, some friends, were going to Thal.’ [MI06d-2]
- b. *kiteek balä zoor d-een-s, anit musäfär täni šukoođ*
 as_much wind power give-F.IPFV-PST more? traveller REFL cloak
pälät-uun-s
 cover-M.SG.IPFV-PST
 ‘The more the wind blew, the more the traveller covered his cloak around him.’ [JK15e]

Perfective forms are already used primarily or exclusively with past reference, but they may still receive the past *-s*. Liljegren (2013) calls the resulting grammatical category ‘pluperfect’, which will be followed here. The Kalkoti pluperfect is used for giving background in narration, as can be seen in example 34a which immediately follows 33a in the narration and further sets

Table 23: Pluperfect forms.

Perfective		Pluperfect	
<i>čuṇil</i>	‘wrote’	<i>čuṇis</i>	‘had written’
<i>bãṭ</i>	‘sat’	<i>bãś</i>	‘had sat’
<i>goo</i>	‘(he) went’	<i>goos</i>	‘(he) had gone’
<i>yaál</i>	‘(he) came’	<i>yaás</i>	‘(he) had come’
<i>driṣ</i>	‘saw’	<i>dris</i>	‘had seen’
<i>buṣilyiit</i>	‘became hungry’	<i>buṣilyiis</i>	‘had become hungry, was hungry’

the scene of the story. The pluperfect is also used to indicate past states corresponding to verbs that inherently indicate a change of state, as in example 34b. If the perfective form ends in a consonant, the *-s* suffix replaces it in the pluperfect, as can be seen in table 23.

- (34) a. *bä gaaḍi=mi bãś*
 1PL.NOM vehicle=on sit.PFV-PST
 ‘We had sat down in the vehicle.’ [MI06d-3]
- b. *zamaán buṣily-ii-s*
 Zaman become_hungry-PFV-PST
 ‘Zaman was hungry.’ [lit. ‘Zaman had become hungry.’] [MJ06a-152]

Liljegren (2013) suggests that the suffix *-s* is etymologically equivalent to the past tense copula *aas* ‘was’, having fused onto the verb as a suffix. This mirrors a development in Gawri, which has a similar past suffix *-š* corresponding to its past tense copula *aaš*. As Liljegren (2013) also notes, the close relatives of Kalkoti also have a separate ‘(present) perfect’ inflectional category, with overt present-tense marking. No such category has been observed in the Kalkoti data.

6.1.4 Irrealis

Kalkoti has a category of irrealis verb forms, which are formed by attaching a suffix to the imperfective verb stem. It appears in some future contexts, but not always. In most elicited sentences, future time is translated using a Kalkoti imperfective verb. When the irrealis is used, it agrees for the person and number of the subject. First and second person plural irrealis forms have not been observed in the data.

- (35) a. *ma guwaá th-am*
 1PL.NOM what do-IRR.1SG
 ‘What should I do?’ [MJ06a-88]
- b. *tu un kitaáb-uni män-ä*
 2SG.NOM this book-PL read-IRR.2SG
 ‘You will have to read these books.’ [MI06b-89]
- c. *un kitaáb=ä män-il=dä ätu púu niin b-ä*
 this book=GEN read-VN=from this boy sleep go-IRR.3SG
 ‘Having read the book the boy will go to sleep.’ [MI06b-111]

- d. *tríidee* *tin* *tam čhin-ään*
 day_after_tomorrow 3SG.ERG tree cut-IRR.3PL

‘The day after tomorrow they may cut down the tree’ [MI06c-116]

There is some variation between the suffixes *-am* and *-um* for 1SG irrealis forms. This variation seems to depend primarily on the speaker and not the base verb, as both *dam* and *dum* are observed for the 1SG irrealis form of ‘give’.

6.1.5 Imperative

There are not many imperative forms in the data. Those that have been found are shown in table 24 with perfective and imperfective forms for comparison. The imperatives present in the data do not show any variation between singular and plural addressees: *čhin* is both ‘(you) cut!’ and ‘(you all) cut!’.

Table 24: Imperative forms.

Perfective	Imperfective	Imperative	English
<i>päšaál</i>	<i>päšuun</i>	<i>päša</i>	‘show’
<i>yaál, yeél</i>	<i>yùun</i>	<i>yä</i>	‘come’
<i>thäál</i>	<i>thuun</i>	<i>thä</i>	‘do’
<i>čhinil</i>	<i>čhinuun</i>	<i>čhin</i>	‘cut’
<i>goo, gee</i>	<i>buun</i>	<i>čoo</i>	‘go’

It seems that when compared to the imperfective stem, the imperative generally either adds a vowel *-a* or *-ä*, or remains identical to the stem. This seems to correspond to whether the perfective is formed with a long or short suffix. It may be the case that the imperative form should be considered the true verb stem, in which case the perfective forms could all be derived through application of a suffix *-il* followed by vowel assimilation: *thä-il* → *thäál*. I have not attempted this analysis in this thesis, however, as the imperative forms of most verbs are not known.

In the case of ‘go’, the imperative form uses a third suppletive stem. A seemingly identical suppletion pattern is present in Gawri (Baart 1999, p. 181), but not Palula (though a cognate to *čoo* does exist, lexicalized in the form of an interjection; Liljegren 2016, p. 93).

6.2 Nonfinite forms

There are at least three nonfinite verb forms in Kalkoti: converbs, and two nominalized or infinitive-like verb forms.

6.2.1 Converb

Kalkoti converbs correspond to what in Indo-Aryan languages is traditionally known as the ‘conjunctive participle’ (Masica 1991, p. 397). This terminology is discussed below. As far as the data shows, Kalkoti converbs are always formed from the imperfective stem through the addition of the suffix *-i*. Example 36a shows a typical use of the Kalkoti converb form, in linking two temporally sequential clauses. The clause containing the converb is enclosed in brackets. The subject *heedär* must be inside the converbial clause, as it does not receive ergative marking (compare example 36b).

- (36) a. [*heedär šii=thä b-i*] *ä xat čuṇ-il*
 Haider house=to go-CVB INDF letter write-PFV
 ‘Having gone inside, Haider wrote a letter.’ [MJ06a-54]
- b. *heedär=ä ä xat čuṇ-i-s*
 Haider=ERG INDF letter write-PFV-PST
 ‘Haider had written a letter.’ [MJ06a-53]

The term ‘converb’ is defined by Haspelmath (1995) as a verb form used for adverbial subordination. This is in opposition to participles, which Haspelmath defines as essentially having an adnominal modifying function, like adjectives. However, as pointed out by Bickel (1998), this term appears to conflate multiple types of verb forms. In many languages (particularly in South Asia), the subordinating, modifying function of the converb is secondary, and the ‘clause chaining’ function is primary. This appears to be the case in Kalkoti on the basis of examples such as 36a, and clause chaining is further discussed in section 7.3.2.

However, Kalkoti converbs can also participate in other constructions that are more akin to complex predicates than to clause chaining. Example 37 shows the converb *gi* ‘having taken’ combining with the basic motion verbs meaning ‘come’ and ‘go’, to produce the more specialized meanings of ‘bring’ and ‘carry’ (see also example 13a on page 27). This kind of ‘compound verb’ construction is also common in Indo-Aryan languages, and exists in a very similar form in Palula (Liljegren 2016, p. 244).

- (37) a. *soo tipä wää g-i y-ùun*
 3SG.NOM now water take-CVB come-IPFV.M.SG
 ‘He is bringing water (here) now.’ [MI06c-25]
- b. *míšaal-úm sunduu bázaar=par g-i gee*
 man-OBL.PL box? market=to take-CVB go.PFV.M.PL
 ‘The men carried the boxes to the market.’ [JK15d-38]

The construction in example 37a may be on its way to fusing into a single lexicalized verb *giy-* ‘bring’. Such an analysis is chosen by Baart (1999, p. 194) for related forms in Gawri: Baart considers the Gawri *giyaag* ‘to bring’ to be an irregular causative form of *yaag* ‘to come’.

6.2.2 Infinitives and verbal nouns

There is evidence for two further nonfinite forms, infinitives and verbal nouns. Infinitive verbs are generally identical to the verb stem, sometimes with a vowel change. Verbal nouns are formed with suffix containing *l*. They are sometimes identical to the perfective form, but often not. The terminological distinction between infinitives and verbal nouns follows Liljegren (2016, p. 269, 271) regarding two functionally very similar forms in Palula. The infinitives and verbal nouns encountered in the data are shown in table 25.

Table 25: Infinitives and nominalized verb forms.

Perfective	Imperfective	Infinitive	Verbal noun	English
<i>čhinil</i>	<i>čhinuun</i>	<i>čhin</i>	<i>čhinil</i>	‘cut’
<i>goo, gee</i>	<i>buun</i>	<i>bä</i>	<i>bäyil</i>	‘go’
<i>yaál, yeél</i>	<i>yuun</i>	<i>yee</i>	<i>yääl</i>	‘come’
<i>khaál</i>	<i>khuun</i>	<i>khee</i>		‘eat’
<i>driṣ</i>	<i>päṣuun</i>	<i>piš</i>		‘see’
<i>mānil</i>	<i>mānuun</i>	<i>min</i>		‘say, speak’
<i>pīil</i>	<i>piluun</i>	<i>pil</i>		‘drink’
<i>thääl</i>	<i>thuun</i>		<i>thääl</i>	‘do’
<i>bil</i>	<i>būun</i>		<i>bääl</i>	‘become’
<i>čuṇil</i>	<i>čuṇuun</i>		<i>čuṇil</i>	‘write’
<i>waat</i>	<i>wuun</i>		<i>waal</i>	‘come down’

Infinitives are mainly found together with the auxiliary verb *b-* ‘can’, as in example 38a. Verbal nouns are found in a broader range of situations. This includes appearing with the copula to indicate obligation as in example 38b, and with other auxiliary verbs like *daw-* ‘want to’ as in example 38c. The objects of verbal nouns can be indicated in genitive case.

- (38) a. *zamaán geél khee b-uun=ää*
 Zaman bread eat.INF can-IPFV.M.SG=Q
 ‘Can Zaman eat the food?’ [MJ06a-34]
- b. *ma=tä bäyil aas*
 1SG.ACC=DAT go.VN COP.PST
 ‘I had to go.’ [MJ06a-92]
- c. *soo kaar=ä thääl daw-uun*
 3SG.NOM work=GEN do.VN want-IPFV.M.SG
 ‘He wants to work.’ [MJ06a-189]

6.3 Valency modification

There are at least two valency-altering operations: the passive, and the causative.

6.3.1 Passive

The passive is marked by the suffix *-j* after the verb stem. This form is rare in the data, and is only found in a few clauses where it was specifically elicited.

- (39) a. *tam čhin-j-il*
 tree cut-PASS-PFV
 ‘The tree was cut down.’ [MI06c-123]
- b. *angrizii män-j-il*
 English speak-PASS-PFV
 ‘English was spoken.’ [MJ06a-155]

6.3.2 Causative

Explicitly causative clauses are also rare in the data. One elicited causative form is marked by *-aw*, shown in example 40. This same suffix apparently appears in the verb *şaw-* ‘to dress (someone)’, derived from *ş-* ‘to put on (clothes)’, but it is not clear how productive this derivation is.

In the construction in example 40, an introduced ‘causer’ argument becomes the subject of the verb, and the original subject, the ‘causee’, is followed by the word *şaa*. The terminology of causer and causee is from Dixon (2000). Judging by Liljegren’s (2016, p. 325) analysis of a very similar construction in Palula, *şaa* may itself be a grammaticalized form of the verb *şaw-* ‘to dress (someone)’ mentioned above.

- (40) a. *mi tās şaa tam čhin-aw-aál*
 1SG.ERG 3SG.ACC CAUSEE tree cut-CAUS-PFV

‘I made him cut the tree.’

[MI06c-122]

This pattern, whereby the causee gets a special marker, is one of several possibilities that Dixon (2000) identified for causatives of transitive clauses, but not the most common. However, Kalkoti exhibits other causative patterns, too. Table 26 shows imperfective and perfective forms of the verbs meaning ‘see’ and ‘show’. The latter is apparently derived from the imperfective stem of the former, and the imperfective forms appear to be identical.⁷ There is no overt *-aw* morpheme as with *čhin-aw-aál*. Another difference is that the verb ‘show’ acts syntactically like a ditransitive verb akin to ‘give’ (see section 7.1.3), and the causee is marked with the clitic *=thä* ‘to’, and not *şaa*.

Table 26: Distinction between ‘see’ and ‘show’ in perfective forms.

	IPFV.M.SG	PFV
‘see’	<i>pāşuun</i>	<i>driş</i>
‘show’	<i>pāşuun</i>	<i>pāşaaál</i>

⁷Some data suggests that *pāşuun* ‘(he) shows’ may have a high pitch on the second syllable, distinguishing it from *pāşuun* ‘(he) sees’ by way of tone. However, this contrast would require further investigation to confirm.

7 Clauses

Kalkoti syntax is generally head-final, and this is the case in the structure of clauses, too. The verb is almost always the last part of its clause. The only exceptions to this appear to be with certain subordinate clauses (see section 7.3.1).

The discussion of clauses is divided into three sections: section 7.1 categorizes Kalkoti clauses according to their argument structure, section 7.2 describes negation and question formation, and section 7.3 describes the different ways clauses are combined.

7.1 Clause types

This section presents a categorization of Kalkoti clauses into five types. The first three types are based on intransitive, transitive, and ditransitive verbs. The fourth type is based on the copula, which cannot be grouped together with other verbs. The fifth type is quotative, and does not include any kind of verb or copula.

7.1.1 Intransitive

Some examples of intransitive verbs are *äs-* ‘laugh’ and *jäl-* ‘burn’ as seen in example 41. These can formally be distinguished from transitive verbs in that their subject does not receive ergative marking in perfective clauses.

- (41) a. *piṭeeš äs-il*
little_girl laugh-PFV
‘The little girl laughed.’ [JK15d-57]
- b. *šii jäl-een*
house burn-IPFV.F
‘The house is burning.’ [JK15d-67]

The intransitive basic motion verbs *y-* ‘come’ and *b-* ‘go’ are highly morphologically irregular (see e.g. section 6.1.2 and table 25), and are the only verbs that exhibit agreement in their perfective forms. They are also used in several idiosyncratic constructions, including as the second part of a ‘compound verb’ as illustrated in section 6.2.1.

The verb *y-* ‘come’ is used in some constructions together with a dative experiencer, as shown in example 42. The ‘sensation’ noun (e.g. ‘cold’, ‘sleep’) is the subject and controls the verb agreement. This kind of construction is common throughout Indo-Aryan (Masica 1991, p. 346) and traditionally known as a ‘dative subject’ construction, though the subjecthood of the dative noun phrase is at most partial. The dative experiencer is generally subject-like in that it can control reflexive pronouns (Masica 1991, p. 354). This is true of Kalkoti, too, as illustrated by example 56a in section 7.1.4.

- (42) a. *puṭooš=thä šidäl y-uun*
boy=DAT cold(M) come-IPFV.M.SG
‘The child is feeling cold.’ [MJ06a-70]
- b. *ma=ṭä niin y-iin*
1SG.ACC=DAT sleep(F) come-IPFV.F
‘I am sleepy.’ [MJ06a-77]

There is a similar, but structurally different, construction with *b-* ‘go’. In example 43, the subject must be *zamaán* and not *niin* in order to explain the masculine verb agreement.

- (43) a. *zamaán niin b-uun*
 Zaman sleep GO-IPFV.M.SG
 ‘Zaman is sleeping.’ [MJ06a-149]

There is also an intransitive construction where the noun *agä* (likely meaning ‘rain’) is used together with the verb *b-* ‘become’ to form a weather predicate.

- (44) a. *biyaál bääḍ agä b-il*
 yesterday much rain? become-PFV
 ‘Yesterday it rained a lot.’ [MJ06a-32]

Hunger and thirst are expressed by the intransitive verbs *bušily-* and *trušily-*, respectively (both of which form the perfective with the T-verb ending *-iit*). However, the basic meanings of those verbs do not seem to be stative ‘be hungry/thirsty’, but rather expressing a change of state: ‘become hungry/thirsty’. In example 45, the perfective form is used to indicate present hunger, and past hunger is indicated with the pluperfect form.

- (45) a. *sumeerä bušilyiit*
 Sumaira become_hungry.PFV
 ‘Sumaira is hungry.’ [lit. ‘Sumaira has become hungry.’] [MJ06a-72]
- b. *zamaán bušilyii-s*
 Zaman become_hungry.PFV-PST
 ‘Zaman was hungry.’ [lit. ‘Zaman had become hungry.’] [MJ06a-152]

7.1.2 Transitive

Some examples of transitive verbs are *kh-* ‘eat’ and *päš-* ‘see’ as seen in example 46. Transitive verbs require their subject to be marked for ergative case in perfective forms.

- (46) a. *zamaán=ä geél kh-aál*
 Zaman=ERG bread eat-PFV
 ‘Zaman ate the food.’ [MJ06a-72]
- b. *tu umä=dä räni šii päš-uun=ää*
 2SG.NOM here=from 3PL.GEN house see-IPFV.M.SG=Q
 ‘Can you see their house from here?’ [MJ06a-178]

This kind of ‘split ergativity’ based on aspect is common in Indo-Aryan, and also has parallels in many other language families (Dixon 1994, p. 100). Table 27 illustrates the ergativity split using sentences from the MJ06a data that have the same subject ‘he’.

Pronominal subjects provide the most reliable tests for ergative patterns in Kalkoti. This is because at least one Kalkoti consultant uses plain oblique forms for ergative nouns, without

Table 27: Intersection of aspect and transitivity.

	Perfective	Imperfective
Intransitive	<i>soo yaál</i> 'He came.'	<i>soo khusuun</i> 'He is coughing.'
Transitive	<i>tä ä phit máril</i> 'He killed a fly.'	<i>soo wää piluun</i> 'He is drinking water.'

the overt marker =*ä*, and the oblique case is only distinct from the direct case in plural nouns (see section 5.4.1).

Kalkoti has several constructions where a semantically nonspecific transitive verb (e.g. 'do', 'give') combines with a more semantically specific noun to form a complex predicate. This is commonly known as a 'light verb' or 'conjunct verb' construction Liljegren (2016), though the term 'light verb' is sometimes also used in a broader sense, including 'compound verbs' such as the ones described in section 6.2.1 (Butt 2010). These constructions are commonly used in Indo-Aryan languages, especially as a strategy to expand the inventory of predicates, as lexical verbs may be a closed class (Butt 2010). Example 47 shows two such light verb constructions in Kalkoti.

- (47) a. *biyaál mi [traam thää́l]*
 yesterday 1SG.ERG work do.PFV
 'I worked yesterday.' [MI06c-50]
- b. *čikār [šiiš d-iin]*
 bird flight give-IPFV.F
 'The bird flies.' [MJ06a-67]

Even though the syntactic object is part of the predicate, this construction is still transitive, since the subject is ergative with perfective light verbs (example 47a). Some light verbs appear to vary with full intransitive verbs, however.

- (48) a. *soo [čikaar d-iin]*
 3SG cry give-IPFV.F
 'She is crying.' [MJ06a-13]
- b. *pée čikār-il*
 girl cry-PFV
 'The girl cried.' [JK15d-83]

7.1.3 Ditransitive

Ditransitive predicates in Kalkoti distinguish themselves by having an argument marked by the dative clitic =*thä* in addition to a standard object. The most typical ditransitive construction is with the verb *d-* 'give', but there are other verbs such as *päš-* 'show' which also have the same syntactic pattern. The dative 'indirect' object apparently always precedes the standard 'direct' object.

- (49) a. *tä ä deer lädrúj-um=thä troo tangur dit*
 3SG.ERG a time? boy.PL-OBL.PL=DAT three pear give.PFV
 ‘Then he gave the boys three pears.’ [JK15g]
- b. *pée tani ústaaz=thä fuṭú päs-aál*
 girl REFL teacher=DAT photo show-PFV
 ‘The girls showed pictures to the teacher.’ [JK15d-35]

In the Kalkoti ditransitive constructions, the direct object (*troo tangur* and *fuṭú* in the examples above) has the semantic role of a theme, while the argument that receives different (dative) marking has the semantic role of a recipient. This corresponds to the ‘indirective’ pattern of ditransitive alignment (Malchukov et al. 2011), which is the norm in Indo-Aryan languages (Masica 1991, p. 367).

7.1.4 Copula clauses

The Kalkoti copula has two forms: *in* ‘is’ and *aas* ‘was’, and does not show any kind of agreement. Its basic use is to link a subject noun to predicative adjectives, nouns, or adverbial elements, as in the following examples.

- (50) a. *dàrin šíšeél in*
 ground dry COP.PRS
 ‘The ground is dry.’ [JK15d-68]
- b. *soo prufisär aas*
 3SG professor COP.PST
 ‘He was a professor.’ [MJ06a-48]
- c. *čhätree dār=dii in*
 umbrella door=at COP.PRS
 ‘The umbrella is next to the door.’ [MJ06a-43]

Omission of the present tense copula *in* occurs in some situations (but the past tense copula is never omitted). This phenomenon is known to occur in some Indo-Aryan languages (Masica 1991, p. 337), including Palula (Liljegren 2016, p. 304). Like in Palula, the omission of the Kalkoti copula in affirmative clauses appears to be restricted to equational sentences such as 51a, where the complement is a noun. In negative sentences, the present tense copula is always replaced by the negation particle *nä* (see section 7.2.1 for discussion).

- (51) a. *soo prufisär*
 3SG professor
 ‘He is a professor.’ [MJ06a-47]
- b. *soo uma nä*
 he here NEG
 ‘He is not here.’ [MJ06a-216]

The copula can be used in several other constructions, which require a case-marked noun phrase, or more generally a postpositional phrase, instead of the copula subject. Together with a postpositional phrase indicating location, the copula can be used in an existential sense, as shown in example 52.

- (52) a. *un kāmra=mi du čidiriđ aas*
 this room=in two window COP.PST
 ‘There were two windows in this room.’ [MJ06a-131]

The copula is also involved in a specialized construction for predicative possession. The possessor is indicated either with the clitic =*dii* ‘at, next to’, or in genitive case (using the genitive clitic =*ä* for nouns, and special forms for pronouns). These options seem to be interchangeable in at least some situations. This is illustrated by example 53, showing a sentence from an elicitation task that was performed with two different speakers, who used different markers for the possessor.

- (53) a. *täsi púu=dii trä šii in*
 3SG.GEN boy=at three house COP.PRS
 ‘His son has three houses.’ [MJ06a-140]
- b. *täsi púu=ä trä šii in*
 3SG.GEN boy=GEN three house COP.PRS
 ‘His son has three houses.’ [MI06b-140]

The two constructions are shown with pronoun possessors in example 54. For nouns, the genitive marker =*ä* is homophonous to the ergative marker. The marker used in the possession construction can nevertheless be shown to be genitive, as it can be replaced by genitive pronouns (which differ from ergative pronouns), as shown in example 54b.

- (54) a. *ma=dii tipä wax nă*
 1SG.ACC=at now time NEG
 ‘I don’t have time now.’ [MJ06a-100]
- b. *täsi ä péé aas*
 3SG.GEN INDF girl COP.PRS
 ‘She had a daughter.’ [MJ06a-101]

Two essentially identical constructions are found in Palula (Liljegren 2016, p. 310), where the construction with genitive case is associated with inalienable possession. The situation in Kalkoti appears to be similar, as predicative possession of body parts and family members (typically considered inalienable) always uses the genitive in the available data. It is not an absolute rule, however, as both constructions exist when describing alienable possession.

- (55) a. *mi du theer in*
 1SG.GEN two hand COP.PRS
 ‘I have two hands.’ [MJ06a-137]
- b. *thi kiteek lädkúr in*
 2SG.GEN how_many boy.PL COP.PRS
 ‘How many children do you have?’ [MJ06a-96]

The copula, like certain intransitive verbs, can form part of a ‘dative subject’ construction, such as the ones in example 56. Example 56a shows the partial subjecthood of the dative subject, as it can control the reflexive possessive pronoun *täni*.⁸ The copula is also involved in a construction indicating obligation, which is formed with a dative subject and a nominalized verb.

- (56) a. *lädkúr=thä täni mälgar yaád in*
 boy?=DAT REFL friend memory COP.PRS
 ‘The boy remembered his friend.’ [MI06b-109]
- b. *tu=tä un kitaáb män-il in*
 2SG.ACC=DAT this book read-VN COP.PRS
 ‘You will have to read this book.’ [MJ06a-89]
- c. *ma=tä khur=dä bäyil aas*
 1SG.ACC=DAT foot=from go.VN COP.PST
 ‘I had to go on foot.’ [MJ06a-91]

7.1.5 Quotative clauses

Quoted speech is introduced by the uninflecting word *alu*, which can be used on its own, without a verb. It can also be used for subordination, together with an overt verb (see section 7.3.1). It bears a resemblance to the Gawri counterpart *äro* (Baart 1999, p. 147).

- (57) a. *rin alu nä*
 3PL.NOM QUOT no
 ‘They said no.’ [JK15d-22]
- b. *soo täs=thä alu [ma daan nä gäd b-uun]*
 3SG.NOM 3SG.ACC=DAT QUOT 1SG.NOM tooth NEG remove.INF can-IPFV.M.SG
 ‘He said: I do not know how to pick out the teeth.’ [MJ06c-12]

7.2 Negation and questions

This section moves beyond affirmative declarative clauses and describes the formation of negative clauses, polar questions, and wh-questions.

⁸The translation of the sentence has a singular subject ‘boy’, though the form used in the Kalkoti sentence is *lädkúr*, which is typically plural. The plural oblique marker *-um* is absent, however, indicating that *lädkúr* really is used in a singular sense in this sentence.

7.2.1 Negation

Negative clauses with a lexical verb use the particle *nä* placed before the verb. There is no special prohibitive form for negative imperatives, which use *nä* in the same way as declarative clauses.

- (58) a. *ma nä jän-uun*
 1SG NEG know-IPFV.M.SG
 ‘I don’t know.’ [MJ06a-81]
- b. *phuṇ nä čhin*
 flower NEG pluck.IMP
 ‘Don’t pluck the flowers.’ [MJ06a-213]

The particle *nä* on its own functions as the negative form of the present-tense copula, as in example 59a. The past tense copula combines regularly with negation as *nä aas* (though this combination is rare in the data as an accidental gap in elicitation; see example 63b on page 54 for one example). However, example 59b from the very end of the Pear story narrative uses *nä* on its own with what appears to be past-tense reference.

- (59) a. *ma=dii atu kitaáb nä*
 1SG.NOM=at that book NEG
 ‘I don’t have that book.’ [MJ06a-127]
- b. *ä čädeer tangur nä*
 INDF basket pear NEG
 ‘[... and saw that] a basket of pears was not there; was missing.’ [JK15g]

The observed forms of the Kalkoti copula (including in its functions to mark existential predicates) distinguished by tense and polarity are summarized in table 28.

Table 28: Observed forms of the Kalkoti copula or existential marker.

	Present	Past
Affirmative	<i>in, Ø</i>	<i>aas</i>
Negative	<i>nä</i>	<i>nä aas, nä ?</i>

The ability of the standard negation particle *nä* to stand on its own as a ‘negative existential’ marker distinguishes Kalkoti from both Palula and Gawri. Palula and Gawri use straightforward cognates to *nä* for verbal negation (Palula *na* and Gawri *nä*). However, negative existence in Palula and Gawri is expressed using contracted (but distinct) forms of the negation particle + the copula: Palula *náinu* (Liljegren 2016, p. 413) and Gawri *naat* (Baart 1999, p. 157).

The Kalkoti situation where an independent negative existential marker is identical to a verbal negator has parallels in several other languages, both in IA and outside of it: Croft (1991) cites Marathi as a synchronically similar example. According to the ‘negative existential cycle’ proposed by Croft (1991), systems of this kind arise when a negative existential gets reinterpreted as a marker of verbal negation, replacing an older verbal negator, which is what has

happened in Marathi. However, the Kalkoti particle *nä* (like particles of similar form throughout IA) appears to trace back to the Old Indo-Aryan negation particle *na*, which was used for verbal negation (Masica 1991, p. 390). Given this etymology, it seems more likely that an older Kalkoti construction corresponding to **nä in* ‘is not’ developed to just *nä* through elision of the copula *in*, instead of undergoing phonological fusion as in Palula. This would not be so surprising, given that such elision also occurs in some positive sentences.

7.2.2 Polar questions

Kalkoti forms polar questions with a sentence-final question particle =*ää*. It is a bound morpheme that follows the last word in the sentence regardless of its word class, so it is considered a clitic rather than a suffix.

- (60) a. *tu umä=dä räni šii päš-uun=ää*
 2SG.NOM here=from 3PL.GEN house see-IPFV.M.SG=Q
 ‘Can you see their house from here?’ [MJ06a-178]
- b. *geél täyaar in=ää*
 food ready COP.PRS=Q
 ‘Is the food ready?’ [MJ06a-50]
- c. *raathä jumá=ää*
 tomorrow Friday=Q
 ‘Is tomorrow Friday?’ [MI06d-7]

Many IA languages predominantly mark polar questions using intonation alone (Masica 1991, p. 388), but if this strategy is available in Kalkoti, it is not present in the data. The polar questions containing the =*ää* clitic end with a flat mid-low pitch similar to that of declarative statements.

7.2.3 Wh-questions

Wh-questions are formed using an appropriate question word in the place of the questioned constituent. Question words either start with *g* (*guwaá* ‘what’, *gaan* ‘why’) or with *k* (*kii* ‘who’, *kiteek* ‘how much’, *kaan* ‘which’). Some examples of wh-questions are shown in example 61.

- (61) a. *soo guwaá th-uun*
 3SG.NOM what do-IPFV.M.SG
 ‘What is he doing?’ [MJ06a-4]
- b. *kii geél kh-aál*
 who bread eat-PFV
 ‘Who ate the bread?’ [MJ06a-217]
- c. *tu=dii kiteek rupee in*
 2SG.ACC=at how_much money COP.PRS
 ‘How much money do you have?’ [MJ06a-209]

7.3 Multicausal constructions

This section gives a brief overview of the different ways that clauses are combined to form complex structures.

7.3.1 Subordination

Only a few examples of subordinate clauses are present in the data. The types of subordination mentioned here represent very different structures, but they are grouped here as the limited data does not allow a detailed discussion of any one type.

Verbal nouns may head a nominalized clause, which may be embedded in another sentence as a noun phrase. Example 62 shows the nominalized clause *bas=ä yääl*, corresponding to the finite clause *bas yeél* ‘the bus came’.

- (62) *ma* [*bas=ä yääl*]=*een* *uma biš-uun*
 1SG.NOM bus=GEN come.VN =until? here sit-IPFV.M.SG

‘I will keep sitting here until the bus comes.’ [MJ06a-202]

Example 63 shows two examples of relative clauses. The structure of the relative clause follows the ‘relative-correlative’ pattern common in Indo-Aryan (Andrews 2007): the relativized noun (*kitaáb*) is preceded by *kaan* ‘which’ in the subordinate clause, and is represented by a pronoun (*täs, soo*) in the main clause. The subordinate clauses end with *dä*, whose function is not entirely clear.

- (63) a. [*thi theer=mi kaan kitaáb in dä*] *täs ma=tä päš-a*
 2SG.GEN hand=in which book COP.PRS ? 3SG.ACC 1SG.ACC=DAT show-IMP

‘Show me the book which is in your hand.’ [MI06b-221]

- b. [*mi kaan tāluun kh-aál dä*] *soo säyi päčeel nä aas*
 1SG.ERG which rice eat-PFV ? 3SG.NOM well cooked.F? NEG COP.PST

‘The rice that I ate was not cooked well.’ [MI06b-219]

The particle *alu* (which also marks quoted speech; see section 7.1.5) may introduce postverbal subordinate clauses with a connection to quoted speech. Example 64 illustrates this with an example from the North Wind story translation.

- (64) *fäsälá thääl alu* [*räs=daa kii šugood gäd-il too soo tikir*]
 agreement did that 3SG.ACC=from who cloak remove-PFV ? 3SG.NOM strong

‘(They) agreed that the one who took off (the traveller’s) cloak was stronger.’ [JK15e]

There is also a subordinator *aan* which appears once in the data, shown in example 65.

- (65) *heedär=ä driš aan* [*aslam xat čuñ-uun-s*]
 Haider=ERG see.PFV that Aslam letter write-IPFV.M.SG-PST

‘Haider saw Aslam write the letter.’ [MJ06a-55]

7.3.2 Clause chaining

The term ‘(clause) chaining’ is defined by Longacre (2007) as referring to a sentence structure where two independent verbs cannot be joined together in the same sentence. Instead, a ‘chain’ of clauses is used, where only the last (or sometimes, first) is fully independent. Longacre (2007) identifies New Guinea as the linguistic area where chaining structures are most ubiquitous, but they are nevertheless present to some degree in many other languages, including Indo-Aryan ones. Clause chaining in Palula has been described by Liljegren (2016, p. 350), and Kalkoti appears to have a similar system. Non-final Kalkoti clauses are typically headed by converbs, which can be used both with declarative and imperative final verbs. This is illustrated with two elicited sentences in example 66.

- (66) a. [*heedār šii=thä b-i*] *ä xat čuŋ-il*
 Haider house=to go-CVB a letter write-PFV
 ‘Haider went inside and wrote a letter.’ [MJ06a-54]
- b. [*un tam čhin-i*] *räs čun-a*
 this tree cut-CVB 3SG.ACC split-IMP
 ‘Cut this tree and split it!’ [MI06c-121]

Clause chaining is common in narration. Example 67 shows two sentences from the start of the JK15g narration, both of which have an initial clause with a converb accompanied by *tée* ‘and’. The placement of *tée* inside the initial clause is deliberate and corresponds to the pauses made by the narrator, which in both sentences followed *tée*.

- (67) a. [*ä meš kiç gän-i tée*] *tangur thunä-s*
 INDF man sack? tie-CVB and pear collect.PFV?-PST
 ‘A man having a sack tied (to his back) was plucking pears.’ [JK15g]
- b. [*šileđ nām-i tée*] *tä du čädeer tangur pur-il*
 ladder put-CVB and 3SG.ERG two basket pear fill-PFV
 ‘Placing a ladder he filled two baskets with pears.’ [JK15g]

8 Discussion

This chapter discusses the thesis project as a whole and gives suggestions for future work.

8.1 Methodological limitations

This study has not involved any fieldwork or extensive data gathering, instead mainly relying on previously gathered data. This method brings with it some obvious limitations. Some data (U23a) was gathered with a native Kalkoti speaker during the project, but this was only a small fraction of the data used in the project and could only mitigate the limitations to a small degree.

The issues with this method include the following. I have at several times been unsure in the glossing of certain lexemes due to their low frequency. I have also sometimes had trouble finding examples with minimal contrasts, which has been particularly noticeable in the analysis of the tone system. Some inflectional forms, such as imperatives, are too infrequent in the data to be certain about their morphological formation. As most of the data consists of elicited translations (from Urdu) to Kalkoti, the precise semantics of most Kalkoti lexemes and inflectional categories is not possible to investigate with this data.

Another limitation is the fact that all the Kalkoti language consultants are male. Language usage often differs (sometimes drastically so) between different demographic groups, but the degree to which that affects Kalkoti has not been possible to investigate in this thesis.

Despite the limitations, the method has made this project feasible within the scope of this Bachelor's thesis, and the data has been sufficient to make some progress in understanding the linguistic structure of Kalkoti.

8.2 Discussion of synchronic results

Throughout the chapters 4 to 7, the grammatical analysis of Kalkoti has been presented along with examples from the Kalkoti data to support the analysis. Each part of the analysis is discussed where it is presented, and compared with other languages where relevant. Not every part of the analysis is equally certain, and where the analysis is uncertain, that has been noted. Although I have tried to prevent it, it is still possible that some data has been misinterpreted, especially considering that the analyses have not been checked by a native Kalkoti speaker. The transcription of tone throughout the thesis is especially uncertain, and a more careful study involving minimal pairs would be required to be more certain in the analysis of tone.

8.3 Historical perspective and relation to other languages

While the historical development of Kalkoti has not been the main focus of this thesis, it is an interesting topic worthy of discussion and further study. In the thesis, some features of Kalkoti have been compared to other languages, mainly Palula and Gawri. The results only corroborate the findings of Liljegren (2009) and Liljegren (2013): that Kalkoti genealogically belongs to the same western Shina group as Palula and Sauji, but has been heavily influenced by Gawri. The Shina roots of Kalkoti can be seen perhaps most clearly in the morphology of nouns and the forms of pronouns, but can be found in most areas of the grammar. Gawri influence is apparent in historical phonological developments, including the development of the tone system, as well as in some morphosyntactic features such as the placement of the indefinite article relative to adjectives, and the morphologization of past tense marking. The vowel changes in gender-agreeing adjectives, as well as the suppletive imperative stem *čoo* 'go!' also appear to

be more similar to Gawri. As for vocabulary, many Kalkoti words show apparent Palula cognates indicating inherited Shina vocabulary, while many others are apparently loaned from Gawri. In several cases, however, the historical Kalkoti phonological changes appear to have brought its inherited word-forms close to or identical to the Gawri equivalents, in which case it does not seem possible to say with certainty if a word has been inherited from Shina or if it has been borrowed from Gawri.

8.4 Future work

As this study has only been a very preliminary investigation of most Kalkoti features, opportunities for future research are abundant. A more detailed study of the tone system would be required in order to have a more certain theory of how it functions. Such a study would need to find minimal pairs for more tone contrasts, and it would also need to test the behavior of different tone classes in different environments in a structured way, to see how tones phonetically affect each other. The morphology of imperatives and nonfinite verb forms also requires further investigation, as does the semantics of the different Kalkoti verb tenses. The structural status of Kalkoti case markers can be further investigated (along the lines of the tests Butt and King (2004) present for Urdu and Hindi). Complex clause structures have only been briefly mentioned in this thesis, and are also a suitable topic for further study. In addition to the mentioned areas of synchronic description, the historical development of Kalkoti including its contact with Gawri can also be explored further. This would likely involve building on the work of Liljegren (2009) in understanding the ancestor language of Palula, Sauji, and Kalkoti, and identifying in more detail which features of Kalkoti are due to Gawri influence.

9 Conclusion

In this thesis, the major phonological and morphosyntactic characteristics of Kalkoti have been described, mostly based on data collected in 2006 and 2015, but also including some data gathered as part of the thesis project.

Some important or otherwise notable phonological findings are the following. Kalkoti consonants are mostly typical for its area, but are distinguished by lacking voiced aspirated consonants and /h/. Consonant clusters are heavily restricted. Kalkoti vowels show an asymmetric length distinction, with six long vowels and four short vowels. Kalkoti has also developed a complex tone system similar to that of the neighboring Gawri language. There are both marked high tones and marked low tones, both of which contrast with the lack of any marked tone.

Some important or otherwise notable morphosyntactic findings are the following. Kalkoti nouns show a contrast between direct and oblique cases only in plural nouns, and more specific case functions are indicated by clitics attached to the oblique form. The ergative and genitive cases are marked by identical morphemes in nouns, a pattern of syncretism that is uncommon in Indo-Aryan languages. Verbs show a core distinction between imperfective and perfective forms. Imperfective forms show agreement with the gender and number of the subject, while perfective forms generally lack agreement. Kalkoti syntax is generally head-final and shows structures typical of Indo-Aryan languages, concerning split ergativity, dative subject constructions, light verbs, and patterns of clause modification.

Multiple characteristics of Kalkoti described in this thesis point to an origin of Kalkoti as a Shina language: pronouns, nominal morphology, much verbal morphology, and much vocabulary all generally agree with the Shina language Palula. However, many features of Kalkoti are shared with Gawri but not Palula: the development of the tone system, some adjectives, placement of the indefinite article, some patterns of suppletion, and much vocabulary in general. This points to extensive contact between Kalkoti and Gawri having brought Kalkoti closer to Gawri over time.

References

- Andrews, Avery D. (2007). Relative clauses. In: *Language Typology and Syntactic Description*. Ed. by Timothy Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 206–236.
- Baart, Joan L. G. (1997). *The sounds and tones of Kalam Kohistani*. Islamabad: National Institute of Pakistani studies, Quaid-i-Azam University.
- (1999). *A sketch of Kalam Kohistani grammar*. Islamabad: National Institute of Pakistani Studies, Quaid-i-Azam University.
- Baart, Joan L. G. and Muhammad Z. Sagar (2004). *Kalam Kohistani texts*. Islamabad: National Institute of Pakistan Studies, Quaid-i-Azam University.
- Bashir, Elena (2003). Dardic. In: *The Indo-Aryan Languages*. Ed. by George Cardona and Danesh Jain. New York: Routledge.
- Bickel, Balthasar (1998). Review article: Converbs in cross-linguistic perspective. In: *Linguistic Typology* 2.3, pp. 381–397.
- Boersma, Paul and David Weenink (2018). *Praat: doing phonetics by computer*. Version 6.0.37. URL: <http://www.praat.org/>.
- Butt, Miriam (2010). The light verb jungle: Still hacking away. In: *Complex predicates: cross-linguistic perspectives on event structure*. Ed. by Mengistu Amberber, Brett Baker, and Mark Harvey. Cambridge: Cambridge University Press, pp. 48–78.
- Butt, Miriam and Tafseer Ahmed (2011). The redevelopment of Indo-Aryan case systems from a lexical semantic perspective. In: *Morphology* 21, pp. 545–572.
- Butt, Miriam and Tracy Holloway King (2004). The status of case. In: *Clause structure in South Asian languages*. New York: Springer, pp. 153–198.
- Cardona, George and Dhanesh Jain (2003). General introduction. In: *The Indo-Aryan Languages*. Ed. by George Cardona and Danesh Jain. New York: Routledge.
- Corbett, Greville G. (1991). *Gender*. Cambridge University Press.
- Croft, William (1991). The evolution of negation. In: *Journal of linguistics* 27.1, pp. 1–39.
- Dixon, Robert M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- (2000). A typology of causatives: form, syntax and meaning. In: *Changing valency: Case studies in transitivity*. Ed. by Robert M. W. Dixon and Alexandra Y. Aikhenvald. Cambridge: Cambridge University Press.
- Haspelmath, Martin (1995). The converb as a cross-linguistically valid category. In: *Converbs in cross-linguistic perspective*. Ed. by Martin Haspelmath and Ekkehard König. Berlin: De Gruyter Mouton, pp. 1–55.
- Knobloch, Nina (2020). A grammar sketch of Sauji: An Indo-Aryan language of Afghanistan. Bachelor's Thesis. Stockholm University.
- Liljegren, Henrik (2009). The Dangari Tongue of Choke and Machoke: Tracing the proto-language of Shina enclaves in the Hindu Kush. In: *Acta Orientalia* 70, pp. 7–62.
- (2013). Notes on Kalkoti: A Shina language with strong Kohistani influences. In: *Linguistic Discovery* 11.1, pp. 129–160.
- (2016). *A Grammar of Palula*. Berlin: Language Science Press.
- (2019). Gender typology and gender (in)stability in Hindu Kush Indo-Aryan languages. In: *Grammatical gender and linguistic complexity*. Ed. by Francesca Di Garbo, Bruno Olsson, and Bernhard Wälchli. Berlin: Language Science Press, pp. 279–328.
- Liljegren, Henrik and Naseem Haider (2011). *Palula vocabulary*. Islamabad: Forum for Language Initiatives.

- Longacre, Robert E. (2007). Sentences as combinations of clauses. In: *Language Typology and Syntactic Description*. Ed. by Timothy Shopen. 2nd ed. Vol. 2. Cambridge University Press, pp. 372–420.
- Malchukov, Andrej, Martin Haspelmath, and Bernard Comrie (2011). Ditransitive constructions: a typological overview. In: *Studies in ditransitive constructions: a comparative handbook*. Ed. by Andrej Malchukov, Martin Haspelmath, and Bernard Comrie. Berlin: De Gruyter Mouton, pp. 1–64.
- Masica, Colin P. (1991). *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Rensch, Calvin R. (1992). Patterns of Language Use among the Kohistanis of the Swat Valley. In: *Languages of Kohistan*. Ed. by Calvin R. Rensch, Sandra J. Decker, and Daniel G. Hallberg. National Institute of Pakistani studies, Quaid-i-Azam University, pp. 3–62.
- Veselinova, Ljuba N. (2006). *Suppletion in Verb Paradigms: Bits and pieces of the puzzle*. Amsterdam: John Benjamins Publishing Company.
- Yip, Moira (2002). *Tone*. Cambridge: Cambridge University Press.

Stockholm University
SE-106 91 Stockholm, Sweden
Telephone +46 (0)8 16 20 00
<https://www.su.se/>



Stockholm
University