

Abstract

A Jewel Inlaid: Ergativity and Markedness in Nepali

Luke Lindemann

2019

Nepali presents with a complex case marking pattern in which ergative case is obligatory in perfective transitive clauses, disallowed in unaccusative intransitive clauses and copular clauses, and varies with the nominative elsewhere. Where ergative marking is variable, its usage correlates with a variety of semantic and pragmatic factors. The purpose of this investigation is to precisely delineate the grammatical domains for which ergative marking is variable and to provide a unified analysis of the semantic and pragmatic factors that correlate with its expression.

The study of pragmatic phenomena requires the implementation of multiple strategies for collecting language data. The data for this investigation come from four converging lines of inquiry: descriptions of the Nepali pattern in the literature, targeted elicitation with thirteen native speakers, the implementation of a grammaticality judgment survey in Kathmandu in 2016, and the analysis of a published corpus of spoken Nepali.

The analysis found ergative marking to be obligatory in perfective main clauses and variable in subordinate clauses. What appears to be active marking in intransitive clauses is analyzed as ergative marking in transitive clauses with covert objects. The only categorical split is the distinction between perfective and non-perfective verb forms. Every other association was found to be non-categorical.

These non-categorical associations include a positive correlation between subjects with inanimate reference and the expression of ergativity in common nouns, and a

negative correlation between first person pronouns and ergativity in the pronominal domain. This follows expected patterns of marking based on the types which are most frequent in discourse. Ergative marking is somewhat associated with highly individuated objects, but not with affected objects.

Ergative marking is positively associated with characterizing or individual-level predicates, kind readings, categorical propositions, and strong construals of quantifiers. There was no correlation found between ergative marking and agency or volitionality.

The unified analysis of these associations contributes to theories of Optional Ergative Marking and to optional case marking systems in general. The main claim is that the Nepali ergative marks an effector of the event described by the clause. This term refers to a participant which is implicated in enacting and effecting the event, but is not necessarily its main controller or instigator. As a component of the ergative case marking system, it has a pragmatic usage, implicating the subject as a participant in a prototypically transitive event. Aspects of this analysis contribute to the general theory of Optional Ergative Marking and its relation to argument proto-roles. Associations between the ergative and prototypical properties of a transitive event arise from the meaning of the ergative marker as an effector. This analysis also provides a straightforward explanation for the lack of volitional correlations in Nepali that we find in other languages with variable ergativity.

Other semantic and pragmatic features are associated with discourse prominence. These include the correlation with categorical propositions and characterizing predicates. Here the associations are attributable to general principles of semantic markedness. Variable ergativity represents the presence of pragmatic implicatures of various strengths. Gradient markedness oppositions can lead to the conventionalization of these associations into semantic entailments. This is demonstrated for English gender marking, the association between ergative marking and semantic properties of the

transitive subject in Nepali, and the association between ergative marking and Nepali perfective verb forms.

A Jewel Inlaid: Ergativity and Markedness in Nepali

A Dissertation
Presented to the Faculty of the Graduate School
of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Luke Lindemann

Dissertation Director: Dr. Claire Bowern

May 2019

Copyright © 2019 by Luke Lindemann

All rights reserved.

Contents

1 Preliminaries	1
1.1 Introduction to the Problem	2
1.2 Nepali and its Speakers	5
1.2.1 A Brief History of the Language	7
1.2.2 The Languages of Nepal	8
1.3 Ergativity and Basic Alignment Patterns	10
1.3.1 Alignment Splits	15
1.4 Relevant Features of Nepali Grammar	17
1.4.1 Nominal Morphology and Case	18
The <i>-le</i> postposition	19
The <i>-lāi</i> postposition	20
1.4.2 Verbal Morphology	21
Copular Clauses	22
1.5 Ergativity in Nepali	23
1.5.1 The development of Ergativity in Indo-Aryan	25
2 Methodologies	28
2.1 The Necessity of a Multifaceted Approach	30
2.2 Elicitations	32
2.2.1 Procedure	33
2.2.2 Advantages and Limitations of Elicitation	35
2.3 Grammaticality Judgment Survey	36
2.3.1 Survey Respondents	40
2.3.2 Advantages and Limitations of the Kathmandu Survey	40
2.4 Corpus Analysis	43
2.4.1 The NNSP Corpus Sample	43
2.4.2 Annotation Procedures	45
2.4.3 Advantages and Limitations of Corpus Analysis	53
2.5 The Problem of Lect	54
3 Theories	59
3.1 General explanations of the Nepali Pattern	59
3.1.1 Case Marking in Traditional Grammar	60
3.1.2 Disambiguation	63
3.1.3 Animacy	64

3.1.4	Stage-Level and Individual-Level Predication	67
	Stage-Level Predication in Nepali	70
3.1.5	Perfectivity	72
3.1.6	The Intransitive Domain	74
3.1.7	Ergative Marking in Particular Structures and Tenses/Aspects	76
	Hypothetical Future	76
	Optative	76
	Definite Future	77
	Modal Constructions	77
3.1.8	Multiple Factor Analysis	78
3.1.9	Summary	79
3.2	General Theories of Optional Ergativity	80
3.2.1	Overview	80
3.2.2	Terminology	81
	Ergative Case	81
	Typology of Split Case Marking systems	82
3.2.3	Discriminative Function	83
3.2.4	Transitivity	86
	The Transitivity Hypothesis	86
	Argument Proto-Roles	92
	Causal Structure	96
3.2.5	Markedness	101
	The Nominal Hierarchy	107
	A Formal Implementation of the Nominal Hierarchy	112
	Markedness Extensions: Prototypes, Reversals, and Hierarchies	116
	Criticisms of Markedness	118
3.2.6	Discourse Prominence	120
	Focus, Topic, and the Question Under Discussion	123
	Intonation, Word Order, and Discourse Particles	127
	Categorical Propositions	130
	Prominence and Markedness	135
4	Observations	138
4.1	Ergative marking and the interpretation of the Event	140
4.1.1	Ergativity and the Nepali Verb Form	140
	The Perfective Domain	142
	The Present Imperfective Domain	144
	The Past Habitual Imperfective	146
	The Definite Future	147
	The Hypothetical Future, Optative, and Imperative Mood	149
	Subordinate Clauses	151
4.1.2	Intransitive Clauses	154
	Unaccusative Predicates	158
	Unergative Semelfactives/Verbs of Emission	160
	Ergativity and Volitionality in the Intransitive Domain	162

Telicity and Motion Verbs	163
Unergative Atelic Predicates and “Do” Light Verbs	166
Ergativity and Transitivity	169
4.1.3 Ergativity in Copular Clauses	169
Brief observations on the dual copula system	171
4.1.4 Summary of the domains of ergative marking	174
4.1.5 Individual-Level Predication	175
Judgments on Ergative Marking and Habitual Interpretations	177
4.1.6 Transitivity and the Event	186
4.2 Ergative marking and the Subject	188
4.2.1 The Nominal Hierarchy	189
Inanimate Common Nouns	193
4.2.2 Agency and Volitionality	196
4.2.3 Kind Readings of the Subject	198
4.2.4 Sets and Strong Construals of Quantifiers	202
4.2.5 Summary	204
4.3 Argument Realization and Case Frequency in the NNSP Sample . . .	205
4.4 Ergative marking and the Object	211
4.4.1 Ergative marking and the Realization of the Object	212
4.4.2 Affectedness	216
4.4.3 Connectedness to the Object	217
4.4.4 Summary	218
4.5 Ergativity and the Discourse	219
4.5.1 Topic, Focus, and Word Order	219
4.5.2 Categorical Propositions	223
4.5.3 QUD, Definiteness, and Aboutness	225
4.5.4 Ergative Marking and the (Implied) Question	229
4.5.5 Summary	232
4.6 Overall Summary of Feature Correlations	232
5 The Syntactic Profile of Nepali Ergativity	235
5.1 Morphological and Syntactic Ergativity	236
5.1.1 Subjecthood Diagnostics	237
Reflexivization	237
Coordination	239
Control	240
Syntactic ergativity in Dyirbal	241
5.1.2 Agreement	242
5.1.3 Lack of Covert Ergative Case in Adjectives	242
5.2 Theories of Syntactic Ergativity	244
5.2.1 Structural Case, Inherent Case, and Dependent Case	244
5.2.2 Structural versus Inherent Case	247
Nonfinite Clauses	249
5.2.3 Marantz’ Ergative Case Generalization	250
5.2.4 Legate’s Typology of Ergative Case	253

Lack of Accusative Morphology	255
Caseless DPs	256
Cases in Nonfinite Contexts	257
Multiple Absolutives	258
Agreement	259
5.2.5 Conclusions	259
6 Discussions	261
6.1 A Prototype Analysis of the Nepali Ergative	262
6.1.1 Peripheral Usages of <i>-le</i>	263
6.1.2 Agents and Effectors	265
6.2 On Markedness and Variability	270
6.2.1 Opposition with a Zero Form	271
6.2.2 Obligatory Marking and Grammatical Context	272
6.2.3 Gradient Markedness in English Gender Marking	277
6.2.4 Gradient Markedness on the Transitive Subject	280
6.3 Variable Ergativity and Prominence	285
6.3.1 Predicate Prominence	287
6.4 Ergativity and Event Structure	290
6.4.1 Ergative Marking in Transitive Clauses	293
6.4.2 The Connection between Ergativity and Perfectivity	296
7 Conclusions	300

List of Figures

1.1	Nepali-speaking Regions of South Asia	6
1.2	L1 Speakers of Language Varieties in Nepal (2011 Census)	9
1.3	Provinces of Nepal	9
1.4	Accusative and Ergative Alignment, adapted from Dixon (1994: 9) . .	11
1.5	Neutral, Tripartite, and Active-Inactive Alignment	13
1.6	Incidence of Ergativity in the World's Languages (data from WALS online)	14
1.7	Ergative Case-Marking in a Language with a Perfective/Imperfective Split	17
1.8	Nepali Case Marking on Common Nouns	18
1.9	Nepali Case Postpositions	19
1.10	Nepali Verb Forms	21
1.11	Nepali Copulas	23
1.12	Ergative marking in Nepali (according to Li 2007)	24
2.1	Nepali-speaking Elicitation Consultants	32
2.2	Kathmandu Survey: Item C3	38
2.3	Kathmandu Survey: Item C3 (English Translation)	38
2.4	Kathmandu Survey Respondent Summary	41
2.5	NNSP Annotation Example (from interview V001001001)	44
2.6	Interviewer Data from the NNSP Sample Interviews	45
2.7	Birth Regions of Nepali Speakers	55
2.8	Ergative Marking in Imperfective Clauses by Speaker	58
3.1	Nepali <i>Kāraka</i> : Argument Roles (Pradhān 1944)	60
3.2	Nepali <i>Vibhakti</i> : Nominal Cases (Pradhān 1944)	62
3.3	Hopper and Thompson (1980)'s Transitivity Prototype	86
3.4	Dowty (1991)'s Proto-roles	93
3.5	General vs. Narrow Meaning of Unmarked; based on Andrews (1990: 154)	103
3.6	Aspects of Marked and Unmarked Forms (Battistella 1996: 71)	104
3.7	The Nominal Hierarchy as represented in Dixon (1994: 85)	108
3.8	Ergative and accusative splits in four languages	110
3.9	Hindi Differential Object Marking; based on (Aissen 2003: 459) . . .	114
3.10	Summary of the Prototypical Subject and Object	118
3.11	Summary of Features from McGregor (2010)	135

4.1	Ergative Marking Results by Verb Form	141
4.2	Survey Results: ERG/NOM with Perfective Transitive Clauses	143
4.3	Survey Results: ERG/NOM with Simple Present Transitive Clauses .	146
4.4	Survey Results: ERG/NOM with Present Continuous Transitive Clauses	146
4.5	Survey Results: ERG/NOM with Definite Future Transitive Clauses .	148
4.6	Survey Results: ERG/NOM with Hypothetical Future Transitive Clauses	150
4.7	Corpus Results: ERG/NOM in Subordinate Clauses	154
4.8	Ergative Marking Results for Intransitive Clauses	157
4.9	Survey Results: ERG/NOM with Unaccusative Predicates	159
4.10	Survey Results: ERG/NOM with Verbs of Emission	161
4.11	Survey Results: ERG/NOM with Telic Intransitives	164
4.12	Survey Results: ERG/NOM with Unergative Atelic Verbs	167
4.13	Survey Results: ERG/NOM with Copulas	170
4.14	Summary of the Domains of Ergativity in Nepali Clauses	175
4.15	Survey Results: ERG/NOM with Simple Present Transitive Clauses .	182
4.16	Ergative Marking Results in Survey Section C	183
4.17	Survey Results: Individual Strategies for Ergative Marking in Section C	183
4.18	Corpus Results: Individual Strategies for Ergative Marking in Section C	185
4.19	Ergative Correlations and the Event	187
4.20	Survey Judgments for Subject Type	191
4.21	Survey Results: ERG/NOM by Subject Animacy	192
4.22	Percentage of Ergative Marking on Overt Transitive Subjects	193
4.23	Arguments with Inanimate Reference in Different Argument Positions	194
4.24	Survey Results for Questions with Kind Readings (Section G)	201
4.25	Ergative Correlations and the Subject	204
4.26	Transitive Subject Types in the NNSP Sample	206
4.27	Transitive Object Types in the NNSP Sample	207
4.28	Intransitive Subject Types in the NNSP Sample	208
4.29	Copular Subject Types in the NNSP Sample	208
4.30	Oblique Argument Types in the NNSP Sample	209
4.31	Number of Expressed Arguments in Transitive Clauses	210
4.32	Survey Judgements for Transitive Clauses with Overt/Elided Objects	214
4.33	Corpus Results for Object Realization	215
4.34	Ergative Correlations and the Object	218
4.35	Survey Results for Section B	221
4.36	Argument Positions in Transitive Present Clauses	222
4.37	Ergative Marking Results for Categorical Propositions	230
4.38	Ergative Marking Results and Case on the Survey Question	231
4.39	Overall Summary of Ergative Feature Correlations	233
6.1	Properties of Transitive Subjects and Instruments	266
6.2	Nepali and Hindi Case Markers in Croft's Conceptual Space for Participant Roles	270
6.3	Noun Phrase Number Marking in English and Nepali	274
6.4	Scale of Markedness	280

Glossing Conventions and Abbreviations

For the transliteration of Nepali I employ a modified version of the International Alphabet of Sanskrit Transliteration (IAST). I endeavor to adhere to modern Nepali phonology over traditional orthography. Thus there is no vowel length distinction except for *a/ā*, and the traditional *anusvara* is variably represented as a nasal vowel or consonant.

1	first person	MIR.PRES	mirative present
2	second person	NEG	negation
3	third person	NOM	nominative
ABL	ablative	NON.FIN	non-finite
AOR	aorist (Tibetan glosses)	OBL	oblique
ACC	accusative	OPT	optative
ARCH.PRES	archaic present	PASS	passive
BEN	beneficiary	PERF	perfective
CAUS	causative	PL	plural
CLF	classifier	PRES	simple present
COND	conditional	PST	past
CONJ	conjunctive adverbial	PST.HAB	past habitual
CONT	continuous	PST.PERF	past present
CT	contrastive topic	RED	reduplicant
COP	copula		
DAT	dative		
DEF.FUT	definite future		
DEM	demonstrative		
ERG	ergative		
EGO	egophoric (Tibetan glosses)		
F	feminine	DAM	Differential Agent Marking
FOC	focus	DSM	Differential Subject Marking
FUT	future	MIA	Middle Indo-Aryan
GEN	genitive	NIA	Modern Indo-Aryan
GNOM	gnomic (Tibetan glosses)	NNSP	Nepali National Spoken Corpus
HON	honorific	OEM	Optional Ergative Marking
HYP.FUT	hypothetical future	OIA	Old Indo-Aryan
IMP	imperative	QUD	Question Under Discussion
IMPF	imperfective		
INSTR	instrumental		
LOC	locative		
M	masculine		
MIR.PST	mirative past		

Acknowledgments

Many thanks to the members of my dissertation committee: Claire Bowern, Jim Wood, Kristine Hildebrandt, Paul Kockelman, and Veneeta Dayal. Without your expertise, coaching, and patience this dissertation would not exist.

Many thanks also to my Nepali friends whose judgments and opinions form the core of this work: Sushant Banjara, Bibek Basnet, Rijan Maharjan, Timila Dhakwa, Prashanta Kharel, Anobha Gurung, Kamal Sharma, Min Gurung, Sabin Pyakural, and Uddhab Pyakurel.

I am indebted to those members of the Yale linguistics faculty past and present who have advised me and offered their comments and support at different stages throughout this process: Ashwini Deo, Larry Horn, Maria Piñango, Stephen Anderson, and Robert Frank. I am also grateful to the following linguists who have given me their invaluable advice and opinions: Tikaram Poudel, Madhav Pokharel, Tyler Peterson, Saartje Verbeke, Dan Raj Regmi, Mark Turin, Hans Boas, and Prem Phyak.

I would also like to thank my research assistants Biplob Acharya and Sajju Tamang and my tutor Geeta Manandhar for their invaluable contributions to my fieldwork in Nepal, as well as the members of the Boudha Coterie: Christopher Geissler, Mikaela Chase, Zara and Loki.

Finally, thank you to those members of the Yale graduate student community who commented on versions of this work over the years: Matthew Tyler, Joshua Phillips, Rikker Dockum, Rashad Ullah, Alyria Harris, Sara Sánchez-Alonso, and Sarah Babinski.

*To Hans, for wash-bears and oak-cats,
To Ashwini, for the Sauraha eye exam,
To the Khatris, for Sangam,
To Claire, a lighthouse for the unmoored,
To Mom, who showed me how,
To Dad, that grain of sand not yet a pearl*

Chapter 1

Preliminaries

गौरीको मञ्चु लीला—महल चहकिलो
रत्नशोभाऽभिराम ।

'हिमाल', कविशिरोमणि लेखनाथ पौड्याल

Gauri's beautiful palace of play,
bedecked with bright lovely jewels...

The Himalaya, Lekhnath Paudyal
(translation by Hutt 1988)

1.1 Introduction to the Problem

The Nepali language presents with a complex pattern of subject case marking that is unique among the languages of the Indo-Aryan family, although variations on this pattern are found in other languages throughout the Himalayas and in other ergative languages around the world.

- (1) a. *sāno khetā-le eutā bhyāgutā pokhari-mā dekh-yo*
small child-ERG one.CLF frog pond-LOC see-PERF.3.SG
'The small child saw a frog in the pond.' [AG]
- b. *sunita bāthi chin*
Sunita clever.F COP.PRES.3.SG.F
'Sunita is clever.' [TD]
- c. *surya/surya-le cricket khel-cha*
Surya/Surya-ERG cricket play-PRES.3.SG
'Surya plays cricket.' [BA]

In (1a) the verb *dekhnu* ‘to see’ is marked with a perfective form (typically construed in the past tense). The subject, *khetā*, is marked by the ergative case postposition (-*le*).¹ (1b) is a present tense copular clause. Here the subject, *Sunita*, is bare. In (1c), the verb *khelnu* ‘to play’ is in the simple present tense verb form. The translation given is the habitual “Surya plays cricket,” but depending on the context it might also refer to an ongoing event (“Surya is playing cricket”) or a future-oriented event (“Surya will play cricket”). In this sentence the ergative marker *-le* may be either present or absent without affecting the grammaticality of the clause.

In Nepali, there are grammatical domains in which ergative marking is **obligatory** on subjects, as in (1a), regions in which it is **disallowed**, as in (1b), and regions in which it is **variable**, as in (1c). Where it is variable, ergative marking correlates with a variety of semantic and pragmatic factors. These domains of variable ergative

1. The initials after each quote refer to the elicitation consultant with whom I discussed the particular sentence in question (see section 2.2).

marking are not marginal to the language: they represent the rule rather than the exception. The purpose of this investigation is to delineate these grammatical domains and provide a unified explanation of the semantic and pragmatic factors that correlate with the ergative.

This analysis makes two claims about ergative marking in Nepali. The first claim concerns the particular meaning of the *-le* marker, which also marks instrumental arguments and reason clauses. I argue that all of these usages of *-le* mark an **Effector** of the event described by the clause. This term refers to a participant which is implicated in enacting and effecting the event, but is not necessarily its main controller or instigator. As a component of the ergative case marking system, *-le* has a pragmatic usage, implicating the subject as a participant in a prototypically transitive event. This claim applies particularly to the Nepali language, although aspects of this analysis will be applicable to other languages with variable ergativity. In some of these languages, ergative marking correlates with agency or volitionality, but this is not what we find in Nepali. This typological difference is a natural consequence of the analysis; the Nepali ergative is a marker of an effector rather than a marker of a prototypically transitive subject.

The second claim is more broadly applicable to variable case marking systems in which a morphological form varies with its absence. These are also known as optional case marking (OCM) systems. Such systems are amenable to an analysis in terms of a markedness asymmetry, in which “optionality” represents the presence of pragmatic implicatures of various strengths. These implicatures can become conventionalized as semantic entailments, leading to domains in which marking is either obligatory or disallowed. This analysis predicts that marking in an OCM system will always correlate with increased discourse prominence, and in the context of Nepali ergativity this leads to ergative marking being associated with topicality, definiteness, and characterizing or individual-level predicates.

The remainder of the current chapter consists of background material on ergativity and relevant information about Nepali. Section 1.2 is an overview of the Nepali language and the speakers of Nepali in the linguistic context of the Himalayas. Section 1.3 is a discussion of ergativity and split-ergative patterning. In section 1.4, I discuss relevant features of the Nepali grammar, particularly the nominal case marking system and verbal morphology. In section 1.5, I give a basic description of the Nepali ergative pattern and briefly describe the development of ergativity in the Indo-Aryan family.

Chapter 2 is a discussion of the methodologies I employed for data collection. The study of pragmatic phenomena requires the implementation of multiple strategies for collecting language data, as I discuss in 2.1. Sections 2.1-2.4 consist of a detailed overview of each of the methodologies I employed: targeted elicitations with native speakers, a written grammaticality judgment survey, and a discourse analysis from a corpus of transcribed spoken Nepali.

Chapter 3 profiles the theoretical underpinnings of this investigation. The first part of this chapter consists of an overview of the explanations of the Nepali pattern that have been given in the literature. In particular, I discuss explanations based on Disambiguation, Animacy, Individual-Level Predication, and Telicity. These explanations each cover particular domains of the grammar rather than providing a unified explanation for the pattern. In the second part of chapter 3, I discuss theories of variable ergativity, which is often referred to as Optional Ergative Marking (OEM). In particular, I examine OEM and its relation to three broader topics in the literature. The first is Transitivity as it relates to argument prototypes and force-dynamic causal structure. The second is Markedness as a cornerstone of ergative theory. The third is Discourse Prominence and in particular the notion of categorical propositions.

In chapter 4, I present the results of the investigation. This chapter is structured in terms of the various feature associations between ergative marking and properties of the event (4.1), semantic properties of the subject (4.2), the types of arguments

and their frequency in the corpus (4.3), semantic properties of the object (4.4), or properties of the discourse (4.5). I conclude that all of these features either result from emphasizing the subject as an effector of a transitive event or as an element which is discourse-prominent.

In chapter 5, I examine the ergative case marking system and its relationship to the syntax. I argue that ergativity is limited to case morphology and has a minimal effect on the syntactic structure of Nepali, even in comparison to related languages like Hindi. I suggest that a full analysis of syntactic ergativity in Nepali should rely on a dependent case analysis.

In chapter 6, I present the details of a prototype analysis of the Nepali ergative. In section 6.1, I argue for the notion of *-le* as the marker of an Effector, and discuss its relationship with the feature association mentioned in chapter 4. In 6.2, I discuss the relationship between markedness and variability, and argue that markedness represents a cline of opposition. I analyze gradient markedness with the example of English gender marking, and then apply this to gradient markedness on the Nepali subject. Section 6.3 is a discussion of discourse prominence and its relation to categorical propositions and characterizing predicates. In section 6.4, I discuss the grammaticalization of obligatory associations related to ergative marking and event structure.

1.2 Nepali and its Speakers

Nepali (Glottocode: nepa1254, ISO 639-3: npi) is a Northern Zone Indo-Aryan language spoken in Nepal, Bhutan, and India (Hammarström et al. 2018). Nepali and three related languages spoken to the west (Dotyali, Jumli, and Palpa) make up the Eastern Pahari macrolanguage family.² Nepali is spoken by about 24 million

2. The Glottolog refers to this macrolanguage (Glottocode: east1436, ISO 6393: nep) as “Eastern Pahari” and the Ethnologue refers to it as “Nepali (macrolanguage).” I will use the term “Nepali” to refer to the language and all of its dialects and “Eastern Pahari” to refer to the macrolanguage group. I use “Nepali” as a demonym for residents of the country of Nepal and “Nepalese” as a general

people total.

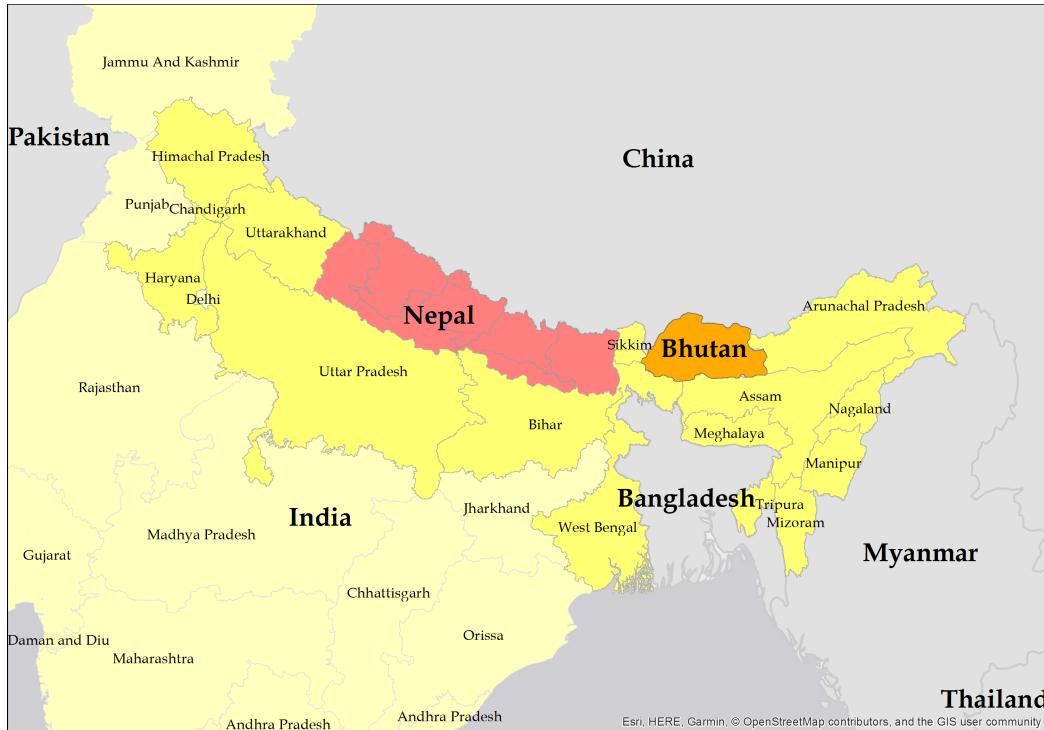


Figure 1.1: Nepali-speaking Regions of South Asia

In Nepal, Nepali is the official national language as well as the main language of government and education (Simon and Fennig 2018). It is spoken by around 21 million people throughout Nepal, of which 12.3 million (or 47% of the country's population) are L1 speakers (Malla 2012).³ In India, Nepali is spoken by about 2.9 million people. Figure (1.2) depicts the regions of South Asia in which Nepali is spoken: Nepal (red), Bhutan (orange), and surrounding Indian states (dark yellow), particularly Sikkim and the Darjeeling region of northern West Bengal.⁴ It is an official state language in the Indian states of Sikkim and West Bengal. In the country of Bhutan, there

3. This number is an aggregate of Census responses of Nepali or any Nepali dialect, of which Acchami, Baitadeli, Bajhangi, Dailekhi, Dadeldhuri, Bajureli, and Darchuleli were listed in the census. About 11.8 million speakers (45%) gave “Nepali” as their mother tongue.

4. These and subsequent maps were created with Esri’s ArcGIS software and the ESRI World Topographical basemap (Esri 2018).

currently reside approximately 85,000 speakers of Nepali (Simon and Fennig 2018). Approximately 100,000 Bhutanese Nepali speakers (who share cultural and historical ties with Nepalese communities elsewhere) were expelled from Bhutan in the late 20th century, leading to a refugee crisis and the emigration of Nepali-speaking Bhutanese refugees to Nepal, India, the United States, and several European countries (Hutt 1996). Additionally, there are significant populations of Non-resident Nepalis who work and reside in other countries. The 2011 Nepal Census records some 2 million who reside abroad (7% of the population), with significant populations in India, the Arab States of the Persian Gulf, Malaysia, Europe, the United States, and Australia (Malla 2012).

1.2.1 A Brief History of the Language

Nepali was originally the language of the Khas ethnic group from the hill regions of Western Nepal. Other names for the language that are in common usage in Nepal include *Khas Kurā* (meaning the language of the Khas) and *Pahari* or *Parbatiya* (referring to the hill area). The name *Nepāl* originally referred only to Kathmandu Valley. The valley has been a dominant cultural and political force in the region for thousands of years due to its control over trade routes through the Himalayas to Tibet. In the latter 18th century, Khas from the Gorkha hill region conquered and unified the kingdoms and hill tribes across modern Nepal and India and supplanted the (primarily Newari-speaking) former rulers of the kingdoms of Kathmandu Valley. From their capitol in Kathmandu, the Khas established Nepali as a lingua franca and language of royal patronage, and in modern times Nepali is the official language of the state.

Certain aspects of Nepali have more of an affinity with the languages of Western India and particularly Rajasthan than they do with geographically adjacent languages like Hindi and Bihari (Wright 1877, Grierson 1904a). From ancient times there have

been waves of migration into the hills from India, and these groups of people mixed with the Khas. The most noteworthy of these waves came during the Mughal invasions of the 15th century, when there were migrations from the areas around modern-day Rajasthan. Grierson, who notes a particularly close relationship between Nepali and the Mewari/Marwari dialects of Rajasthan, theorizes that these newcomers from Rajasthan mixed with local populations and supplanted the local language.

However, Wallace (1982) and Hutt (1988) object to this claim, arguing that Nepali already existed in the region and has been minimally influenced by Western Indo-Aryan languages. In fact, the earliest known Nepali language inscriptions date to the 13th century (Poudel 2008). The extent to which migration from Rajasthan and language mixture has had an impact on the grammar of Nepali is an open question. Grierson (1904a) also describes many aspects of Nepali grammar which are unusual for Indo-Aryan languages (including the peculiar usage of the ergative case, as well as grammatical declensions and honorific constructions), and which he believes to be due to influence from the Tibeto-Burman languages of the Nepal. This is not surprising considering the linguistic diversity of Nepal and the surrounding Himalayan region.

1.2.2 The Languages of Nepal

About half the population of Nepal speaks Nepali as a first language. The 2011 Nepal Census records 122 other languages that are spoken as a mother tongue in Nepal. In Figure (1.2) I have tabulated the number of speakers for major language families and groups represented in Nepal. About 82% of the population speaks Nepali or another Indo-Aryan language (of which Maithili, Bhojpuri, and Tharu have over a million speakers each). Many of these languages are spoken in the Terai, the southern plains which border India. Tibeto-Burman languages comprise the second biggest group, representing 17% of the population. Tamang is the only Tibeto-Burman language with more than a million speakers, and is spoken in the central part of the country.

Nepali Variety	Number of L1 Speakers	Percentage of Population
Nepali (including dialects)	12,333,525	46.6%
Other Eastern Pahari Languages	506,572	3.0%
Language Families	Number of L1 Speakers	Percentage of Population
Indo-European	21,742,298	82.1%
Tibeto-Burman	4,592,014	17.3%
Austro-Asiatic	51,912	0.2%
Dravidian	34,829	0.1%
Other	73,508	0.3%
Largest Minority Languages	Number of L1 Speakers	Percentage of Population
Maithili (Indo-Aryan)	3,092,530	11.7%
Bhojpuri (Indo-Aryan)	1,584,958	6.0%
Tharu (Indo-Aryan)	1,529,875	5.8%
Tamang (Tibeto-Burman)	1,353,311	5.1%
Newari (Tibeto-Burman)	846,557	3.2%
Magar (Tibeto-Burman)	788,530	3.0%

Figure 1.2: L1 Speakers of Language Varieties in Nepal (2011 Census)

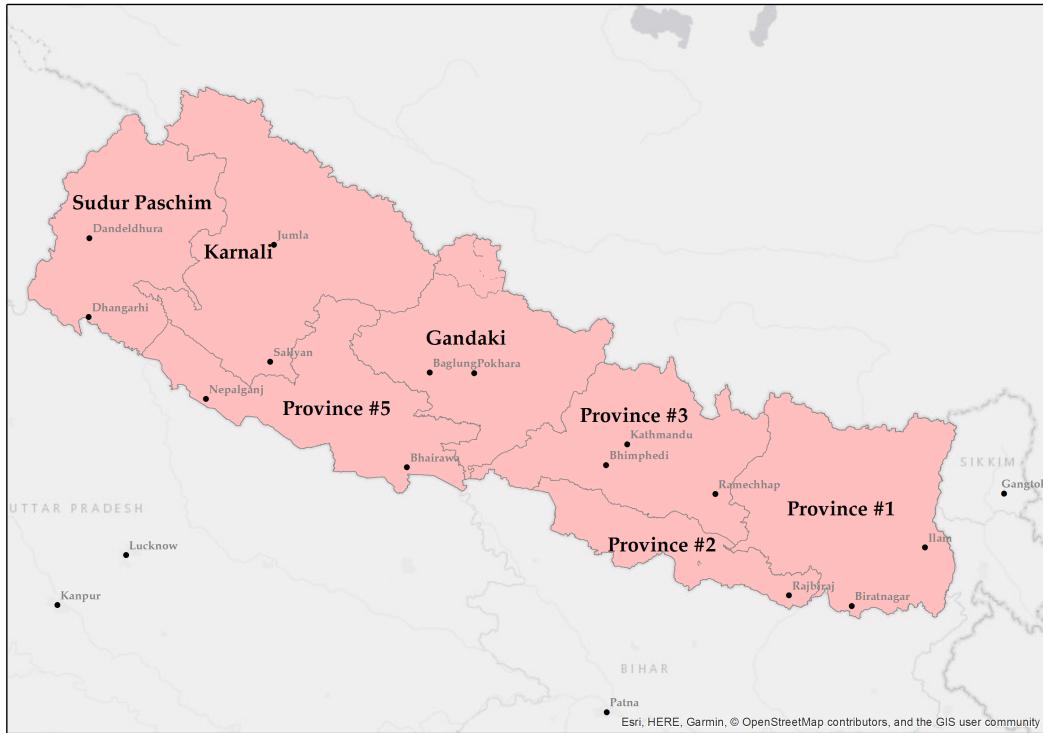


Figure 1.3: Provinces of Nepal

There are significant populations of Newari speakers (particularly around Kathmandu Valley), as well as Kiranti languages (Kiranti, Rai, and Limbu) in the eastern hill regions. Gurung and Magar are spoken in the west. In the far west, there is a dialect

continuum between Nepali and other Eastern Pahari languages (Grierson 1904a). The greatest dialectal diversity of Nepali also exists in this region.⁵

There is widespread language shift due to the dominance of Nepali. Many of the languages of Nepal have become endangered, partially as a result of the intentional policies of the Nepali government in the latter half of the 20th century, although there are also political efforts to preserve linguistic diversity in the country (Turin 2007, Eagle 2008).

I now turn to a discussion of basic alignment patterns and ergative-accusative splits before focusing on the Nepali alignment system in detail.

1.3 Ergativity and Basic Alignment Patterns

There are many ways for a language to mark the grammatical relations of the arguments in a clause. In a canonical transitive clause, the core arguments are the **transitive subject (S_t)** and the **direct object (O)**. In a canonical intransitive clause, the sole core argument is the **intransitive subject (S_i)**.⁶

- (2) a. *Sita saw Ram.*
 $S_t \quad V \quad O$

- b. *Sita laughed.*
 $S_i \quad V$

The majority of the world's languages have some means of morphologically demarcat-

5. These provinces were formed in 2015 in accordance with the Constitution of Nepal, and replaced the previous system of Administrative Zones and Development Regions. As of February 2019, three of them have been given official names. (Patrika 2018, Deuba 2018).

6. I am adopting the usage of these symbols ($S_i/S_t/O$) from Keenan (1984). Two other common conventions are S/A/O (as in Dixon 1994) and S/A/P (as in Comrie 1978). I have chosen Keenan's convention because I wish to emphasize an abstract representation of syntactic categories without making a claim about semantic roles (Agent and Patient). I do not intend to imply by this usage the theoretical position that Subject is a universal category in every language, although there is evidence for a unified (S_t and S_i) subject in Nepali (as I discuss in the Syntax section). Nor am I implying that nominative-accusative alignment is the default pattern among the world's languages and that ergative-absolutive is a marked alignment.

ing core arguments (Siewierska 2013b). The two most common strategies are nominal case-marking (morphological case or adpositions), and verbal crossreference (verbal agreement in person/gender/number). A language may use either or both of these strategies.

Nominative-accusative alignment (or just accusative alignment) refers to a system in which S_i and S_t pattern together in opposition to O. **Ergative-absolutive alignment** (or just ergative alignment) refers to a system in which S_t and O pattern together in opposition to S_i . Figure (1.4) illustrates the basic pattern. In an accusative language the O is given a special accusative case marking while S_i and S_t are typically unmarked, and in an ergative language S_t is given a special ergative case marking while S_i and O are unmarked.⁷

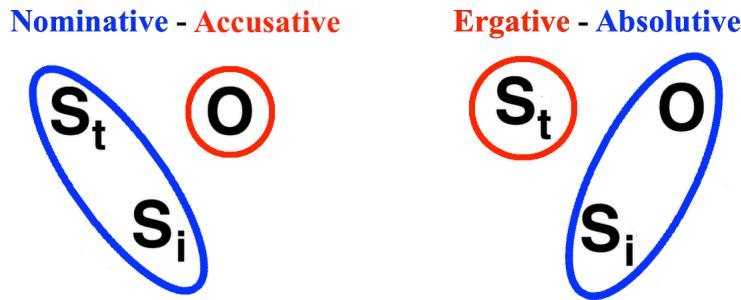


Figure 1.4: Accusative and Ergative Alignment, adapted from Dixon (1994: 9)

The following example of nominative-accusative patterning comes from the Indo-Aryan language Bangla:

- (3) a. *sītā* *rām-ke* *dekh-elo*
 sita.F.NOM ram.M-ACC see-PERF.3.F.SG
 'Sita saw Ram.'⁸
- b. *sītā* *hās-elo*
 sita.F.NOM laugh-PERF.3.F.SG

7. Overt nominative case-marking in accusative languages is rare; overt absolutive case-marking in ergative languages is rarer (Comrie 2013a).

8. These two Bangla examples were provided to me by Rashad Ullah.

‘Sita laughed.’

Example (3a) is a transitive sentence with two arguments, *Ram* and *Sita*. Example (3b) is an intransitive sentence with a single argument, *Sita*. The transitive subject S_t of (3a) and the intransitive subject S_i of (3b) are both in the **nominative** case, which is unmarked in Bangla. The transitive direct object O of (3b) is in the **accusative** case and is marked by the case marker *-ke*. Thus the language follows a nominative-accusative case-marking pattern. Furthermore, the verb agrees with the S_t in (3a) and the verb agrees with the S_i in (3b). Thus verbal agreement also follows a nominative-accusative pattern.

By contrast, consider these sentences in a related Indo-Aryan language, Hindi:

- (4) a. *rām-ne cidiyā dekh-ī*
ram.M-ERG sparrow.F.ABS see-PERF.F.SG
'Ram saw a sparrow.' (Deo and Sharma 2006: 376)
- b. *rām hās-e*
ram.M.ABS laugh-PERF.M.SG
'Ram laughed.' (Deo and Sharma 2006: 376)

For the transitive sentence in (4a) the case-marking pattern is reversed: there is marking on the S_t , **Ram**, and no marking on the O, **cidiyā**. This is an ergative-absolutive pattern. The S_t is in the **ergative** case, while the O is in the (unmarked) **absolutive** case. Furthermore, the verb agrees with the O, **cidiyā**. For the intransitive sentence in (4b), the single argument S_i is in the (unmarked) absolutive case, and the verb agrees with the intransitive subject (S_i).

There are other possible alignment patterns: a **neutral alignment** is one in which all three arguments are grouped together such that they are not demarcated from each other. Conversely, a **tripartite alignment** is one in which all three pattern differently such that S_i receives absolutive marking, S_t receives nominative marking, and O receives ergative marking.

Finally, there are languages in which the subject of an intransitive clause S_i is split in two: in clauses for which the argument is agent-like (S_a) the argument patterns with S_t , and in clauses where the argument is more patient-like (S_p) the argument patterns with O. This is the **active-inactive alignment** pattern.⁹ These additional three alignment patterns are schematized in Figure (1.5).

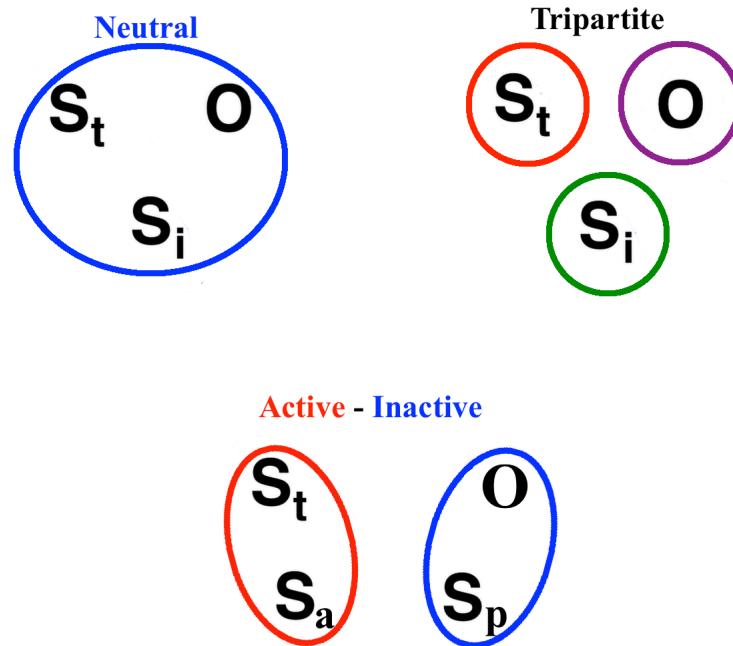


Figure 1.5: Neutral, Tripartite, and Active-Inactive Alignment

This is a basic typology of alignments. In fact, many languages are mixed systems, employing one pattern in some domains and another pattern in others (a topic which will be discussed in the next subsection). Depending upon the clause, Nepali case-marking may follow a Nominative-Accusative, Ergative-Absolutive, Active-Inactive, or Tripartite alignment. As a general rule, I refer to any morphological form which uniquely demarcates the transitive subject S_t as “ergative.”

Furthermore, in the domain of case marking, there is usually only one morpho-

9. Other terms for this alignment pattern are agentive-patientive alignment, stative-active alignment and Split-S alignment. A subtype is Fluid-S alignment, in which either case is possible, but the usage of one or the other has an effect on the semantic interpretation of the clause.

logically overt case with an opposition between a morphologically marked form and a morphologically unmarked form.¹⁰ In an accusative language, the O will take a special (accusative) case marker or inflection while the S_t and S_i retains a default or zero form. In an ergative language, the ergative S_t will take a special (ergative) case marker or inflection while the O and S_i retains a default or zero form. There are different theoretical positions on whether the absolute and nominative represent the same abstract case in such languages (as discussed in the Syntax section).

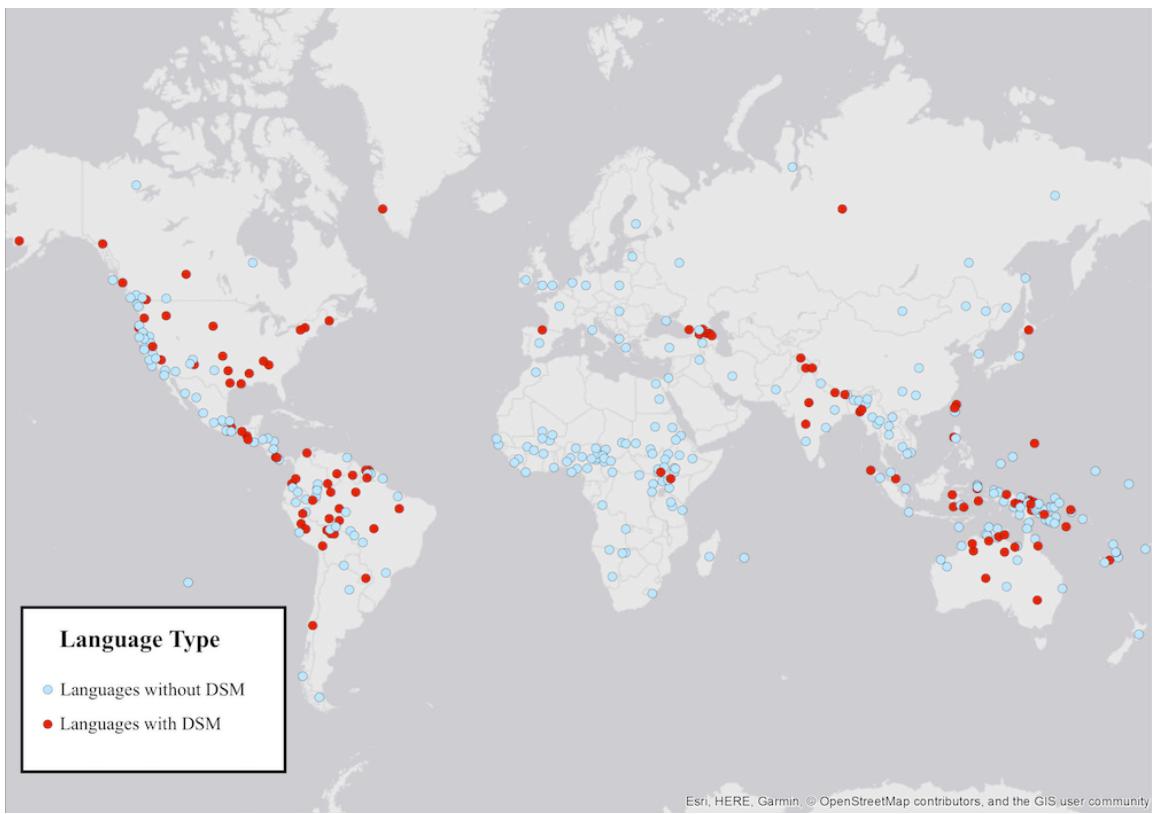


Figure 1.6: Incidence of Ergativity in the World's Languages (data from WALS online)

The map in Figure (1.6) depicts the geographical distribution of languages with ergative alignment based on data from the World Atlas of Language Structures (WALS).¹¹ Ergative languages are marked in red. Ergative alignment is found in

10. Throughout this dissertation, markedness values are color-coded. Red indicates a marked (in this case morphologically marked) form, and blue indicates an unmarked form. This illustrates that Neutral, Inactive, Nominative, and Absolutive are typically default, zero forms.

11. The data obtained from WALS online is an aggregate of data from the chapters “Alignment

about 35% of the 382 languages sampled, and is particularly concentrated among the languages of Australia, Papua New Guinea, South Asia, and the Americas. It is broadly a feature of Indo-Aryan languages of India and the Himalayas, of which Nepali is a member, and of many of the Tibeto-Burman languages of the Himalayas.

1.3.1 Alignment Splits

Most languages which exhibit ergative alignment do so for only part of the grammar. They may have ergative alignment in some transitive clauses and accusative alignment in others. For example, the case morphology may follow an ergative pattern while verbal cross-reference follows an accusative pattern. Or the case morphology may follow an ergative pattern for some clauses and an accusative pattern in others.

Alignment splits are typically conditioned by some property of the clause or its constituents. These are the common sources of alignment splits (Dixon 1994: 70):

- (1) Verbal Tense-Aspect-Mode (perfective/imperfective, past/nonpast, affirmative/negative)
- (2) Properties of the noun phrase referent (noun/pronoun, person, number, animacy, referentiality, definiteness)
- (3) Clausal properties (main/subordinate)

A common locus of the ergative split is based on aspect: ergative morphology is typically restricted to verbal forms with perfective aspect. This is the quintessential alignment pattern for many Indo-Aryan languages (Bangla is an exception, being nominative-accusative throughout). Hindi and Nepali are two languages with align-

of Case Marking in Full Noun Phrases” (Comrie 2013a), “Alignment of Case Marking in Pronouns” (Comrie 2013b), and “Alignment of Verbal Person Marking” (Siewierska 2013a). I employ a quite liberal definition of ergativity to include any language with ergative, active, or tripartite case-marking and any language with ergative, split, active, or hierarchical (inverse) verbal agreement, hence, any form of differential subject marking. While this clearly lumps together disparate grammatical phenomena, it provides a better visualization of the overall distribution of ergativity than can be seen from any one variable.

ment splits based on perfective aspect. This pattern is demonstrated for Hindi with the example below:

- (5) a. *rām-ne cidiyā dekh-ī*
ram.M-ERG sparrow.F.NOM see-PERF.F.SG
'Ram saw a sparrow.' (Deo and Sharma 2006: 376)
- b. *sita rām-ko dekh-t-ī h-ai*
sita.F.NOM ram.M-ACC see-IMPF-F.SG be-PRES.3.SG
'Sita sees Ram.' (Deo and Sharma 2006: 376)

In (5a), the verb form is perfective and agrees with the object. There is ergative case-marking (*-ne*) on the subject of the sentence, *Ram*. This is an ergative alignment pattern. In imperfective clauses like (5b), the alignment is nominative-accusative. The verb form is imperfective and agrees with the subject. There is accusative case-marking (*-ko*) on the object of the sentence, *Ram*. This is an accusative alignment pattern.

This is the “classic split” of modern Indo-Aryan languages: “quasi-ergative case-marking and agreement in the Perfective only, vs. nominative-accusative patterns in non-Perfective tenses”(Masica 1993: 342). Figure (1.7) is a schematization of the domain of ergativity in a language with such a split.

The term “alignment split” implies that ergative morphology and accusative morphology are never found in the same clause: we should only find ergative case in perfective (transitive) clauses and we should only find accusative case in imperfective (transitive) clauses (and nominative case elsewhere). In fact, accusative morphology may have independent conditions on emergence. Silverstein (1976: 123) refers to this as a distinction between **global splits**, in which ergative case is affected by the presence or absence of accusative case (or vice versa), and **local splits**, in which ergative case and accusative case are independently conditioned.

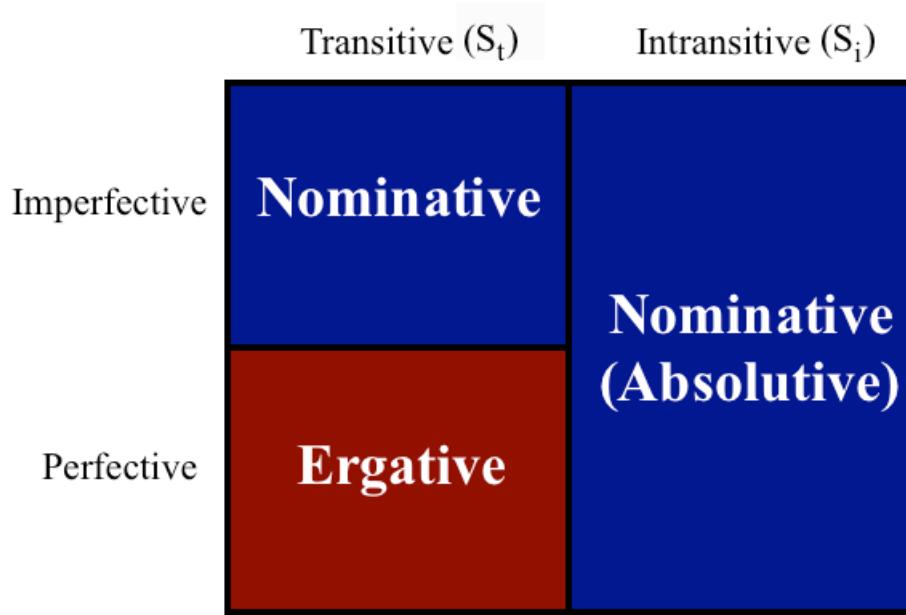


Figure 1.7: Ergative Case-Marking in a Language with a Perfective/Imperfective Split

1.4 Relevant Features of Nepali Grammar

This section is a description of Nepali nominal case morphology and verbal morphology. The major reference works consulted here include Grierson (1904a), Turner (1931), Acharya (1991), Masica (1993), and Schmidt (1993). Nepali is a non-tonal, head-final language with default SOV word order. Clause constituents are marked with case postpositions. Verbs are marked for tense-aspect-mood through successive suffixes, and there is consistent verbal cross-reference with the S_t or S_i argument in all tenses; verbal affixation marks person, number, gender, and honorificity. Verbal crossreference follows a straightforwardly nominative-accusative alignment, but there is a split accusative-ergative patterning in the expression of the S_t/S_i (as either NOM or ERG) and of the O (as either NOM or ACC).

1.4.1 Nominal Morphology and Case

Nouns are optionally marked for plurality with the suffix *-haru*. Gender in modern Nepali is restricted to a masculine/feminine distinction on a small set of animate nouns, as in *choro/chori* ‘son/daughter’.¹²

Case is expressed on common nouns by postpositions which do not inflect for person, number, or gender (Figure 1.8).¹³ The nominative case is the unmarked form. Many of the personal pronouns have suppletive forms, particularly in the ergative and genitive (Figure 1.9).¹⁴

Case	Form
Nominative	-∅
Accusative	- <i>lāi</i>
Dative	- <i>lāi</i>
Ergative	- <i>le</i>
Instrumental	- <i>le</i>
Genitive	- <i>ko</i>
Locative	- <i>mā</i>
Ablative	- <i>bāṭa</i>

Figure 1.8: Nepali Case Marking on Common Nouns

There is a syncretism between the ergative and instrumental case markers, as well as a syncretism between the accusative and dative.

12. Because the default case is masculine and feminine morphology is the marked case, I will not gloss masculine gender unless it is relevant to the discussion.

13. The exception to this rule is the genitive, which has a plural/oblique form *kā* and a singular feminine form *-ki*.

14. Here I have included only the mid-grade honorific forms of the 2nd and 3rd personal pronouns for simplicity. Grammars variously ascribe three to five levels of honorificity. Acharya (1991) has three levels: Low Grade Honorific: (2nd *tā*, 3rd *tyo/u*); Mid Grade Honorific (2nd *timī*, 3rd *unī*); High Grade Honorific (2nd *tapāī*, 3rd *wahā*). There are also particular forms for addressing the royal family.

	Nominative	Ergative	Accusative	Genitive
1st SG	<i>ma</i>	<i>maile</i>	<i>ma-lāi</i>	<i>mero</i>
1st PL	<i>hāmi</i>	<i>hāmi-le</i>	<i>hāmi-lāi</i>	<i>hāmro</i>
2nd SG mid	<i>timī</i>	<i>timī-le</i>	<i>timī-lāi</i>	<i>timro</i>
2nd PL mid	<i>timī-haru</i>	<i>timī-haru-le</i>	<i>timī-haru-lāi</i>	<i>timī-haru-ko</i>
3rd SG mid	<i>u</i>	<i>us-le</i>	<i>us-lāi</i>	<i>us-ko</i>
3rd PL mid	<i>uni-haru</i>	<i>uni-haru-le</i>	<i>uni-haru-lāi</i>	<i>uni-haru-ko</i>

Figure 1.9: Nepali Case Postpositions

The *-le* postposition

As a postposition on core arguments of the clause, *-le* marks **ergative case**. It is obligatory in perfective transitive clauses like (6). It varies with the nominative in imperfective transitive clauses (7), as well as in certain clauses which are arguably intransitive like (8).¹⁵

- (6) *bhāguta dekh-era sāno kheta-le sam-āu-na khoj-yo*
 frog see-CONJ little boy-ERG catch-CAUS-INF try-PERF.3.SG
 ‘Seeing the frog, the little boy tried to catch it.’ [AG]
- (7) *mero bahini/bahini-le kitāb lekh-chin.*
 my little.sister/little.sister-(ERG) book write-PRES.3.F.SG
 ‘My little sister writes books/ is writing a book / will write a book.’ [AG]
- (8) *kukur-haru/kukur-haru-le rāti bhak-tyo*
 dog-PL/DOG-PL-(ERG) night bark-HAB.3.SG
 ‘Dogs would bark at night.’ [TD]

It is homophonous with the **instrumental case** marker *-le*, which attaches to adjunct noun phrases which bring about effect the event described in the clause. This instrumental usage may co-occur in a clause with an ergative marker, as in the example below:

- (9) *maile camcā-le bhāt khā-ē*
 I.OBL-ERG spoon-INSTR rice eat-PERF.1.SG

15. I follow the convention of writing out both forms separated by a slash if variability is present. In some cases I emphasize the lack of variation by marking a form as ungrammatical or infelicitous. This convention makes it visually easier to represent alternations that involve oblique forms.

‘I ate rice with a spoon.’ [TD]

The *-le* postposition may also mark verbs that are either nonfinite or on verbs marked with the past participle *-eko*. These are subordinate clauses which give reasons or causes. I follow Butt and Poudel (2007) in considering these to be a form of instrumental marking.

- (10) a. *pāhunā āu-na-le ma timro bihā-mā jā-na pā-ina*
guests come-INF-INSTR I your wedding-LOC go-INF
get-PERF.1.SG.NEG
‘Because of guests’ coming, I could not go to your wedding.’ (Butt and
Poudel 2007: 10)
- b. *khās bhan-na-le ma-lāi thāhā chaina*
truly speak-INF-INSTR I-OBJ knowledge cop-3.SG.NEG
‘Truthfully, I don’t know.’

The *-le* postposition alternates with the accusative in certain modal constructions, particularly those that use the verb *pārnu* (“fall”/“need”):

- (11) a. *rām-lāi āp khā-nu par-cha*
ram-ACC mango eat-INF need-PRES.3.SG
‘Ram needs to eat mangoes.’ [AG]
- b. *rām-le āap khā-nu par-cha*
ram-ERG mango eat-INF need-PRES.3.SG
‘Ram likes to eat mangoes.’ [AG]

The *-lāi* postposition

The *-lāi* postposition marks the direct object of a transitive verb. It varies with nominative marking.

- (12) a. *Bibek-le birālo-(lāi) dekh-yo*
Bibek-ERG cat-(OBJ) see-PAST.3.SG
‘Bibek saw a cat.’ [TD]

Imperfective Forms	
simple present (PRES)	✓-cha
present continuous (CONT)	✓-dai-cha
past habitual (PST.HAB)	✓-thyo
archaic present (ARCH.PRES)	✓-dā-cha
definite future (DEF.FUT)	✓-ne cha
hypothetical future (HYP.FUT)	✓-lā
progressive (PROG)	✓-i-rah-eko cha
optative (OPT)	✓-os
Perfective Forms	
perfective (PERF)	✓-yo
present perfect (PRES.PERF)	✓-eko cha
past perfect (PST.PERF)	✓-eko thyo
present mirative (PRES.MIR)	✓-e-cha
past mirative (PST.MIR)	✓-e-thyo

Figure 1.10: Nepali Verb Forms

In a ditransitive verb with a direct object and an indirect object, the direct object is unmarked and the indirect object is obligatorily marked with *-lāi*, which I consider to be an expression of dative case.

- (13) *Bibek-le Timila-lāi phul di-yo*
 Bibek-ERG Timila-DAT flower give-PAST.3.SG
 ‘Bibek gave Timila a flower.’ (indirect object) [TD]

1.4.2 Verbal Morphology

Tense and aspect information and nominal cross-reference is marked via successive suffixes on the Nepali verb root. There are inflectional tenses, and there are periphrastic tenses which consist of a verb root marked with a participle and followed by a copula. Every tense has separate affirmative and negative declensions.

Verbal cross-reference is always with the S_t or S_i , whether it is marked in the ergative or the nominative. Verbs agree in person, number, gender, and honorificity, although number and gender features are often omitted.

Figure (1.10) lists the verb forms in third person mid-honorific singular form.

The forms can be broadly grouped together into perfective and imperfective based on semantic and morphological grounds. The **Perfective forms** share a number of properties. They denote perfective aspect and typically refer to events in the past. Many are periphrastic, and they all contain a reflex of the Sanskrit *yā*-construction (realized as *-yo/-e-* in Nepali). And for all these forms the **ergative case is obligatory in the transitive clauses.**

The imperfective tenses do not contain reflexes of the Sanskrit *-yā* construction (with the exception of the progressive). The Simple Present has pronominal affix markers which are similar in form to the present tense copula but have different inflections. The past habitual has a pronominal affix markers which are similar in form to the past tense copula. Other tenses include aspectual markers that attach to the verb root or periphrastic constructions. For all clauses in these tenses **ergative case is variable in transitive clauses.**

Copular Clauses

Nepali has two copulas in the present tense: *ho* and *cha*. Grammars of Nepali tend to characterize *ho* and *cha* as separate present-tense instantiations of the verb *hunu* ‘to be’, with *ho* described as ‘identificational’ and *cha* as ‘existential’ (Schmidt (1993), Acharya (1991: 154-155)). Butt and Poudel (2007) characterize *ho* as an **individual-level** copula and *cha* as a **stage-level** copula. I discuss this distinction in greater detail in section 4.1.3. The verb forms of the copular *hunu* are summarized in Figure (1.11) in third person mid-honorific singular form.

Many of the verb forms in the previous subsection historically derive from periphrastic constructions consisting of a verb root or deverbal adjective and a copula. Those forms which end with a full copula (definite future, progressive, present perfect, past perfect) are periphrastic in the modern language. Either *cha* or *ho* may be used,

Form	Gloss
<i>hunu</i>	nonfinite
<i>cha</i>	copula 1
<i>ho</i>	copula 2
<i>thiyo</i>	past (imperfective)
<i>bhayo</i>	perfective
<i>huncha</i>	future
	‘to be’
	‘is’
	‘is’
	‘was’
	‘became’
	‘will be’/‘tends to be’

Figure 1.11: Nepali Copulas

as well as past and future tense forms. Other forms are historically derived from periphrastic constructions with copulas (simple present, past habitual, archaic present, present and past mirative) but are inflectional suffixes in the modern language.¹⁶

1.5 Ergativity in Nepali

Nepali has an ergative pattern that is unique among Indo-Aryan languages. Verbal morphology always references the subject (S_t or S_i), whether the subject is case marked or not. So in this domain the alignment is straightforwardly nominative-accusative. In the domain of nominal morphology, the pattern is more complex. Figure (1.5) is a visualization of the domains of ergative marking based on properties of the verb. It is derived from the conclusions of Li (2007). Ergative marking on the S_t is required in perfective transitive clauses and variable in imperfective transitive clauses. In the intransitive domain, there is a split between agent-like subjects (Unergatives) and patient-like subjects (Unaccusatives). Ergative marking is disallowed on the S_i of unaccusative intransitives. For unergative intransitives, the pattern depends upon the lexical semantics of the particular verb. Depending upon the verb, all three patterns are possible: ergative marking may be obligatory, disallowed, or

16. For example, the forms with suffixes have a separate negative paradigm: compare *cha* ‘is’ and *chainā* ‘isn’t’; *gar-eko cha* ‘has done’ and *gar-eko chainā* ‘hasn’t done’; but *gar-cha* ‘does’ and *gar-daina* ‘does not do’.

variable.¹⁷

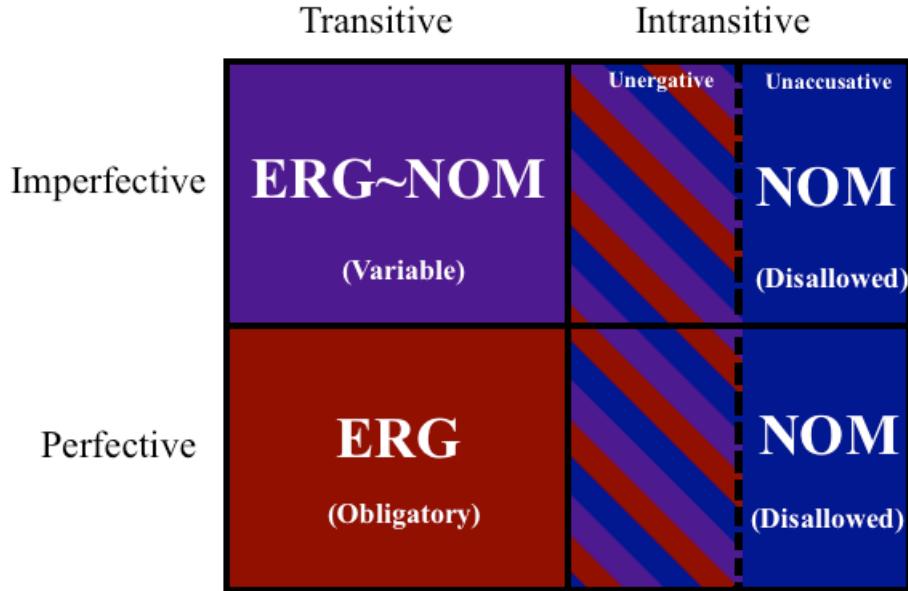


Figure 1.12: Ergative marking in Nepali (according to Li 2007)

An early account of the variable Nepali pattern comes from George Grierson's monumental *Linguistic Survey of India* (Grierson 1904a). This was a comprehensive survey of the languages of India undertaken by Grierson on behalf of the Government of India, a thirty-year project which documented 364 languages and dialects. On the unusual usage of the ergative (which he refers to as the Agent case), Grierson considers its unusual character to derive from contact with its neighbors:

In the fact that the verb is not changed by the object, we see the influence of Tibeto-Burman languages. The influence is still plainer in the colloquial language, which in this respect differs markedly from the literary style. In the colloquial language, the agent case may optionally be employed before any tense of a transitive verb whether derived from the past participle

17. The conclusion I reach in section 4.1.2 is simpler: I argue that ergative marking is completely disallowed in all intransitive clauses. I also discuss the overall variability ergative marking in subordinate clauses in section 4.1.1 and in copular clauses in 4.1.3.

or not, in fact it is more customary to employ it than to employ the nominative.

(Grierson 1904a: 26)

Grierson, who notes that the usage is “emphatic,” is describing variable ergative marking in Nepali, which is not generally a property of Indo-Aryan languages. Languages of this type are not very widespread, consisting of 10% of the morphologically ergative languages in McGregor (2010)’s language survey, and about 3% of all languages in Fauconnier (2011)’s language sample. But they are particularly common in the Tibeto-Burman languages of the Himalayan region (McGregor 2010, Chelliah et al. 2011). Language contact between Nepali and Newari in particular is implicated in morphological changes by Bendix (1974), Genetti (1994), and Verbeke (2011).¹⁸

1.5.1 The development of Ergativity in Indo-Aryan

The ergative pattern in Nepali fits into the broader milieu of ergative systems in the Indo-Aryan languages, which are inherited from ergative alignment shift that occurred in the history of the family. The most common historical explanation for the emergence of split-ergativity is that it arose from the reanalysis of a deverbal passive-like periphrastic construction in Sanskrit into a general perfective form. This construction was not the main form of producing passive constructions in Sanskrit, but the form was essentially a passive periphrastic construction with a clear perfective sense (Anderson 1977: 177). As in an ergative clause, the agreement was with the O and the S_t was marked by additional morphology.

During the development of the Old Indo-Aryan languages into the Middle Indo-Aryan languages, much of the tense and case morphology eroded or fell out of use completely. In Hindi and Nepali, the formerly inflected instrumental case collapsed into other cases to form a general oblique case, which has almost completely eroded

18. Meakins (2009) provides a particular example of shift in ergative alignment due to language contact.

into the nominative in modern Nepali. Older tenses fell out of use, such that the constructed form above became the only way to express perfective aspect in MIA. The periphrastic construction was reanalyzed as an active, perfective tense with ergative morphology.

The result of this reanalysis was that perfectives developed ergative morphology. For many of the modern Indo-Aryan languages there is a perfective split in both verbal agreement and nominal morphology.¹⁹ There have been various developments across the Indo-Aryan family that temper with this system, either with the agreement pattern or the form of the ergative (Masica 1993: 341-345).

In Nepali, Asamiya, Shina and Bangla there was a shift from object agreement to subject agreement in the perfective domain. In Bangla, the ergative case marking has completely eroded and there is no ergative case-marking at all. Indo-Aryan languages like Kashmiri retain vestiges of the old Sanskrit form which inflects for gender and number. Other languages like Asamiya retain a phonological descendent of the marker (-e) that does not inflect for gender and number. Hindi, Punjabi, Marathi, and Nepali all reinforced the ergative case with a non-inflectional instrumental postposition (*ne* in Hindi, Punjabi, and Marathi, and *le* in Nepali). For Marathi and Punjabi, the ergative and nominative have collapsed into the same form for first and second person pronominal forms (Deo and Sharma 2006).

Wallace (1982, 1985) chronicles the development of ergative alignment in Nepali over multiple periods of its history. Ergative marking was reintroduced into the language in two waves as new verb forms were created from participials (hence the similar form of the present mirative and present perfective forms). However, Poudel

19. Butt (2001) provides an alternative viewpoint to this sequence of events, which is that OIA and its modern descendants “all used (and continue to use) a complex system of case marking that includes non-nominative marking on subjects and case alternations to express consistent semantic differences” (Butt 2001: 106). Under this view, there was never a grand structural ergative realignment throughout the language, because case marking and verb agreement alternations have always been used to impart particular semantic meanings like agentivity.

(2008) disputes certain aspects of this analysis, particularly in assigning earlier dates to the origin of Nepali and the creation of case marking postpositions.

The historical developments that led to the modern-day Nepali are significant because they created grammatical associations between particular verb forms and ergative case marking. These associations represent the categorical boundaries of obligatory/disallowed marking that delimit regions of variable marking. However, the primary focus of this investigation will be on the synchronic system. I now turn to a discussion of the methodologies I employ in data collection.

Chapter 2

Methodologies

The factors which condition the expression of case in Nepali are varied and correlate with subtle properties of the discourse. This requires approaching the problem of data collection from multiple angles. The observations and conclusions of this study are based upon four converging lines of inquiry. The first line of inquiry consists of descriptive and theoretical accounts of ergative patterning in the **literature**. This includes descriptions of the Nepali language in grammars and dictionaries of Nepali (in English and Nepali), linguistic descriptions of the Nepali ergative system, and theoretical literature on split-ergativity, differential and optional case marking, semantic markedness, transitivity, and discourse prominence. In the Theories chapter I summarize these proposals and discuss their advantages and deficiencies. Secondly, I conducted **elicitation sessions** with native speakers of Nepali. From 2013 to 2019 I worked with a total of thirteen speakers, collecting judgments on various aspects of the Nepali case system. In the Elicitation section below I go into more detail about the process and discuss the general elicitation procedure I developed. Third, I conducted a **grammaticality judgment survey** at Tribhuvan University at Kirtipur in Kathmandu, Nepal. The goal of this survey was to explore a variety of topics that had arisen from elicitation sessions, as described in the Survey section. Finally,

I annotated and analyzed a sample of the Nepali National Spoken Corpus (NNSP) which was produced by Nelralec in 2006. I annotated 67 minutes of spoken Nepali, focusing on the realization of core arguments in various clause types. I discuss this procedure in detail in the Corpus section. The results of this research are represented throughout the Observations chapter.

It may not be immediately obvious why it is necessary to examine this problem from multiple perspectives simultaneously. This multifaceted research plan arose organically from difficulties I encountered studying the problem. When I began working with consultants, I quickly discovered that judgments were subtle and heavily dependent upon the discourse context. In many specific instances consultants could not detect any difference of meaning between a minimal pair, and judgments varied both from speaker to speaker and from session to session.

I developed the survey in order to see whether the intuitions expressed by consultants in elicitation sessions bear out in the judgments of other Nepali speakers. This also allowed me to better control the discourse context. The judgments consisted of question-answer pairs, which allowed for the question under discussion to be explicit. While the survey provided useful data (particularly about judgments relating to semantic factors), it was less successful in determining judgments relating to more subtle pragmatic factors.

On the suggestion of survey respondents and linguists at Tribhuvan University, I began an analysis of naturally-occurring speech to examine the judgments and intuitions shared by elicitation consultants and survey respondents. This allowed me to make quantitative claims about the expression of case marking in particular grammatical and semantic contexts. Corpus analysis also fed directly into the elicitation research. I examined particular discourses with consultants and discussed intuitions about the expression of case in particular contexts. In this way the results obtained from the corpus analysis, survey results, and elicitation sessions informed each other

and allowed for a more precise understanding of case marking within a broad range of grammatical and discourse contexts.

2.1 The Necessity of a Multifaceted Approach

Other researchers of pragmatically-conditioned ergativity have noted the difficulty of collecting data through speaker elicitation. For example, Holisky (1987) writes the following in her analysis of Tsova-Tush: “As I worked with my consultant, verbs moved back and forth from group to group; the lists could not be verified completely, because they kept changing. Had I had a different consultant, asked about these words on a different day or in a different way, when she was in a different frame of mind, the lists would be somewhat different” (Holisky 1987: 114).

In discussing ergative/nominative alternations in Nepali, Butt and Poudel (2007) write that when “questioned about the semantic difference between minimal pairs as in (9), native speakers of Nepali tend to look at you quite helplessly - a difference is felt, but what that difference is cannot usually be articulated” (Butt and Poudel 2007: 4). I suspect that many native English speakers will exhibit a fair approximation of that same feeling of helplessness when questioned about the meaning difference involved with a dative shift alternation:

- (14) a. I threw Shauna the ball.
- b. I threw the ball to Shauna.

Dative shift alternations and Nepali ergative/nominative alternations share some similarities. They both consist of variant strategies of argument realization that correlate with a host of subtle pragmatic and semantic effects which present challenges for data collection and generalization (cf. Green 1974, Krifka 1999, Bresnan et al. 2007).

Pragmatically-conditioned ergativity is unusual in the Indo-Aryan family, for which the verb form and properties of the argument will categorically determine the morphological expression of case.¹ But it is common enough in the Tibeto-Burman languages of the Himalayas. In their introduction to a collection of essays on Optional Ergative Marking in Tibeto-Burman, Chelliah et al. (2011) write that “native speakers cannot easily explain the use of the agentive marker because of the number of related and nuanced implications its use has” (Chelliah et al. 2011: 6). This sentiment is echoed by DeLancey (2011) and Lidz et al. (2011) in the same volume. They both note that direct elicitation has led to misinterpretations of the ergative pattern, partly because of influence from the contact language. Based on these elicitation results, researchers typically give description of the language which seem straightforwardly split-ergative with a split conditioned by perfective aspect. But in fact this is not how people use these markers when speaking.

Chelliah et al. (2011) and Willis et al. (2011) argue that the best way to study Optional Ergative Marking is to combine direct elicitation with analysis of natural speech (e.g. through participant observation) and grammaticality judgments. Chelliah et al. (2011) write that the “researcher must elicit these constructions indirectly through native speaker explanations of culturally-bound situations and relationships” (Chelliah et al. 2011: 6).

I have attempted to follow this advice. In the next subsections I describe the approaches I took to these three methodologies in detail, and then I compare the advantages and limitations of each.

1. There are some exceptions, such as the nominative/ergative alternations in certain intransitive verbs of Hindi (discussed in the Theories section).

2.2 Elicitations

I worked with thirteen total Nepali speakers from 2013 to 2019. I spent an extended period of time with eight of these speakers: Anobha Gurung, Timila Dhakhwa, Rijan Maharjan, Biplob Acharya, Sajju Tamang, Geeta Manandhar, Bibek Basnet, and Sushant Banjara. All of these speakers were born and raised in Nepal and speak it natively, although many of them currently live in the United States. Most of the speakers who were born and raised in Kathmandu Valley (Province 3), but I interviewed speakers from other regions to get a basic understanding of the way that these patterns may vary geographically. Figure (2.1) tabulates the demographic data for every consultant. Throughout this dissertation, I will refer to these consultants by their initials given in this chart.

Name	ID	Sex	Birth Region	Residence	Native Language	Parents' L1
Biplob Acharya	BA	Male	Province 3	Province 3	Nepali	Nepali
Sushant Banjara	SB	Male	Province 1	USA	Nepali	Nepali
Bibek Basnet	BB	Male	Province 1	USA	Nepali	Nepali
Timila Dhakhwa	TD	Female	Province 3	USA	Nepali	Nepali, Newari
Anobha Gurung	AG	Female	Province 3	USA	Nepali	Gurung, Nepali
Min Gurung	MG	Male	Gandaki	Gandaki	Gurung, Nepali	Gurung
Prashanta Kharel	PK	Male	Province 3	USA	Nepali	Nepali, Newari
Rijan Maharjan	RM	Male	Province 3	USA	Nepali, Newari	Nepali, Newari
Geeta Manandhar	GM	Female	Province 3	Province 3	Nepali, Newari	Newari
Sabin Pyakurel	SP	Male	Province 3	Province 3	Nepali	Nepali
Uddhab Pyakurel	UP	Male	Province 3	Province 3	Nepali	Nepali
Kamal Sharma	KS	Male	Province 1	USA	Nepali	Nepali
Sajju Tamang	ST	Female	Province 3	Province 3	Nepali	Nepali

Figure 2.1: Nepali-speaking Elicitation Consultants

I began working with Anobha Gurung and Prashanta Tamang while they were graduate students at Yale University, initially for an Indo-Aryan linguistics seminar taught by Ashwini Deo in the Fall of 2013 and with a directed independent study on Nepali in the Spring of 2014. I subsequently applied to the Yale DILS (Directed Independent Language Study) program, and worked with Yale graduate students Timila Dhakhwa (Spring of 2015, on Nepali) and Rijan Maharjan (Fall of 2015, on Nepali and Patan dialect Newari). I began meeting with another Yale graduate

student, Bibek Basnet, beginning in the Fall of 2017. I applied for another DILS specifically to work on a corpus analysis in Fall 2018. For this project I worked with another Yale graduate student, Sushant Banjara.

These interviews were supplemented by two field visits to Kathmandu, in 2014 and 2016. In July and August of 2014, I did fieldwork with Dr. Ashwini Deo in India. During the latter part of August, I traveled to Nepal to Lalitpur District where I had taught as a Fulbright scholar in 2010-2011, and spoke with Uddhab Pyakurel and Sabin Pyakurel. From June to September of 2016, I did fieldwork in Kathmandu, which included interviews with two graduate students, Biplob Acharya and Sajju Tamang, and targeted Nepali and Kathmandu dialect Newari tutoring from Geeta Manandhar. I also discussed my research project with Dr. Madhav Pokharel and Dr. Dan Raj Regmi at Tribhuvan University and Dr. Tikaram Poudel at Kathmandu University.

2.2.1 Procedure

For a typical elicitation session, I would begin with a list of hypotheses about the effect of the ergative marker in a particular context. For each hypothesis, I would devise a minimal pair of sentences with the ergative and without. I would describe a general situation in which the sentence might be uttered by the speaker, and then I would ask the speaker to respond to the situation in Nepali. I would record the speaker's response. I would ask whether the other forms were grammatically possible, the extent to which they sounded awkward in this context, and how the form might be better in a different context.

For example, the simple present verb form (PRES) may be used to describe events which are habitual, ongoing, or future-oriented. I was interested in whether the ergative form of the subject would be preferred or dispreferred with future-oriented readings of the simple present verb form. I presented the following context to PK: he

and I are both eating at a cafe and I accidentally knock my drink onto the floor. I start to clean up the spill myself, but the waiter comes by with a towel and says he'll take care of it. I asked PK what specifically the waiter would be likely to say, and he responded with (15a):

- (15) a. *ma puc-chu* !
I clean-PRES.3.SG !
“I’ll clean (it)!” [PK]
- b. *maile puc-chu* !
I.ERG clean-PRES.3.SG !
“I’ll clean (it)!” [PK]

PK noted that (15b), on the other hand, sounded unnatural to him in that situation. I asked about other ways that the sentiment might be expressed and PK offered a sentence with the hypothetical future verb form: *maile pucchāula!* “I shall clean it!” He preferred the ergative over the nominative with the hypothetical future in this case. However, he also noted that the hypothetical future form sounds stilted in the precise situation described, in which the waiter is speaking quickly but politely to stop me from getting my hands dirty.

Compare this method to a direct elicitation method in which the researcher simply requests a Nepali translation of the English sentence “I will clean it.” Without a context, there are several ways to describe a future-oriented event and multiple possible verb forms in Nepali that can all express the English sentence in a particular context. The response may contain an overt subject or it may be elided. If the subject is overt, the researcher might record an ergative marker on the subject but does not know whether it is obligatory or varies with the nominative.

Alternatively, we can compare this to a method in which the researcher presents two minimal pairs, *ma pucchu* and *maile pucchu*, and asks which one sounds more natural. Here we might learn that the ergative varies with the nominative for most speakers. The most common response I received in comparing minimal paired sen-

tences was “They both are correct.” But divorced from a particular context it is impossible to tell what, if anything, is contributed by the ergative.

Furthermore, giving the speaker the opportunity to think about and describe the implications of using one form over another provides valuable qualitative data. Speakers are much more likely to have an opinion about the difference in meaning when given a particular context. And while speakers do not have a perfect knowledge of their own grammar and may sometimes make incorrect assumptions, their intuitions are a crucial source of insights that can be tested in later elicitation sessions.

When a statement is a response to a direct question, many speakers expressed the intuition that the pattern of case marking in the response follows the question: an ergative in the question means an ergative in the answer. This led me to frame the topic of a particular statement in terms of a question under discussion, which informed my choice of discourse model (as discussed in the Theories chapter in the section on Discourse Prominence).

With four consultants (AG, UP, SP, and GM), I also elicited narratives using storyboards. These are picture books that consultants are asked to narrate in their own words (Burton and Matthewson 2015). For storyboards I used *A Boy, a Dog, and a Frog* (Mayer 2003) and the Woodchopper narrative from Totem Field Storyboards (TFS Working Group 2011). While I found this technique useful for eliciting naturalistic data and example contexts, I found narratives to be less useful as a source of data than the recorded conversations from the NNSP. This is largely because speakers tend to use the past tense to narrate the events of the story, in which the most common verb forms were perfective and ergative marking is obligatory.

2.2.2 Advantages and Limitations of Elicitation

The elicitation of grammaticality judgments allows the researcher to target specific hypotheses. It also gives the researcher the greatest amount of control in manipulating

discourse context by describing the scene or setting up an explicit question under discussion. The one-on-one approach is ideal for developing new hypotheses, because conversation about the nuances of meaning in a particular context naturally leads to new insights and intuitions from the consultant. Grammaticality judgments are also a source of negative data, i.e., ungrammatical utterances, which cannot be found by restricting data collection to naturally-occurring speech.

However, working with a small number of consultants makes it less possible to uncover patterns of geographic and/or social variation. This is particularly problematic for elicitation of pragmatic phenomena, in which speakers may already have difficulty judging differences in meaning that do not necessarily correlate with grammatical-ity but rather with strategies of information packaging. My consultants are mostly university-educated and under the age of thirty, because I worked extensively with graduate students at Yale University, and this may introduce a bias in the results (I discuss this in the final section of this chapter). Secondly, there is the question of whether elicitation judgments accurately represent the way speakers use language in practice. Elicitation must be supplemented with analysis of naturally-occurring speech and judgments should be corroborated by looking at a wider pool of speakers.

2.3 Grammaticality Judgment Survey

During the summer of 2016, I had the opportunity to conduct research on Nepali and Newari in Kathmandu, Nepal. The United States Educational Foundation Nepal (USEF-Nepal) connected me with two student researchers (BA and SJ) who were tremendously helpful to my research. In addition to providing me with their own judgments, they assisted me in developing, translating, and disseminating a grammatical-ity judgment survey to students at the Department of Education at Tribhuvan University in Kirtipur.

This Kathmandu Survey was intended as a pilot survey to explore the expression of ergativity in the imperfective domain of Nepali. The survey consisted of 58 items presented in written Nepali. Respondents were prompted to judge the grammaticality of a statement along a five-point Likert scale. This methodology and design was inspired by Meyerhoff (2008)'s usage of grammaticality judgment surveys in the field. Respondents were given the following instructions:

“नेपालीहरु आफ्नो दैनिक जीवन मा कसरी बोल्छन भनेर हामिले यो सर्वेक्षण गरदैछौं। तपाईंहरुले यहाँ दिएको नम्बरहरु लाई एक देखि पाँच सम्म कुनै एउटा वाक्य मा ‘tick’ लगाउनु होला । ‘५’ भनेको तपाईं आफैले बोलने गर्नुभएको व यो शब्दहरु अरु कसै बाट सुनु भएको । ‘१’ भनेको तपाईंले यो शब्द आफ्नो दैनिक जीवन मा कहिले सुन्नु भएको छैन । हामी व्याकरण शब्दमा चाहिं केहि रुचि राख्दैनौं तर के हामीले बोलेको कुरा स्वाभाविक सुनिन्छ ?”

“We are conducting this survey on how Nepalis speak in their daily lives. Please tick one of the numbers between one and five on each question. ‘5’ means that you use sentences like this and have heard them spoken by others. ‘1’ means that you have never heard a sentence like this in your daily life. We are not looking for what is grammatically correct but rather what sounds natural.”

Each item consisted of question-answer pairs, for which a question is followed by two possible responses that differ only in whether the subject is nominative or ergative. The overt question allowed there to be some control over the context of utterance so that statements are not simply presented in isolation. Figure (2.2) depicts Item C3 as it appeared on the Survey and Figure (2.3) is an English translation of this item. The survey consisted of eleven sections, each of which was designed to test a particular theory about ergative marking. These theories will be described in greater detail

सुर्य साका हारि हो कि?

a. होइन सुर्य मासु खान्छ ।

१ २

३

४

५

b. होइन सुर्यले मासु खान्छ ।

१ २

३

४

५

Figure 2.2: Kathmandu Survey: Item C3

Surya is a vegetarian, right?

a. No, Surya eats meat.

१ २

३

४

५

b. No, Surya-ERG eats meat.

१ २

३

४

५

Figure 2.3: Kathmandu Survey: Item C3 (English Translation)

in the Theory section.

A: Inanimate Subjects - The responses in this section consisted of transitive imperfective clauses in which the S_t has inanimate reference. The prediction is that the nominative form should be judged ungrammatical.

B: Focus - Here the responses consisted of transitive imperfective clauses in which either the subject or object was focalized in the given the context, examining the extent to which the ergative form is preferred when the subject is focalized.

C: Individual-Level Predication - This section consisted of six items which were paired. Between a pair, the questions differed but the responses were the same. The questions either set up a context in which the response has a stage-level or individual-level interpretation.

D: Elided Objects - In this section the object was elided in each response to examine the argument that ergative marking should be preferred in such cases.

E: Aboutness - To examine whether ergative marking must be associated with topicality, the questions in this section were all of the type “What about X?”

F: Unergative Intransitives - The responses in this section were unergative intransitive clauses for which ergative marking is predicted to occur.

G: Generic Readings - The responses in this section consisted of generic statements about kinds. The prediction is that the ergative is preferred.

H: Sets - In this section the subject referents were picked out of a set of possible entities introduced by the question, either through quantificational determiners or definite descriptions. I expected ergative marking to be preferred here.

I: Thetic/Categorical Propositions - This section consisted of six items which were paired. Between a pair, the questions differed but the responses were the same. The questions either set up a context in which the response has a thetic or categorical interpretation.

J: Copulas - The responses in this section were copular clauses, in which ergative marking is predicted to be disallowed.

K: Unaccusative Intransitives - The responses in this section were unaccusative intransitive clauses, in which ergative marking is predicted to be disallowed.

Each section contained five or six items each (C and I, which consisted of double question pairs, had six items each). The order of item presentation was made in the

following way: I created two survey forms with different orderings, and gave half the respondents one survey form and the other half the other.

For each survey form, the order of the response presented as (a) or (b) was randomized for each item (that is, whether the ergative or nominative form was presented first in a particular item). The items themselves were also presented in a pseudo-random order: I split the survey into five trials, putting one item from each section A-K into each trial, and randomizing the order of items within the trial.²

2.3.1 Survey Respondents

I collected 28 responses (17 male, 10 female, ages 20-56) from students and professors in the Education and Linguistics Departments at Tribhuvan University in Kathmandu. The survey forms requested information on birth district, current district of residence, age, sex, languages the respondent speaks, and languages the respondents' parents speak. Figure (2.4) tabulates this data for each respondent.

I recorded the responses for each question into an R dataframe, and associated each response with the speaker, trial and survey form, as well as other relevant information such as the particular subject type and object type in the question, the verb form and whether or not honorifics were used. I also recorded notes that were made by respondents at the end of the survey.

2.3.2 Advantages and Limitations of the Kathmandu Survey

The use of surveys allows the collection of judgments from a larger number of native speakers, and the results are quantitative so they can be examined for particular correlations. Furthermore, the question-answer methodology gives the researcher some control over the discourse context, and eliciting judgments for both nominative and

2. Randomization was determined by coin flip due to my limited access to Internet and electricity at the time.

ID	Birth Region	Residence	Age	Sex	Languages	Parent Languages
1	Province 5	Province 3	26	Male	Nepali	Nepali
2	Province 3	Province 3	25	Male	Nepali	Nepali
3	Gandaki	Province 3	48	Male	Nepali	Nepali
4	Gandaki	Province 3	33	Female	Nepali, English, Hindi	Nepali
5	Province 5	Province 3	26	Male	Tharu	Tharu
6	Gandaki	Province 3	28	Female	Nepali	Nepali
7	Karnali	Province 3	45	Male	Nepali, English, Hindi	Nepali
8	Province 3	Province 3	41	Male	Nepali, English	Nepali
9	Province 3	Province 3	20	Male	Nepali, English	Nepali, English
10	Province 3	Province 3	31	Female	Nepali, English	Nepali
11	Province 3	Province 3	26	Male	Nepali, Tamang	Nepali, Tamang
12	Province 3	Province 3	24	Female	Nepali	Nepali
13	Province 3	Province 3	50	Male	Nepali	Nepali
14	Province 3	Province 3	26	Female	Nepali	Nepali
15	Province 3	Province 3	56	Male	Nepali, Hindi	Nepali, Magar
16	Province 3	Province 3	26	Female	Nepali	Nepali
17	Province 3	Province 3	25	Female	Nepali	Nepali
18	Province 1	Province 3	46	Male	Nepali, English, Hindi, Dhimal	Nepali, Hindi
19	Province 1	Province 3	35	Female	Nepali	Nepali
20	Province 3	Province 3	49	Male	Nepali, English	Nepali
21	Province 3	Province 3	25	Male	Nepali	Nepali
22	Sudurpaschim	Province 3	38	Male	Baitedeli, Doteli, Nepali	Baitedeli, Doteli
23	Province 3	Province 3	27	Female	Nepali	Nepali
24	Province 3	Province 3	24	Female	Nepali	Nepali
25	Sudurpaschim	Province 3	26	Male	Baitadel, Doteli	Baitadel
26	Province 1	Province 3	53	Male	Nepali, Nangali, Hindi	Nepali
27	Province 5	Province 3	46	Male	Nepali, English	Nepali
28	Province 3	Province 3	unknown	Female	Newari, Nepali Hindi, English	Newari, Nepali

Figure 2.4: Kathmandu Survey Respondent Summary

ergative minimal pairs gives both positive and negative evidence for a correlation. If both judgments are consistently ranked low, then this can be an indication that there is a problem with the trial set-up (i.e., grammatical mistakes or unnatural contexts).

This type of survey is helpful for studying semantic phenomena and categorical distinctions related to the interaction between morphology and case expression. For example, the judgments were quite distinct in picking out the relationship between ergative marking and inanimate subjects, copular clauses, and perfective aspect. But the survey was not ideal for testing pragmatic phenomena like topicality, categorical propositions, and individual-level predicates. The results here were fairly inconclusive. This is partly because these are not issues of grammaticality but rather of usage; both the nominative and ergative form are grammatical, but one might be used to emphasize a particular aspect of discourse. Respondents generally ranked both forms

highly. However, this is a significant finding in itself. The lack of a strong correlation between, topicality, categorical propositions, and individual-level predicates tells us that the correlations that have been made in the literature are not categorical but rather represent gradient tendencies.

There is once again a bias toward young university-educated students, and two respondents (22 and 25) were not native speakers of Nepali but of related Eastern Pahari languages. There is also more diversity in the birth region of the respondents. I address dialect diversity in the final subsection of this chapter.

Comments that respondents gave on the survey design, as well as discussions with linguists at Tribhuvan University (and later with SB), exposed some weaknesses in the survey design. There were typographical mistakes in the written survey (particularly in terms of the usage of commas and conventions regarding the spacing of orthographic words). More importantly, a few respondents suggested that some of the question-response pairs did not sound natural. This is partly because speakers will tend to omit the subject if it is heavily topicalized, but I needed to have overt subjects in each response. There is also an issue of question-response, in which the case form on the subject of the question tends to be followed in the response. SB also noted that the purpose of the survey was to examine spoken conventions, but the written responses did not contain the discourse particles and “ornamentation” that would be representative of spoken Nepali. Overall the results of the Kathmandu Survey should be regarded with more skepticism than results obtained from the other methodologies, but in general they provide extra corroborating evidence in the Observation chapter.

Respondents suggested that I supplement my study with naturally-occurring data from narratives and conversations to get a better picture of the usage of the ergative. This led me to include a corpus analysis.

2.4 Corpus Analysis

2.4.1 The NNSP Corpus Sample

The Nepali National Spoken Corpus (NNSP) is a collection of recordings and transcriptions of spoken Nepali produced in 2006 by Nelralec (Nepali Language Resources and Localization for Education and Communication). It is maintained by the European Languages Resources Association (ELRA). The NNSP corpus contains 31 hours of material from 115 recordings of conversations in a variety of social settings. The corpus is broadly categorized by seventeen types of conversational activities. These include shopping and bargaining for clothes, work conversations, radio interviews, fortune-telling, medical visits, and thesis defenses (Yadava et al. 2008).

The corpus material consists of the audio files from these recordings, transcriptions, and participant metadata. The transcriptions are written phonetically in the Devanagari script. Each speaker turn is associated with its speaker ID code, and slashes separate each turn into utterances. Interruptions due to background noises are marked, as are relevant speaker gestures, and instances of overlapping speech. Figure (2.5) is a section of the annotation from interview V001001001.

The data for the analysis presented here comes from a sample of the NNSP corpus. I analyzed four interviews of 67 minutes of total dialogue between fourteen total speakers representing 2845 separate clauses. I annotated each clause by clause type, verb type, and properties of the core arguments (I will discuss that process in detail in the following subsection).

The four interviews came from the first two activity categories: Shopping and Discussion. The first interview (**V001001001**) consisted of a conversation between a salesperson and two customers bargaining over the price of a sweater in a market in Kathmandu. The second interview (**V001001004**) consisted of a conversation between a salesperson and two customers bargaining over various clothing items in a

\$B: नआउने भएर मात्र हो / नभए मैले देर पाठाइदिन्थैं / घाटा लगाउन मिल्दैन हैन / नाफा नखान मिल्छ घाटा खान मिल्दैन
 \$C: हैन मिल्दैन मिल्दैन
 \$A: खोल्न मन लाएन <1>1
 @ <1 पृष्ठभूमिबाट>1
 \$B: घाटा खान मिल्दैन / नाफा नलिन चैं मिल्छ / तर्पैले मलाइ अब घाटै लगाउने सुर गर्नु भो
 \$C: हैन म ह्वाँ आइरने / देख्नु भो नि तपाइँ
 \$B: देख्ने
 \$C: हेरै कुरा गर्नु म
 \$B: देख्ने मैले
 \$C: तेङ्ग भएर
 \$B: मैले देख्ने

Figure 2.5: NNSP Annotation Example (from interview V001001001)

market in Kathmandu. The third interview (**V001002003**) is a conversation between two academics discussing language standardization and their opinions on the state of the modern Nepali language. The fourth interview (**V001002005**) consisted of an outdoor conversation between five participants, mostly employees at a wildlife resort in the Terai region who take tourists on jungle walks to see rhinoceroses and tigers.

I concentrated on the first two activity categories because they consisted of wide-ranging discussions of ongoing events with fairly little narration of past events. This is important because I am interested in the way that core arguments are marked in imperfective clauses, and so the focus of my analysis was on dialogue which discussed ongoing or future-oriented events. I also wanted to avoid certain categories for practical reasons: the fortune-telling, radio/television interview, and thesis defense sections are likely to be somewhat complicated by the presence of stylized registers and terminology. My initial goal was to completely annotate every interview in these first two categories (twenty interviews, about 298 minutes total), but due to time

constraints the analysis is based upon the four interviews listed here.

Interview	ID	Age	Sex	L1	L2	Province	District	Clauses
V001001001	M1	25	Male	Nepali		Gandaki	Syangja	158
V001001001	M2	20	Male	Nepali		Province 2	Rautahat	276
V001001001	M3	25	Male	Nepali		Province 1	Jhapa	212
V001001004	F4	22	Female	Nepali	English/Hindi	Province 3	Kathmandu	260
V001001004	M7	25	Male	Nepali (Pacchami)		Province 5	Palpa	795
V001001004	M9	21	Male	Nepali		Province 3	Kathmandu	199
V001002003	M13	65	Male	Nepali (Purbali)		Gandaki	Kaski	189
V001002003	M14	64	Male	Dotyali		Sudurpaschim	Dadheldura	114
V001002005	M7	25	Male	Nepali (Pacchami)		Province 5	Palpa	795
V001002005	M24	43	Male	Nepali	English	Province 3	Chitwan	373
V001002005	M25	NA	Male	Nepali		NA	NA	1
V001002005	M26	38	Male	Tharu	Nepali	Province 3	Chitwan	222
V001002005	M27	NA	Male	NA		NA	NA	2

Figure 2.6: Interviewer Data from the NNSP Sample Interviews

The fourteen speakers represented in these interviews were identified by ID number. The speaker demographics are tabulated in Figure (2.6). I have changed the stated region of birth to conform to the modern provinces. One speaker (M7) was present in two interviews (V001001004 and V001002005). Two speakers (M25 and M27) only contributed a few utterances in the interviews and their demographic information was not recorded. Two speakers (M7 and M13) listed Nepali as their L1 (listed as *mātri bhāshā* ‘mother tongue’) and also specified a particular Nepali dialect (Purbali and Pacchami). There were two speakers who were not L1 speakers of Nepali: M14 listed another Eastern Pahari language (Dotyali) and M26 listed Tharu (a Central Zone Indo-Aryan language spoken primarily in the Terai).

2.4.2 Annotation Procedures

I began work on the NNSP sample in the Spring of 2017. I worked primarily with BB during the Fall of 2017 and Spring of 2018. To get a basic understanding of nominative/ergative alternations in the corpus, I searched through the entire Discussion section and looked at examples of pronouns with and without ergative marking. I focused on pronouns, and particularly first and second person pronouns, because they represent a closed set. Thus it was easier to compare the proportion of nominative

and ergative forms.

I would then examine the transcription of the recording in which a particular utterance takes place, and I would discuss the context of the conversation with BB: the ten utterances preceding and following the clause of interest. Here is a shortened example from interview V001002005. In this section of the interview M7 and M24 are discussing their business expenses. Here I have tried to preserve the phonetic transcription, with shortened forms of words, as they were written in Devanagari:

- (16) a. *yā̄̄ tapi tāpi-lā̄̄ talab-dekhi bāek ta kei*
here you.HON you.HON-DAT salary-ABL apart FOC something
di-daina ?
give-PRES.3.SG.PL ?

M7: ‘So here they don’t give you anything apart from salary?’

- b. *m tei ho tyo ta*
hm like.that COP that FOC

M24: ‘Yes, that’s how it is.’

- c. *pheri kasto huncha bhane ma jyādā hin-ne kām*
again how COP.FUT.3.SG say.NON.FIN I more walk-NON.FIN work
par-cha hin-ne kām par-epachi kharca jyādā
need-PRES.3.SG walk-NON.FIN work need-AFTER expense more
bad-cha ni tyo ta
increase-PRES.3.SG PRT that FOC

M24: ‘So this is how it is. I have to do traveling (lit. walking) work, and when you have to do traveling work the expenses increase.’

- d. *kharca bad-cha*
expense increase-PRES.3.SG

M7: ‘Expenses increase.’

- e. *m jasto etā khā-yo utā khā-yo*
hm like this.direction eat-PERF.3.SG that.direction eat-PERF.3.SG
utā gayo
that.direction go.PERF.3.SG

M24: ‘Yes, like: eat here, eat there, go there.’

- f. *ā̄̄*
ah

M7: ‘Ah.’

- g. *tesai-le* *gar-dā*
like.this.EMPH-ERG do-WHILE

M24: ‘That is why.’

- h. *mero ta* *yā* *aile kharca* *hu-ni* *bhan-eko* *din-mā* *tin*
my FOC here now expense COP.NON.FIN say-PERF day-LOC three
se *rupyā jati-ko* *phon gar-chu* ***maile***
hundred rupee how.much-GEN phone do-PRES.1.SG I.ERG

M7: ‘My daily expenses here are that I make 300 rupees worth of phone calls a day.’

- i. *e*
eh

M7: ‘Eh.’

- j. *mobāil-bāta*
cell.phone-ABL

M24: ‘From my cell phone.’

- k. *ā*
ah

M7: ‘Ah.’

The utterance of interest is (16h), in which the speaker uses the ergative form of the first person pronoun (*maile*) with a transitive verb in the simple present tense (*phon garnu* ‘to telephone’). In discussing this utterance with BB, I noted that the subject is postposed and asked whether the nominative form is possible in the given conversation and whether doing so appears to change the intended meaning. BB expressed the intuition that using the nominative form would sound a little strange if the subject is postposed but would be acceptable if the subject is expressed before the verb. I noted that the reading here is habitual, but in another context the same clause could express a future-oriented event, and BB noted that both the ergative and nominative would be acceptable whether the meaning is habitual or future-oriented. The reading can also be ongoing in the simple present, and BB noted a difference

between *ma phon garchu* ‘I am telephoning.’ and *maile phon garchu*: “with *maile phon garchu* I am thinking *kaslāi?* (‘whom?’) Who is receiving the impact? With *ma phon garchu* this is not the case.”

In the Fall of 2018 and Spring of 2019, I worked primarily with SB to annotate every utterance in each of the four interviews. First, we listened to the audio recording of the interview. I used R to create a dataframe of utterances by speaker, and I wrote an R script to convert the Devanagari script into Roman letters for my own readability. SB and I worked through the transcription to provide English glosses for each of the utterances. I spoke with SB about various grammatical points that arose and made notes on relevant topics. Then, I segmented each utterance by clause and made the following annotations:³

Clause Order: (1, 2, 3, 4...)

Interview Number: V001001001, V001001003, V001002003, or V001002005

Speaker: Speaker order as listed in the transcription (A, B, C, D, E). In my analysis I associated these codes with the proper Speaker ID.

Utterance (Devanagari): The original text

Utterance (Transliterated): Romanized version of the text

Gloss: Basic English gloss with a focus on conveying the intended meaning introduced by the clause in question. In sentences with multiple embedded clauses I ended

3. For this process I used Microsoft Excel Workbooks which I read into R as a dataframe for analysis.

the gloss with a comma to convey that the clause is part of a whole.

Verb: The full form of the particular verb or verb construction contained in the clause. These were tagged as OMIT if the verb was clearly elided in a clause with propositional content, and NA if the given utterance did not contain a verb at all, as with a phatic expressives of agreement or interjections like *ā* or *e*.

Verb Type: TRANS (transitive predicates, including ditransitive predicates), IN-TRANS (intransitive predicates), and COP (copular clauses of all types).

Verb Tense: Main clause verb forms: I subsumed all present forms under the tag PRES (PRES, CONT, ARCH.PRES, as well as present modal forms with incorporated verbs like *-idin-*, *-ihāl-* and *-isak-*). The other main clause tags were FUT (FUT), HYP (HYP), PAST HAB (PST.HAB), PAST (referring only to the past tense copula *thyō*), and OPT (OPT, which occurred only in copulas). I also tagged imperative verb forms with IMPER. I subsumed all perfective tenses under the tag PERF (PERF, PRES.PERF, PST.PERF, PRES.MIR, PST.MIR, including perfective modal forms with incorporated verbs like *-idin-*, *-ihāl-* and *-isak-*, and including perfective nominalizations with *-eko*). Subordinate Clause forms: The tag ADV includes verb forms that end in the adverbial forms (*-era*, *-dā(kheri)*, *-epacchi*, *-epani*, etc.). The tag NONFIN includes verb forms that end in nonfinite markers (*-na*, *-nu*, *-ne*) and are not part of a larger verbal construction.

Verb Construction: MODAL (any modal periphrastic constructions: *saknu* in the ability construction *X-na saknu*, *parnu* in the obligation construction *X-nu parnu*, *manparnu*, *lāgnu*, and *manlāgnu*), COND (conditional constructions with *bhane*, e.g. *X-yo bhane*), NOM (any perfective nominalizations with *-eko* or imperfective nom-

inalizations with *-ne*), PASS (passive constructions). The few instances of complex combined constructions were doubly tagged as PASS COND, PASS NOM, MODAL COND, etc. Clauses which did not contain a Verb Construction were tagged NA.

Subject: The Subject Argument (S_t in a transitive clause, S_i in an intransitive clause, the subject of a copular construction). Tagged OMIT if the subject was elided.

Subject Case: The most frequent cases were ERG (ergative), NOM (nominative), and NA (if omitted). There were scattered examples of ACC (accusative, if the subject of a nominalization is the object of a main clause) and CLAUS (if the subject is a clause).

Subject Type: Pronominals: 1SG (first person singular pronoun), 1PL (first person plural pronoun), 2SG (second person singular pronoun; including both genders and honorific forms), 2PL (second person plural pronoun; including both genders and honorific forms), 3SG.ANIM (third person singular pronoun with animate reference; including both genders and honorific forms), 3.PL.ANIM (third person plural pronoun with animate reference; including both genders and honorific forms), 3.SG.INAN (third person singular pronoun with inanimate reference), 3.PL.INAN (third person plural pronoun with inanimate reference). This category also marks elided subjects, for which person, number, and animacy are retrievable from the discourse context. Nominals: ANIM (animate), and INAN (inanimate).

Object: The O argument of a transitive clause.

Object Case: ACC (accusative), NOM (nominative), NA if omitted (or if the given clause is not transitive).

Object Type: Same categories as in Subject Type.

Other: The goal argument in a ditransitive clause, the dative-marked experiencer of a copular clause, or an instrumental argument.

Case: DAT (DAT *-lāi*), INSTR (INSTR *-le*), NOM (nominative).

Notes: Additional notes about the meaning of particular clauses or grammatical constructions.

I created some general principles for tagging Nepali speech for an analysis of argument realization in clauses. Many of these principles had the effect of delimiting what I consider to be a clause for this analysis. There are several constructions in Nepali which are technically separate clauses but which cannot contain arguments, so I did not annotate them.

- (1) As stated above, expressive phatics, including expressions of agreement and approval like *hajur* and *hai*, were not considered to be clauses and were marked NA. This included forms of the *ho* copula (*ho*, *hoina*, *huncha*) if they were used to signal approval or agreement rather than as a predication.
- (2) I did not consider tag questions to be separate clauses, but rather part of the main clause. If a verb was simply repeated for emphasis I did not consider it to be part of a new clause.
- (3) I considered honorific verb forms (*X-nu huncha*, *X-nu bhayo*) to consist of honorific marking on the verb rather than analyzing them as multi-clausal constructions.

- (4) I did not consider common fixed expressions (*jehos* ‘whatever’, lit. ‘whatever may be’, *tyas bhannale* ‘by saying that’, etc.) to be separate clauses.
- (5) In general I did not consider grammaticalized usages of *bhannu* ‘to say’ to be separate clauses. This includes the usage of *bhaneko* before a topic roughly in the sense of the English ‘speaking of X’, which I consider to be a discourse marker. As mentioned above, I consider *bhane* to be a separate construction when used as a conditional. I consider *kina bhane* (‘because’, lit. ‘saying why’) to be a conjunction.
- (6) Clauses in other languages (English, Hindi, and Sanskrit) were put in quotes in the gloss and left unanalyzed. Nouns in other languages which were otherwise embedded in Nepali clauses were treated as loanwords.
- (7) For copular forms, both *cha* and *ho* were tagged PRES, *bhayo* was tagged PERF, *thyo* PAST, *hunthyo* PAST HAB, *hune* NONFIN, and there are various ADV forms like *bhaepacchi*, *hudo*, etc. I considered *holā* with a single argument to be OPT, but if incorporated in a regular clause I considered it to be an evidential marker. While it might be argued that *rahecha* acts as a mirative copula in modern Nepali, I considered it to be the intransitive verb *rahanu* ‘to stay’ in the MIR.PRES verb form (the same source as the *-iraha-* of the PROG verb form).
- (8) If a subject or object was elided and its referent was known from context, I made a note of it.
- (9) Some verb forms like *cāhinu* ‘to want’ could be analyzed either as passive forms or as impersonal verbs which take dative-marked experiencer subjects and have default singular agreement. I tagged them as passive constructions for simplicity.
- (10) The transcription consisted of shortened forms indicative of rapid spoken language. Some common examples of this were *X-esi* for *X-epacchi*, *X-yaa* for

X-eko cha, and *X-diu* for *X-idinus*. I frequently relied on SB’s interpretation to distinguish between, for example, present perfect and present mirative verb forms in rapid speech.

- (11) Because of the relatively free word order in Nepali, and because arguments are frequently elided if coreferential with an argument in another clause, it can be impossible to tell which clause an overt argument belongs to in a complex sentence. In these cases I tagged the argument as overt in both positions.

2.4.3 Advantages and Limitations of Corpus Analysis

Corpus analysis examines actual speech rather than relying on the judgments speakers have about their own language. Thus it can provide corroboration of intuitions obtained from elicitation, and going through recordings with elicitation consultants is a useful technique for gathering new insights and judgments. Furthermore, it is a primary source of quantitative data on the realization of case marking in different grammatical contexts. It allows us to study the frequency of a particular case-marking pattern in proportion to the overall realization of arguments.

However, analysis of naturally-occurring speech does not give us negative data; if we do not find a particular construction, it does not tell us that such a construction cannot exist. Furthermore, annotation is a time-consuming activity, and the grammatical contexts of interest occur in a minority of the overall clauses. For example, in examining the frequency of nominative versus ergative case in simple present tense clauses, I restrict my focus to simple present tense clauses with overt subjects that are not part of modal constructions. Out of the 2845 total clauses, 54 clauses fit those criteria.

The thirteen speakers represented in this corpus sample are mostly male and university-educated, and they are from many different areas of Nepal. Only one listed a language other than Nepali as a first language, but two additionally list dialects of

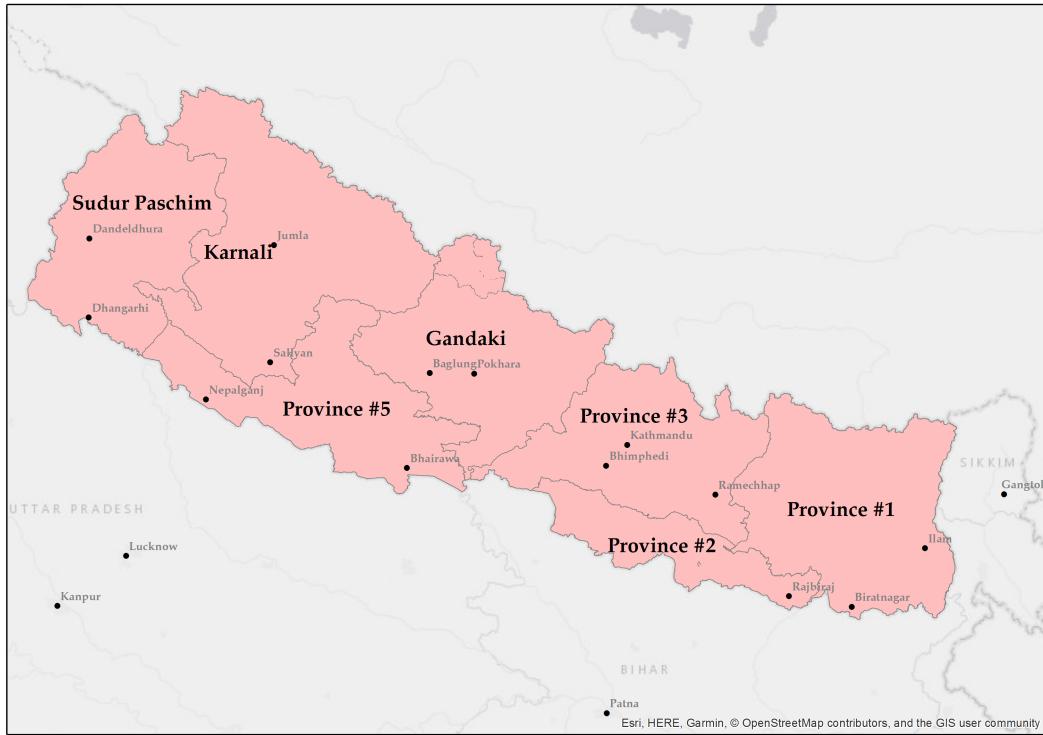
Nepali that they speak. Thus we might suspect an effect of dialect variation, which I discuss in the next section.

2.5 The Problem of Lect

Variable ergativity in Nepali is conditioned by multiple pragmatic and semantic factors. Considering the dialect diversity in Nepali, it is reasonable to suspect that variable ergativity might have different manifestations in different dialects and vary by social, ethnic, and geographical factors. A significant proportion of Nepalis are L2 learners of the Nepali language, the variable ergative pattern in Nepal may have developed from contact with Tibeto-Burman languages in which this type of pattern is more common. However, among the 53 Nepali speakers whose statements and judgments form the basis of this analysis, there is fairly little diversity. While individual speakers may differ in their precise judgments on particular pragmatic and semantic phenomena, the broad pattern is the same: the ergative is obligatory in the perfective and variable in the imperfective where it correlates not with agency/volitionality but with the subject being emphasized as an effector of the event.

In Figure (2.5) I have tabulated the birth regions by Province of every elicitation consultant, survey respondent, and speaker from the NNSP sample. My focus of investigation has been primarily on speakers from Kathmandu Valley in Province 3, which is the regional dialect studied by Clark (1963), Abadie (1974), and Li (2007). While Province 3 predominates in this sample of Nepali speakers, there is a fair amount of regional diversity, particularly among the Corpus speakers.

As noted in the first chapter, the greatest amount of dialect diversity is in Sudur Paschim and Karnali. In the east of the country (the first three provinces) there is much less dialect diversity. The general social perception of dialect in Nepal is that speakers in the far east speak the most “pure,” prescriptively correct Nepali, while



	Elicitation Consultants	Survey Respondents	Corpus Speakers
Province 1	SB, BB, KS	18, 19, 26	M3
Province 2			M2
Province 3	BA, TD, AG, PK, RM GM, SP, UP, ST	8, 9, 10, 11, 12, 13, 14, 15 16, 17, 20, 21, 23, 24, 28	F4, M9, M24, M26
Gandaki	MG	3, 4, 6,	M1, M13
Province 5		1, 4, 27	M7
Karnali		7	
Sudur Paschim		22, 25	M14

Figure 2.7: Birth Regions of Nepali Speakers

in the West (and in the southern part of the whole country) the speech is thought to have been corrupted by influence from other languages. Because people from all over the region move into Kathmandu Valley, the speech is more varied, though not as varied as in the West.

There is very little in the literature that pertains to ergative patterning in dialects other than that of Kathmandu. Ahearn (2001b) writes that the usage of ergativity in the imperfective is prominent in the Nepali-speaking Magar community she studies, although this usage was first noted by Grierson (1904b), whose consultants were mainly from the Darjeeling region of India along Nepal's eastern border (Nepal was

closed to foreign researchers at the time). Clark (1963) notes that this usage is prescriptively incorrect but common in colloquial speech. UP (a principal and former schoolteacher in Lalitpur district) and other educators have confirmed to me that the usage of the ergative in the imperfective is generally seen as prescriptively incorrect and emblematic of Nepali spoken by ethnic groups who traditionally spoke their own languages. Dr. Madhav Pokharel mentioned to me that ergative alternations in the imperfective are not found in the speech of educated Brahmins and Chhetris, although he did find it to be acceptable with intransitives like *nācnu* ‘to dance.’⁴ Rather, this is indicative of the speech of Nepalis from the *Magarrat*, in the traditional lands of the Magar ethnic group west of Kathmandu. SB had an observation about the converse pattern. While ergative marking is generally required in the perfective, SB noted that usage of the nominative in perfectives is emblematic of speech in the southern Terai, both among communities that traditionally speak other languages like Tharu and Maithili and among Brahmin-Chhetris in the region.

Despite the fact that the majority of my consultants were university-educated, and many of them were Brahmin-Chhetri and born in the East, the usage of ergativity in the imperfective is found in the speech of every one of them, and it is also found in the speech of every one of the corpus speakers.⁵

Figure (2.5) illustrates the preponderance of ergativity in imperfective clauses in the Kathmandu Survey and in the corpus sample. For the Kathmandu survey, I tabulated the average rating of the ergative form for every question in the simple present tense for each speaker. There are slight differences in the overall ratings,

4. Brahmin and Chhetri were the top two rungs of a legally-codified caste hierarchy in Nepal until the Constitution of 1990, and the Nepali language as spoken by Brahmins and Chhetris of the Khas ethnic group is considered to be the standard.

5. Throughout my research there was only one generalization that I could confidently correlate with regional dialect: MG, a native speaker of Gurung from Gandaki, was the only consultant to categorically disallow ergative marking in the intransitive verbs *nācnu* ‘to dance’ and *bhoknu* ‘to bark.’

but the average is high (around 4.5 out of 5), and only one speaker has an average rating below 3.⁶ In the corpus sample, the percentage of ergative-marked subjects was tabulated among all transitive imperfective clauses with overt subjects. Ergative-marked subjects are found in the speech of every respondent, with proportions ranging from 29%- 67% with an average of 48%.⁷

Taken together, this strongly indicates that all speakers of the Nepali dialects represented here allow the ergative in imperfective clauses. Furthermore, it is not immediately apparent that there are distinctive patterns associated with a particular region, caste, or mother tongue. Further research will be required to tease apart the complicated issue of language variation. However, my overall impression is that while speakers from different regions and language backgrounds may differ somewhat in the strength they give to particular pragmatic and semantic features (such that some may have grammaticalized, say, a rule that inanimate subjects must be marked ergative while for others it may simply be a strong tendency), for each speaker it is the same set of pragmatic and semantic features which are at play.

6. This speaker, 26, is a 53-year-old academic from Province 1, which suggests a prescriptive bias.

7. The lowest values were found with speakers M9 and M13, from Province 3 (Kathmandu) and Gandaki respectively. The highest values were found with speakers M2 and M7, from Province 2 and Province 5.

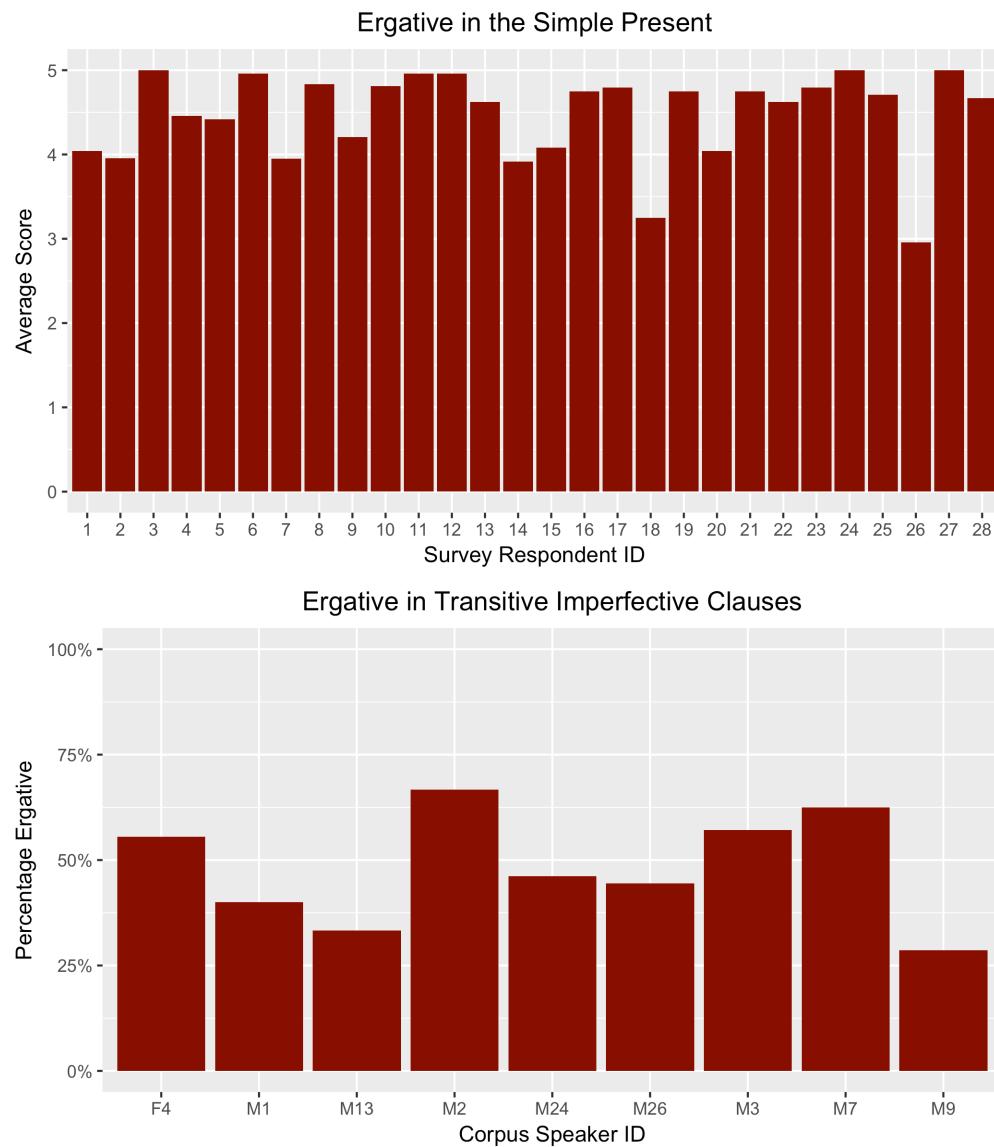


Figure 2.8: Ergative Marking in Imperfective Clauses by Speaker

Chapter 3

Theories

In this chapter, I begin with an overview of the explanations that have been offered for the Nepali ergative pattern. These explanations fit into the broader literature on optional case marking systems. I then overview the major explanations that have been proposed for **Optional Ergative Marking** and discuss them in the context of the general literature.

3.1 General explanations of the Nepali Pattern

Nepali grammars and grammatical sketches (Grierson 1904a, Turnbull 1982, Acharya 1991) do not typically go into detail with regards to the semantic or pragmatic factors that differentiate case marking in the imperfective domain. However, discussions of these differences may be found occasionally in readers for English-speaking students of Nepali (Clark 1963, Matthews 1998, and Hutt and Subedi 1999).

The earliest detailed investigation into the variable ergative marking pattern in Nepali is Abadie (1974), who looked at ergative marking in different main clause tenses and in modal contexts. Li (2007) examines ergative patterning in transitive and intransitive main clauses and concludes that there is a categorical split based on animacy: inanimate subjects take ergative marking in all tenses and aspects. Verbeke

(2011) has an excellent review of the previous work on Nepali ergativity throughout the grammar. In this section, I will first discuss explanations that have been offered for the ergative patterning in transitive main clauses. I will then discuss the use of the ergative in intransitive contexts, and then explanations of the ergative in modal contexts.

3.1.1 Case Marking in Traditional Grammar

Traditional Nepali grammars are based on the model of the ancient Sanskrit grammarians, particularly the fifth-century BCE *Aṣṭādhyāyī* of Pāṇini and its commentaries. In the description of the nominal case system of Sanskrit, Pāṇini notably disassociates the form of a case inflection from its associated semantic role. This distinction is expressed by the Sanskrit terms ***Vibhakti*** (“case”, “division”) and ***Kāraka*** (“role”, lit. “doer”) (Keidan 2011).

Role	English Gloss
<i>Kartā</i>	Agent
<i>Karma</i>	Patient
<i>Karaṇ</i>	Instrument
<i>Sampradān</i>	Beneficiary
<i>Apādān</i>	Source
<i>Sambandha</i>	Possessor
<i>Adhikaran</i>	Location
<i>Sambodhan</i>	Addressee

Figure 3.1: Nepali *Kāraka*: Argument Roles (Pradhān 1944)

The relation between *vibhakti* and *kāraka* can be understood as a distinction between the form of a nominal case inflection on the one hand, and the semantic role of a participant on the other hand.¹ The *vibhakti* are simply named *Prathamā* “first,”

1. Computational versions of this grammatical framework have been applied to text clustering and treebank parsed text corpora. This Paninian Framework may be particularly well-suited to Indic languages because of their relatively free word order and because of mismatches and interplay between semantic roles and syntactic forms. Some contributions in this field include Bharati et al. (1995), Begum (2017), and Sharma and Gupta (2012).

Dvitiyā “second,” etc., while the *kāraka* have descriptive names which indicate the semantic role of the noun. These classifications and terms, which are also found in Whitney (1879) and MacDonell (1927)’s Sanskrit grammars, are applied to Nepali in Pradhān (1944)’s influential early grammar of Nepali.²

Pradhān’s chapter on Nepali case begins with a description of the eight Nepali *kāraka* (Pradhān 1944: 18-25). These roles are depicted in Figure (3.1). The two roles which are most relevant to our purpose are *kartā*, or ‘agent’, and *karan*, or ‘instrument.’ The *kartā* is defined as the participant “which does work or has the ability to do work.”³ The *karan* is defined as the participant “through the aid of which the work is accomplished.”⁴ Pradhān notes that the marker *le* is associated with both roles, but the *kartā* role answers the question “Who did the action?” while the *karan* role answers the question “With what was the action done?”

Pradhān does not mention the optional usage of *le* in the imperfective, although from Grierson (1904a) we know it was a feature of colloquial speech, and Clark (1963) notes that it was considered prescriptively incorrect by traditional grammarians.

Pradhān summarizes the *vibhakti* by associating each with a semantic role and giving corresponding suffixes. Figure (3.2) is a translation of his chart with corresponding English case terminology.

I have given the correspondence to the first *vibhakti* as both Nominative and Ergative, although there is not a conception of ergative case in traditional Nepali grammar. Rather, in discussing the *Kartā* role Pradhān considers *le* to be the marker

2. Tikaram Poudel (p.c.) emphasizes the influence of Parasmani Pradhān’s *Nepālī Vyākaraṇ* (“Nepali Grammar”), which was first published in 1920. There are a few departures from Pāṇini’s original classification: Whitney (1879) notes that Pāṇini has seven *vibhakti*, but Whitney and Pradhān both include an eighth vocative case. Additionally, Pāṇini does not count the genitive case because it is a relation between a noun and another noun, while the rest are noun-verb relations (Jinitha 2009).

3. “*Kunai kā garna wā garnasakne sangyālāi ‘kartā’ bhandāchan.*” [All translations are my own.]

4. “*Jaskā sahāyatāle kunai kā garincha tyaslāi ‘karan’ bhandāchan.*”

Vibhakti (Case)	Meaning	Marker or Suffix
<i>Prathamā</i> (Nominative/Ergative)	<i>Kartā</i> (Agent)	<i>le/bāta</i> ⁵
<i>Dvitiyā</i> (Accusative)	<i>Karma</i> (Patient)	<i>lāi</i>
<i>Tritiyā</i> (Instrumental)	<i>Karaṇ</i> (Instrument)	<i>le/bāta</i>
<i>Caturthi</i> (Dative)	<i>Sampradān</i> (Beneficiary)	<i>lāi</i>
<i>Pancami</i> (Ablative)	<i>Apādān</i> (Source)	<i>dekhī/dekhīn/bāta</i>
<i>Sasthi</i> (Genitive)	<i>Sambandha</i> (Possessor)	<i>ko/kā/kī</i>
<i>Saptami</i> (Locative)	<i>Sambodhan</i> (Location)	<i>mā</i>
<i>Sambodhan</i> (Vocative)	<i>Bolānu</i> (Addressee)	<i>e... ho</i>

Figure 3.2: Nepali *Vibhakti*: Nominal Cases (Pradhān 1944)

of this role, but it is only used in transitive verbs in the “past” and “future” tenses.⁶ The past tense includes all of my perfective forms, and the future is what I refer to as the Definite Future. Pradhān strongly associates the *vibhakti* with its corresponding *kāraka*, but the suffixes themselves correspond less closely to the *kāraka* roles: *le* is associated with both *Kartā* and *Karaṇ*, while *lāi* is associated with both *Karma* and *Sampradān*.

One could imagine an alternate formulation in which each suffix is taken to represent a single *vibhakti*. In Sanskrit, the main usage of the Instrumental *vibhakti* was to describe an instrument or means by which the action takes place. But it had other uses. It could designate a reason or medium. It could also designate an agent, particularly as the demoted subject of a passive construction (Whitney 1879: 81). In the same way, one could imagine *-le* to be an instrumental marker even when it appears on agents.

However, no traditional grammar of Nepali does this as far as I am aware: S_t and S_i , whether unmarked or marked by *le*, are always considered to be *Prathamā vibhakti* (nominative case).⁷ Clark (1963) is a Nepali reader for English speakers which does

6. “*Kartā kārakko cinha ‘le’ ho. Sakarmak kriyāmā bhut ra bhivashyat kā bujhāūdā kartā kāraksita yo vibhakti prayog huncha. Akarmak kriyāmā tā hundaiña.*” [“The marker of the Agent role is ‘le.’ For transitive verbs which are implied to be in past or future tense, this case is used on the Agent role. For intransitive verbs it is not.”] (Pradhān 1944: 19).

7. Tikaram Poudel, p.c.

conceive of the grammar in this way: the argument marked with *le* is considered to be in the Instrumental case whether or not it is the subject. Grierson (1904a) called this a separate “Agent case,” which is more in line with the Sanskrit tradition.

Most modern English language descriptions of the Nepali language do not engage with the traditional literature on Nepali grammar, although Verma 1976 brings up the *vibhakti/kāraka* distinction in discussing the notion of the subject in Nepali.

3.1.2 Disambiguation

Abadie (1974) is a detailed investigation of Nepali ergative patterning in main clauses and in modal contexts. Abadie argues that the presence or absence of the ergative marker rarely signals a difference in meaning. Rather the ergative marker disambiguates the subject argument in examples like the following in which it would otherwise be ambiguous:

- (17) *yo gāi-le khān-cha*
 this cow-ERG eat-PRES.3.SG
 ‘This cow eats.’ (Abadie 1974: 170)

Without the ergative marker, there are two possible interpretations. The demonstrative could either apply to “cow” or it could apply to a deleted subject, with “cow” acting as the direct object of “eat”:

- (18) *yo gāi khān-cha*
 this cow eat-PRES.3.SG
 ‘This cow eats.’ / ‘This (one) eats cow.’ [TD]

Another example provided by Abadie is the following:

- (19) *birua-haru-lāi hāmi nepāli-(le) jamarā bhan-da-chaū*
 plant-PL-ACC we nepali-(ERG) jamara call-PRES-1.PL
 ‘(This) plant we Nepalis call *jamara*.’ / ‘(This) plant we call *nepali jamara*.’
 (Abadie 1974: 170)

If the ergative marker is used, only the first reading is possible. *Nepali* is unambiguously the subject of the clause. Otherwise, both readings are possible.

This explanation, that the motivation for the ergative marker is to disambiguate arguments, is common for optional ergative patterns. I discuss this in the next section on General Theories of Optional Ergativity. However, it does not explain the presence of ergative/nominative alternations in which there is no ambiguity about the identity of subject argument.

3.1.3 Animacy

Some linguists have argued that the presence of ergative marking in Nepali is at least partially conditioned by the semantics of the NP. In particular, Verma (1976), Pokharel (1998), and Li (2007) argue that subjects with inanimate reference must be marked in the ergative in all tenses:

- (20) a. *dhungā-haru-le jhāl phut-ā-e*
rock-PL-ERG window break-CAUS-PERF.3.PL
'The rocks broke the window.' (Li 2007: 1465)
- b. *dhungā-haru-le jhāl phut-āu-dai-chan*
rock-PL-ERG window break-CAUS-PROG-PRES.3.PL
'The rocks are breaking the window.' (Li 2007: 1467)

In the first example above, we expect the ergative to be required because the transitive verb is in the perfective. In the second example as well, in which the ergative marker would usually optional because the tense is imperfective, the ergative marker is again required because the subject is inanimate.

Li was not the first to notice this tendency of inanimate subjects in Nepali. Verma (1976) noted this as well, and argued that the ergative is required because it is a "secondary" agent in sentences like the following:

- (21) *dudh-le ketā-haru-lāi pos-cha*
milk-ERG child.OBL-PL-DAT nourish-PRES.3.SG

‘Milk nourishes children.’ [SB]

However, Li argues that this represents an exceptionless split in the ergative paradigm: all inanimate subjects are required to be marked irrespective of the tense (in the domain of transitive verbs).

This is significant because it suggests the presence of a second ergative split in the language: a split based on the animacy of the referent NP in addition to the known split based on verbal aspect. Ergative splits based on the **Nominal hierarchy** are common in the world’s languages, and will be discussed at length in the Markedness section of the General Theories of Optional Ergativity chapter below.

However, there are two potential problems with Li’s argument. The first is that it can be difficult to tell whether an inanimate argument is the *subject* in a transitive clause or if it is a instrumental (non-core) argument. Recall that the instrumental marker is also *-le*.

Taking Verma’s example above, it could be argued that the intended meaning was actually “(One) nourishes children with milk.” The only way to differentiate the subject and instrumental is with verbal cross-reference; if the subject is plural then plural agreement should be possible on the noun, but there will be no agreement with an instrumental.

With Verma’s example, “milk” is a mass noun, so it would not generally be marked with a plural. If we change the nourishment to eggs, we run up against the difficulty that the plural would not usually be marked in a generic context (“eggs nourish children”), but we have to change the example to “These eggs will nourish these children.” My consultants found that plural marking on the verb is still infelicitous here:

- (22) *yi andā-haru-le ketā-haru-lāi pos-cha*
these egg-PL-ERG child.OBL-PL-DAT nourish-PRES.3.SG
(#pos-chan)
(#nourish-PRES.3.PL)

‘(Someone) will nourish children with these eggs.’ (not, ‘These eggs will nourish children.’) [SB]

This indicates that we are looking at an instrumental rather than an ergative subject, and it illustrates how difficult it can be to tease them apart. Inanimate subjects are somewhat rare in Nepali, but Poudel (2008) gives an example to illustrate that inanimate subjects are possible:

- (23) a. *yo cābi-le yi dhokā khol-yo*
this.SG key-ERG this.PL door open.TRANS-PERF.3.SG
‘This key opened these doors.’
- b. *yo cābi-le yi dhokā khul-e*
this.SG key-INSTR this.PL door open.INTRANS-PERF.3.PL
‘These doors opened with this key.’

In the first example, “key” is argued to be the subject of the sentence, whereas in the second example “key” is an instrumental. Plurality is expressed with the usage of singular or plural determiners *yo/yi*, and there is verbal agreement in number. In the first example, the verb has singular agreement and “key” is singular. In the second example, the verb has plural agreement to agree with “these doors.”

However, the first example of this sentence pair is still potentially ambiguous with a reading “(Someone) opened these doors with this key.” To prove that “key” is indeed the subject here, one would need to specify a context in which multiple people opened multiple doors using a single key.

Another issue with the animacy argument is that there may be exceptions to the generalization. Verbeke (2011) in particular points to examples in the Nepali literature of sentences with inanimate subjects which are unmarked:

- (24) *sireto mutu ched-lā bhane jasto gar-th-yo*
wind heart pierce-INDEF.FUT QUOT if do-IMPF.3.SG
‘The wind blew (acted) as if it would pierce the heart.’ (Acharya 1991: 191)

The Nominal hierarchy, as described in detail below, includes subdivisions of common nouns and personal pronominal forms. Human referents are ranked higher than non-human referents. Interestingly, Pokharel (1998) suggests that ergative marking is less common for humans, suggesting more of a tendency than a split. Giving the examples below, he writes “the inanimate transitive subjects obligatorily take *-le*, the human subject may or may not take it and the non-human subject more likely chooses it.”(Pokharel 1998: 47)

- (25) *bancara-le rukh dhal-cha*
ax-ERG tree fell-PRES.3.SG
'The ax fells the tree.'⁸ (Pokharel 1998: 47)
- (26) *kamila-(le) cini khān-cha*
ant sugar eat-PRES.3.SG
'The ant eats the sugar.' (Pokharel 1998: 47)
- (27) *mānche-(le) cini khān-cha*
person-(ERG) sugar eat-PRES.3.SG
'The person eats the sugar.' (Pokharel 1998: 47)

Pokharel suggests here that these are tendencies rather than hard categories. So animacy may be a factor in the expression of ergativity, but it remains an open question whether there is a categorical split as Li describes. Furthermore, it may be a secondary factor or an epiphenomenon of some other conditioning factor.

3.1.4 Stage-Level and Individual-Level Predication

Butt and Poudel (2007) argue that *-le* marks individual-level predication. This terminology comes from Carlson (1977b)'s division of predicates into two natural classes, as **stage-level** or **individual-level**. Stage-level predicates typically describe transient or episodic states, while individual-level predicates describe enduring properties.

Consider the following sentences of English:

8. This might be an instrumental usage of *-le*.

(28) Examples of stage-level predicates:

- a. Shristi is distracted.
- b. The zookeeper is feeding the crocodiles.

(29) Examples of individual-level predicates:

- a. Kiran is left-handed.
- b. Leopards are mammals.
- c. Horatio hates to play the harmonica.

The first two examples describe temporary, time-bounded episodes, while the latter three describe properties of an individual or a kind. Note that individual-level predicates may be copular phrases or predicates with a generic interpretation. Carlson argues that individual-level predicates denote properties of an *individual*, while stage-level predicates denote properties of a particular spatio-temporal slice or *stage* of an individual. Others have formalized the distinction in terms of different syntactic constructions (Diesing 1992) or argument structures (Kratzer 1995, Chierchia 1995).

English adjectives like *drunk*, *ready*, *infuriated*, and *dusty* are typically stage-level, while adjectives like *clever*, *redheaded*, *tall* and *reliable* are typically individual-level. Some adjectives can be interpreted as stage-level or individual-level depending on the context: *sick* as a stage-level adjective refers to a lapse in physical health, while *sick* as an individual-level adjective refers to a more permanent property of mental health.

The two types of predication behave differently in particular grammatical contexts of English, as with bare plural subjects (Carlson 1977a) or adverbs of quantification (Kratzer 1995). Furthermore, the distinction has been argued to apply in a wide range of grammatical constructions crosslinguistically (Roy 2013). In particular, multiple languages appear to grammaticalize the distinction with different copulas for individual-level and stage-level predictions:

(30) Spanish (Leonetti 1994: 255,260)

- a. *juan está triste*
juan COP sad
'Juan is sad.'
- b. *maría es muy guapa*
maría COP very pretty.F
'María is very pretty.'

(31) Scottish Gaelic (Ramchand et al. 1997: 204-205)

- a. *tha calum sgith*
be-PRES calum tired
'Calum is tired.'
- b. *is dotair calum*
Cop-PRES doctor Calum
'Calum is a doctor.'

(32) Marathi (Deo 2017)

- a. *karvādh-i-mule lok-a dukkhi āhe-t*
tax.increase-OBL-by people-NOM.N.PL unhappy PRES1-3.PL
'People are unhappy because of the tax increase.'
- b. *dāyet kar-ṇāri lok-a dukkhi as-tāt*
diet do-PART.N.PL people-NOM.N.PL unhappy PRES2-3.M.PL
'Dieters are unhappy.'

It should be noted that in recent years, some linguists have argued that the stage-level/individual-level distinction is too coarse-grained and does not accurately represent the semantics (see in particular Roy (2013) for a more nuanced theory) or discourse properties (for example Deo (2017) discusses the facts in Marathi and Sanchez-Alonso (to appear) in Spanish). But the stage-level/individual-level distinction is relevant for an understanding of the Nepali pattern.

Stage-Level Predication in Nepali

As mentioned in the Introduction section, Nepali has two present tense copulas *cha* and *ho*. Butt and Poudel (2007) invoke the stage-level / individual-level distinction in explaining the distribution of these copulas:

- (33) Nepali (Butt and Poudel 2007: 5)

- a. *saru.bhakta āja khusi chan*
saru.bhakta today happy COP2.PRES.3.M.HON
'Saru Bhakta is happy today.'
- b. *saru.bhakta kabi hun*
saru.bhakta poet COP1.PRES.3.M.HON
'Saru Bhakta is a poet.'

They argue that outside of the copular domain this distinction is preserved with the presence or absence of the ergative marker (Butt and Poudel 2007: 7):

- (34) a. *cālak-le gādi calāu-cha*
driver-ERG car drive-PRES.3.SG
'The driver drives the vehicles.'
- b. *guru gādi calāu-cha*
teacher car drive-PRES.3.SG
'The teacher is driving/will drive the vehicle.'

The context in the first example is that a school's bus driver drives the children everyday: that is his occupation. In the second example the teacher just happens to be driving the bus today because the bus driver is out. The simple present tense of the verb *calāucha* 'to drive' has three possible interpretations: it may have a habitual reading, an immediate present reading, or a future-oriented reading. In this sentence, *-le* marks the reading with individual-level predication, that is, the habitual reading.

Butt and Poudel (2007) note that the ergative is never found with copulas; they are in complementary distribution. Thus the stage-level/individual-level distinction (in the present tense) is preserved with the dual copulas in the copular domain and

with the ergative marker outside of it. Furthermore, the distinction disappears in the perfective for copulas (there is only one copula in the perfective), just as the ergative/nominative alternation disappears for transitive verbs disappears in the perfective.

Similarly, Hutt and Subedi (1999) note that the ergative “can be used to emphasise the subject of a transitive verb in the habitual present tense... if the sentence says that it is a part of the natural order of things for the subject to perform the verb, and therefore states that this is a role that is specific to the subject” (Hutt and Subedi 1999: 116). They provide the following examples:

- (35) a. *kukhurā-le phul pār-cha*
chicken-ERG egg lay-PRES.3.SG
'A chicken lays eggs.' (Hutt and Subedi 1999: 116)
- b. *bāgh-le bākhrā khān-cha*
tiger-ERG goat eat-PRES.3.SG
'A tiger eats goats.' (Hutt and Subedi 1999: 116)
- c. *ghām-le nyāno din-cha*
sunshine-ERG warmth give-PRES.3.SG
'Sunshine gives warmth.' (Hutt and Subedi 1999: 116)
- d. *pakkā bahun-le raksi khān-daina*
proper Brahmin-ERG alcohol eat-PRES.3.SG.NEG
'A proper Brahmin does not drink alcohol.' (Hutt and Subedi 1999: 116)

However, one problem with this theory is that *-le* may also be found in clauses for which the predicate is neither habitual nor individual-level, and in tenses for which there is no ambiguity. Verbeke provides a few examples in which individual-level predication does not seem to be present:

- (36) *rājā-le sodh-e han yas-le pheri ke bhan-dai-che*
king-ERG ask-PERF.3.SG Q she-ERG again what say-CONT-PRES.3.F.SG
'The king asked: What is she saying?' (Verbeke 2011: 165)

- (37) *aphu-le na-bok-era kas-le bok-i-din-cha?*
 oneself-ERG NEG-CARRY-CONJ who-ERG carry-LNK-GIVE-PRES.3.SG
 ‘If I don’t carry it myself, who will carry it?’ (Verbeke 2011: 165)

While the simple present tense is ambiguous between stage-level and individual-level interpretations, there are other imperfective tenses for which the ergative/nominative alternation exists, and Butt and Poudel’s theory is inapplicable. We can make a minimal adjustment to their example to make it unambiguously stage-level. Here the ergative is still in alternation.

- (38) a. *guru-(le) gādi calāu-dai-cha*
 teacher car drive-CONT-PRES.3.SG
 ‘The teacher is driving/will drive the vehicle.’

3.1.5 Perfectivity

Verbeke (2011) and Verbeke (2013) suggest that the use of *-le* outside of the perfective form provides a sense perfectivity or completion to the action of an imperfective verb. The interpretation is that a state of affairs is certain or factual, and that an endpoint is implied.

The use of *-le* in the imperfective context may be employed to convey that the action is conceived of as certain or inherently completed in the mind of the speaker. Because of this, use of *-le* may be particularly associated with rhetorical questions, or with certainty that an event will be completed:

- (39) *kas-le timī-lāi birsa-na sak-cha?*
 who.OBL-ERG you-ACC forget-INF can-PRES.3.SG
 ‘Who can (possibly) forget you?’ (Verbeke 2011: 165)

- (40) *maile humkāra gar-ā-ūla*
 I.OBL-ERG command do-CAUS-FUT
 ‘I will have you do my command.’ (Verbeke 2011: 165)

Verbeke also refers to the completed interpretation as *telic*, which commonly refers to the lexical semantics of predicates which have a distinct endpoint, like *arrive* or *cross the street* as opposed to those which do not, like *dance* or *play cards*. Poudel (2008) gives an example of ergative alternation which he refers to as distinguishing “Accomplishment vs. Non-accomplishment.” The notion appears to be quite similar to Verbeke’s:

- (41) a. *rām-le bihāna-dekhi pāni bhy-ā-i rah-eko cha*
 ram-ERG morning-FROM water bear-CAUS-LNK stay-PERF PRES.3.SG
 ‘Ram has been fetching water since this morning. (and he finished fetching it).’ (Poudel 2008: 8)
- b. *rām bihāna-dekhi pāni bhy-ā-i rah-eko cha*
 ram morning-FROM water bear-CAUS-LNK stay-PERF PRES.3.SG
 ‘Ram has been fetching water since this morning. (and he has not yet finished fetching it).’ (Poudel 2008: 8)

In this situation, the usage of the ergative appears to correlate with whether the activity is ongoing or completed.

This approach and the individual-level predication approach share the intuition that ergative marking has an effect on the interpretation of the event. The intuition is that the case marker itself can be considered part of the verbal aspectual morphology. In discussing the Sanskrit passive-to-perfective reanalysis that originally brought ergativity into the languages, Anderson notes that normally “we think of verbal categories such as tense and aspect as marked on the verb, and not (partly) on the NP, but this by no means necessary.”(Anderson 1977: 336). The instrumental in Sanskrit went from being part of the syntactic passivization construction to being part of the general aspectual marking of perfectivity. By the same token, the later extension of ergativity into the imperfective domain in Nepali is reinterpreted as part of the aspectual machinery of the language.

On the other hand, Verbeke argues that the use of *-le* is optional in that it can

be used to emphasize completion but is never obligatory. Ergative marking here is arguably not semantic but pragmatic; it does not have an effect on the truth conditions of the sentence, but rather conveys certainty on the part of the speaker.

While Verbeke's approach moves forward our understanding of emphasis, it is not necessarily applicable to most or all of the domains in which ergative marking varies with its absence.

3.1.6 The Intransitive Domain

In a canonical ergative pattern, we expect ergative marking to be completely disallowed in intransitive clauses. However, we do find ergative marking with certain intransitive verbs in Nepali. For example, Pokharel (1998) notes that intransitive verbs of “outward energy emission,” typically non-volitional natural bodily processes, always take a subject marked by *-le*:

- (42) *ketā-le khok-yo*
boy-ERG cough-PERF.3.SG
‘The boy coughed.’

Li (2007) surveyed 110 intransitive verbs in Nepali, particularly examining those categories relevant for the distinction between **unaccusative** and **unergative** verbs as listed by Perlmutter (1978). Unaccusative verbs, according to the **Unaccusativity Hypothesis**, are those in which the single core argument is initially generated in object position and moves to subject position. They are contrasted with unergative verbs, in which the single core argument is initially in subject position. Although Perlmutter argues that this distinction is primarily a syntactic one, there are semantic features that are broadly associated with one category or the other. Here are the categories listed by Perlmutter and applied to Nepali by Li:

Unaccusative Verbs	English Examples
Argument semantically a patient	<i>burn, fall, drop, sink, float, soar</i>
Predicates of existing/happening	<i>exist, happen, transpire, occur</i>
Non-voluntary emission of stimuli	<i>shine, sparkle, glitter, glow</i>
Aspectual predicates	<i>begin, start, come, cease</i>
Duratives	<i>last, remain, stay, survive</i>
Unergative Verbs	English Examples
Predicates of willed/volitional acts	<i>work, play, cheat, fly, think</i>
Verbs speaking or animal/human sounds	<i>shout, whisper, roar, bark</i>
Involuntary bodily processes	<i>cough, belch, burp, vomit</i>

Li's finding for Nepali is that the unaccusative verbs categorically disallow ergative marking in all cases (Li 2007: 1468). However, with unergative verbs the situation is more complex. Case marking on subjects varies "according to tenses/aspects, and even speakers." Ergative marking in the perfective is possible or obligatory for all of these verbs,⁹ and varies according to speaker in whether it is possible or obligatory in the imperfective domain.

Li argues that lexical semantics plays a role in the patterning of ergatives in the unergative domain. In particular, **telic** verbs seem to resist ergative marking in Nepali. These are verbs which have a set endpoint, like *āunu* 'come', *jānu* 'go', in contrast to verbs which do not have a set endpoint, like *nācnu* 'dance', *dagurnu* 'run.' Thus, Li concludes that only atelic unergative verbs allow ergative marking. Li also notes that 'optionality' may indicate that the marker imparts subtle semantic differences, but leaves this issue open for future research (Li 2007: footnote 9).

Note that this is the second time we have encountered the term "telicity" in a discussion of ergative patterning in Nepali, although here the usage is closer to the

9. There were two exceptions *udnu* 'fly' and *runu* 'cry' for three of the speakers.

way telicity is commonly used in the literature. Verbeke used the term to refer to an emphasis on the completedness of an event (in transitive clauses). Interestingly, for Verbeke ergative marking emphasized telic aspect, while Li notes that semantically telic verbs disallow ergative marking. I will discuss telicity more thoroughly in the Observations section below.

3.1.7 Ergative Marking in Particular Structures and Tenses/Aspects

Abadie (1974), Pokharel (1998), and Poudel (2008) have each noted the particular behavior of the ergative in certain tenses and modal constructions. The meaning of these tenses may tell us about the contribution of the ergative marker. Abadie in particular notes the behavior of the ergative in construction in tenses which deal with the semantic area of permission or obligation. These include the tenses which I am referring to as the indefinite future, the definite future, and modal constructions of obligation and necessity.

Hypothetical Future

In this tense the ergative marker is optional “for emphasis,” according to (Clark 1963: 164). Abadie notes that with the bare form can be a question, but this cannot be the case with the ergative form.

- (43) a. *ma khā-ū*
I eat-HYP.FUT.1.SG
'Should I eat?'
- b. *mai-le khā-ū*
I-ERG eat-HYP.FUT.1.SG
'It would be nice if I ate.'

Optative

Poudel (2008) simply notes that this verb tense is more amenable to ergative marking

than the simple present in the following context:

- (44) a. *us-le tyo sarpa mār-os*
he.OBL-ERG that snake kill-OPT.3.SG
'(I wish) he could kill that snake.'
- b. *u tyo sarpa mār-cha*
he that snake kill-PRES.3.SG
'He will kill the snake.'

However, Poudel does not mention whether the ergative marker is optional in the first case and whether it correlates with a meaning difference. As I discuss in the Observation section, my consultants found the ergative to be optional in both cases.

Definite Future

Abadie (1974) finds the assertion of Clark (1963), that *-le* is excluded in this tense, to be untrue. Poudel (2008) contrasts the usage of the ergative with the definite future and the nominative form with a future reading of the simple present:

- (45) a. *jiban-le bholi ganit padh-āu-ne chan*
jiban-ERG tomorrow math read-CAUS-DEF.FUT PRES.3.HON
'(It is certain that) Jiban will study math tomorrow.'
- b. *jiban bholi ganit padh-āu-chan*
jiban tomorrow math read-CAUS-PRES.3.HON
'Jiban will study math tomorrow.'

The definite future is not an inflected tense but rather a periphrastic construction with an auxiliary and the verb marked with the *-ne* infinitive.

Modal Constructions

In the introduction section I noted the alternations between ergative and accusative case on the subjects in modal constructions of obligation and necessity. These constructions utilize the verb *parnu*. Abadie argues that the semantic difference here is one of internalized necessity rather than obligation:

- (46) a. *ma-lāi māntri hu-nu par-cha*
 I-ACC priest be-INF must-PRES.3.SG
 'I should be a priest.' (Abadie 1974: 172)
- b. *maile māntri hu-nu par-cha*
 I.ERG priest be-INF must-PRES.3.SG
 'I had to be a minister.' (Abadie 1974: 172)

In the first case, there is an implication of social pressure from others, but with the ergative the pressure comes from within. Abadie discusses ergative marking with several other constructions, and argues that the ergative tends to be required in constructions that have to do with obligations.

3.1.8 Multiple Factor Analysis

Verbeke (2013) argues that none of the analyses given so far fully account for the given data. While perfectivity, focus, individual-level predication, or animacy can be invoked to explain the usage of the ergative in a particular case, none of them can explain every case. There are always exceptions. It is thus likely that there are a range of features which correlate probabilistically with the expression of ergativity.

Verbeke and De Cuypere (2015) run a statistical analysis on a corpus of 355 Nepali sentences compiled from the CRULP corpus (www.crulp.org) with additional samples from published Nepali short stories. They examine the following nine factors:

- (1) Animacy of the subject
- (2) Pronominality of the subject
- (3) Person of the subject
- (4) Honorific marking on the verb
- (5) Presence of the subject argument (whether it is overt or covert)
- (6) Type of the clause (nominal or clausal)

- (7) Marking of the O (whether the object is marked or not)
- (8) Stage/Individual-Level interpretation of clause
- (9) Verb Tense

Comparing two statistical models (a logistic regression analysis and a classification tree model), they find that no single factor in itself accounts for most of the observed variability. They find evidence for four significant variables: Subject Animacy, Tense, Honorificity, and Subject Person. The ergative marker is positively associated with inanimate subjects (the strongest finding), the definite future tense, and non-local subjects. It is negatively associated with honorificity.

Verbeke and De Cuyper argue that these relate to two main functions of the ergative marker: to emphasize the agent role, and to mark semantic features of the verb. They conclude that “the use of an ergative marker in imperfective constructions is not determined by one particular factor alone, but instead motivated by various preferences that operate simultaneously.”(Verbeke and De Cuyper 2015: 19)

3.1.9 Summary

The picture that emerges from this literature is a lack of consensus on a single primary factor or set of factors that correlate with the expression of ergative where it is variable.

In imperfective transitive clauses, the ergative appears to correlate with certain **Properties of the Subject:** Animacy and Person. Ergative marking is more common if not required on inanimate subjects, and is less common on 1st and 2nd person subjects. It may be the case that the ergative emphasizes the agency or volitionality of the subject, because it is associated with interpretations of the indefinite future and in models of obligation with situations in which the subject has more control of their situation.

In imperfective transitive clauses, the ergative appears to correlate with certain

Properties of the Event: Ergative marking seems to correlate with individual-level predicates, or predicates that are interpreted as more complete, including the definite future and all tenses with perfective aspect.

In intransitive clauses, the ergative is completely absent for unaccusatives but present on many unergatives. In the unergative domain there is considerable variability between lexical items, in different tenses, and even among speakers. Ergative marking is more common on atelic verbs.

One way to adequately characterize the patterning of ergativity in Nepali might be to list each of the these factors independently, and perhaps make observations about their strength in relation to each other. For example, do considerations of animacy trump considerations of semantics in particular tenses? However, to my mind this approach ignores the extraordinarily complex interweaving of these factors. How do properties of the subject affect properties of the verb and vice versa? There is a subtle and as yet unexplained relationship between individual-level predication, perfectivity, and ergative marking. Furthermore, we see nearly all of these factors at play to a greater or lesser extent in other ergative languages, as we see in the next section.

3.2 General Theories of Optional Ergativity

3.2.1 Overview

There are four main characterizations of optional ergativity as it appears in the descriptive literature.¹⁰ Many of these theories intersect and overlap with each other.

10. McGregor divides the theories into five categories: Discriminative, Pragmatic, Semantic, Global, and Semiotic. Here I am not discussing the Global category, in which the ergative is correlated with sociolinguistic factors like formality, status, gender, and dialect. I do believe that Optional Ergativity in Nepali can (and most likely does) correlate with these factors, particularly dialect and formality. Grierson (1904a) attributed optional ergativity to Tibeto-Burman influence and stated that it is only found in colloquial and non-literary Nepali, and I have repeatedly encountered the assumption that the high caste and educated do not use the ergative in this way (which

Discriminative Theories argue that the ergative is used particularly when it is needed to discriminate or disambiguate the different arguments in a clause. Theories based on **Transitivity** take the ergative as a marker of high transitivity in a clause. **Markedness Theories** take the ergative to mark atypical or unexpected subjects. Theories based on **Prominence** argue that ergative marking makes salient the referring entity in the discourse.

3.2.2 Terminology

In the literature on split-ergativity, there is a proliferation of terms used to describe the various case marking patterns. Many of these terms have theoretical implications built into them, which may or may not be explicitly acknowledged, and are often used in slightly different ways by different researchers. In this section I give a brief overview of these terms and their generally accepted meanings, and I will attempt to justify the precise terminology that I advocate for in this dissertation.

Ergative Case

The term “ergative” itself is controversial, because it implies the existence of a pattern of case-marking in which S_t are marked in opposition to O and S_i . As we have seen, objects in Nepali sometimes receive their own accusative marking, and intransitive subjects sometimes receive ergative marking. So the opposition is not a strong one. Perhaps “Agent Marking” or “Subject Marking” would be better terms.

Recall that some readers for students used the term “Instrumental” for this case, which has the effect of (a) ignoring a distinction between the ergative *-le* and the instrumental *-le*, and (b) emphasizing the genetic relationship to Sanskrit, that is, the historical origin of the ergative case. Grierson, working on a typology of Indo-Aryan languages before the term ergativity was coined, uses “Agent Case.”

may or may not be true). I believe that these other factors arise as a result of my analysis.

I will continue to use the term ergative to describe any usage of the *-le* marker on a core argument in Nepali, because the Nepali pattern is one particular manifestation of a vast typology of subject-marking patterns in Indo-Aryan languages that ultimately derive from an alignment shift that occurred in the history of the family. The different patterns in the modern Indo-Aryan languages of India and Nepal (arguably) all derive from earlier patterns which were more like canonical ergativity, and referring to the Nepali pattern as ergative connects it with an intimately related phenomenon that is known as ergativity in the literature of hundreds of other South Asian languages.

Typology of Split Case Marking systems

Split Case Marking systems are those in which case-marking on an argument is differentially marked in separate grammatical domains. Typically we are talking about accusative marking on the O and/or ergative marking on the S_t. A split is conditioned either by the nature of the clause (tense, aspect, mood, or main/subordinate status) or the semantics of the marked NP (animacy, pronominal status, person). Nepali has both a **Split Ergative** system with *-le* and a *Split Accusative* system with *-lāi*.

Another frequently-used term is **Differential Case Marking**, which also refers specifically to systems in which the case marking is differentially marked in separate semantically-defined domains. **Differential Object Marking** refers to any system in which there are multiple ways to mark objects in a language, but as used in Aissen (2003) and Fauconnier (2011) it frequently refers to systems in which a particular marker may be present or absent as in Nepali. For the variable case marking on the *subject* of a transitive clause, these authors use the term **Differential Subject Marking** (a term preferred by Verbeke and De Cuyper (2015) for Nepali). Conversely, the term **Differential Agent Marking**, specifically refers to the pattern of marking the S_t of transitive clauses. This excludes the S_i of an intransitive clause. Usage of the term DAM over DSM emphasizes that ergative marking on the S_i is

absent in the language, or else that it should be treated theoretically as a separate phenomenon. Also, McGregor (2010) uses the term **Differential Ergative Marking** specifically to refer to systems in which there are multiple overt ergative markers which are used in separate domains (distinguishing systems in which a marker varies with its absence).

In all the systems discussed so far, the split is conditioned by semantic factors. In other systems (as in many Tibeto-Burman languages), a marker may be absent or present without affecting the meaning of the clause. These are **Optional Case Marking** languages, and as regards the usage of the ergative marker they are termed **Optional Ergative Marking** languages. Here the expression of case is correlated with pragmatic factors, and the occurrence of a particular case marker is not wholly predictable from the grammatical context. Its usage does not depend only on the semantic interpretation or truth conditions of the clause.

In practice, the boundaries between these systems are not always clear. Aissen (2003) focuses mostly on the semantic factors that condition DOM, but notes that there are realms of “optional” marking. Not all languages clearly fit into one system or another: Nepali is a mixed system which shows patterning consistent with DSM in the perfective transitive domain and OEM outside of it.

Wherever it is defined, OEM is often followed by a caveat that the term “optional” is misleading, because it does not indicate the form is in free variation with its absence, but rather that it is pragmatically-conditioned (Chelliah et al. 2011, DeLancey 2011). I will refer to the general phenomenon as OEM or OCM, and in Nepali I will refer to nominative/ergative alternations as variable ergativity or variable case marking.

3.2.3 Discriminative Function

McGregor (2010) notes that this is perhaps the most frequently invoked explanation for Optional Ergative Marking. Abadie (1974) appeals to a discriminative function

for Nepali in arguing that the ergative is used when it is necessary to disambiguate the arguments. Essentially, ergative marking is employed if there is a chance of confusing participant roles. That is to say, the ergative is used if it is necessary to specify which participant is causing the action and which is affected by it. This is a functional explanation for ergative patterning that is given in some form or another as a general explanation for ergativity by Dixon (1994), Comrie (1978), and Garrett (1990).

Keenan (1984) offers support for a functional theory of ergativity from Anderson (1977)'s observation that ergativity is largely restricted to languages with verb-initial or verb-final canonical word order. Keenan (1984) further notes that for these languages, word order by itself is not enough to discriminate the two major participants in cases of subject deletion or fronting. This is not the case for languages with verb-medial word order.

In a canonically SVO language, for example, deletion of the subject or the object (SV or VO) does not lead to ambiguity. Word order alone can distinguish subject deletion from object deletion. Likewise, if the subject is fronted the word order remains SVO, and if the object is fronted (OSV) this is the only configuration with two overt arguments preceding the verb. In a canonically SOV language, however, deletion of the subject or the object leads to potential ambiguity (OV or SV). If the object is fronted there is again a potential ambiguity between SOV and OSV. Keenan argues from this that the perceptual function of ergative case-marking is to discriminate the subject argument from the object argument.

In the context of Optional Ergative Marking, a number of separate observations are explainable in terms of a discriminative function. “Unexpected” or “Unlikely” subjects tend to be marked because there is a greater possibility of confusion. For example, Foley (1986) observes a “disambiguating function” in the ergative marking of many Highland languages of Papua New Guinea. In Dani, ergative marking is necessary in a sentence like *The pig ate him* because this is the opposite of the

expected relation between these two animate referents.

Similarly, nouns with inanimate reference are generally unlikely to be subjects in many languages, and this could explain why many languages require ergative marking on inanimate subjects. This notion of marking less typical subjects is an intrinsic feature of markedness-based theories, which I discuss in detail in the Markedness section below.

I would also add that the potential for confusion is only relevant in the absence of other mechanisms in the language to disambiguate participants. Languages differ in the extent to which verbal agreement with the subject (or object) also serves to disambiguate participants. Nepali robustly marks person features on the verb. So there is never any ambiguity when the subject is local (1st or 2nd person), and a discriminative function could explain why ergative marking is less common on local subjects.

The problem with a purely discriminative description of optional ergativity is that it can be at best only a partial explanation for a larger pattern of case-marking. This is noted by Du Bois (1987) and McGregor (2010). McGregor (2010: 1618) in particular observes that discriminative explanations of optional ergativity in grammars typically provide an example or two in which there would otherwise be ambiguity. But they generally do not address whether optionality is present when there is no possibility of confusion. McGregor gives the example of Yuwaalaraay in the grammar Williams et al. (1980), but this is relevant for Nepali as well. As we have seen, optionality is present in intransitive clauses, and when both the subject and the object are overt and in their canonical order, and when there is no possibility of ambiguity due to verbal agreement in number, gender, or person.

3.2.4 Transitivity

A clause with at least two core participants is transitive and a clause with a single core participant is intransitive. Transitivity, however, can be thought of as a gradient concept that interacts with the expression of ergative case. Below I discuss two influential theories, which are different ways to formulate the concept of transitivity as it relates to the grammatical expression of arguments. The first theory is Hopper and Thompson (1980)'s notion of Transitivity, and the second is Dowty (1991)'s Proto-Roles.

The Transitivity Hypothesis

Hopper and Thompson (1980) propose that **transitivity** is a grammatical property which describes the effectiveness with which an action takes place. Transitivity is a gradient concept, and whether a particular clause has high or low levels of transitivity is determined by the aggregation of multiple features. These features, which are listed below, all relate to the “effectiveness or intensity with which the action is transferred from one participant to another” (Hopper and Thompson 1980: 252):

	High Transitivity	Low Transitivity
Participants	2 or more (S_t and O)	1 participant
Kinesis	eventive predicate	stative predicate
Aspect	telic or perfective	atelic or imperfective
Punctuality	punctual	non-punctual
Affirmation	affirmative	negative
Mode	realis	irrealis
Volitionality of S_t	S_t volitional	S_t non-volitional
Agency of S_t	S_t high in potency	S_t low in potency
Affectedness of O	O totally affected	O not affected
Individuation of O	O highly individuated	O non-individuated

Figure 3.3: Hopper and Thompson (1980)'s Transitivity Prototype

In Figure (3.3) have rearranged the features in the original chart from Hopper and Thompson (1980: 252) to point out that these features broadly fall into three categories: the first five features relate to the features of the event itself, the following two relate to features of the S_t argument, and the final two relate to features of the O. Hopper and Thompson argue that every one of these features is related to the effectiveness by which an event is initiated by one participant and has an effect on a second participant, and the overall transitivity is the result of all of these features taken together. I constructed the following examples to follow a cline from high transitivity to low transitivity:

- (47) a. Cain murdered Abel.
- b. A grizzly bear is destroying the gazebo.
- c. Samuel fears mongooses.
- d. There might not be any more pie.

An event generally requires two or more **participants** to effectively transfer an action, as in (47a), (47b), and (47c). A predicate that describes an **event**, like *murder*, is more clearly transitive than a stative predicate like *fears*. Furthermore, *murder* is **punctual**, meaning that it is generally interpreted as an event that occurs without duration, whereas *fears* and *is destroying* are non-punctual. Under the Aspect category, Hopper and Thompson initially refer to **telicity**, which is a property of the lexical semantics of a predicate. The predicate *destroy* is telic because it has a distinct endpoint, while *fears* does not. However, Hopper and Thompson also consider verbs with **perfective** reference, as in (47a), to have high transitivity,. In either case, the presence of a definite endpoint is considered to be a mark of high transitivity. And a clause which describes an actual event is more transitive than one that describes a hypothetical event or the negation of an event (**mode** and **affirmation** respectively). (47a) has all the features described for the event, (47d) has none, and the others have

some and not others.

For the subject, **volitionality** refers to the subject's intention to carry out the action, and **agency** refers to its ability to do so. Volitionality and Agency mean that the S_t often has animate (if not human) reference. Cain has both the intention and ability to kill Abel (47a). In (47b) the bear may or may not have the intention but certainly has the ability to destroy the gazebo, but Samuel in (47c) does not intentionally fear mongooses, and the inanimate subject of (47d) cannot really be ascribed any volitionality or agency whatsoever.

For the object, **affectedness** and **individuation** are Hopper and Thompson's relevant criteria. A highly transitive event has an effect on the object, most clearly when it comes into being or is destroyed as a result of the event. Only in (47a) is the object dramatically affected by the event. Furthermore, the object is highly individuated, which is further defined by Hopper and Thompson as having the following properties:

Individuated	Non-individuated
proper	common
human, animate	inanimate
concrete	abstract
singular	plural
count	mass
referential, definite	non-referential

Hopper and Thompson argue that an action is most effectively transferred to a patient when it is individuated. A proper noun like *Cain* is more individuated than a common noun like *gazebo* or *mongoose*. We focus more on the effect on humans and animate nouns rather than inanimates like *pie*. Furthermore the effect on a singular noun like

the gazebo is more clearly differentiated than an effect on plural nouns or mass nouns like *pie*, and the effect is stronger when the object is definite rather than indefinite.

A particular clause may have some or all of the features of high transitivity, and taken together these determine how transitive the clause is. The predictions that fall out from this analysis lead to the **Transitivity Hypothesis**:

If two clauses (a) and (b) in a language differ in that (a) is higher in Transitivity according to any of the features..., then, if a concomitant grammatical or semantic difference appears elsewhere in the clause, that difference will also show (a) to be higher in Transitivity. (Hopper and Thompson 1980: 255)

Note that the association between features can be in terms of morphosyntactic forms which are associated with a particular feature or semantic features that are associated with a particular feature. In Nepali, the restrictions on ergative marking provide the best illustration of this hypothesis. In fact, Hopper and Thompson write that there is “a rather typical situation in ergative languages: the canonical ERGATIVE clause signals one, several, or all of the high-Transitivity features, while the ANTI PASSIVE [they use the term to include nominative-accusative patterning in a split system] clause signals one or more of the low-Transitivity features”(Hopper and Thompson 1980: 268) Consider the domain in which ergative marking is required in Nepali:

- (1) There must be an S_t and O (transitive clause) - PARTICIPANT
- (2) The predicate must have perfective aspect - ASPECT

Here we have two features associated with transitivity: ergative marking (indicative of transitive clauses, thus high transitivity) and the perfective aspect of the verb. In the language there is an obligatory association between these two morphosyntactic forms. The Transitivity Hypothesis correctly predicts that both are associated with high

values of transitivity. A counterexample to the Transitivity Hypothesis would be any language in which the ergative case marker is only obligatory in imperfective clauses and disallowed in perfective clauses, because then there would be an association between two features with conflicting values. Now consider the domain in which ergative marking is disallowed in Nepali.

- (1) There must be only one argument (intransitive clause) - PARTICIPANT
- (2) The predicate must be stative - KINESIS
- (3) OR: the S_t must be non-volitional and low in potency - VOLITIONALITY, AGENCY

If intransitive, ergative marking is disallowed on stative predicates, and also on unaccusative intransitives, which have the defining property of having theme subjects which tend to be non-volitional and low in potency. Here we have a cluster of features that constitute obligatory nominative case on the subject, and all of them are associated with low transitivity, as predicted by the Transitivity Hypothesis.

The Transitivity Hypothesis crucially does not make predictions about which of these features will be linked in any particular language. But if they are linked, the prediction is that high transitivity features pair with high transitivity features and low transitivity features pair with low transitivity features. In some languages, as we have seen, ergative marking is associated with highly volitional subjects. In Hindi, we find that intransitive subjects have a volitionality alternation, with the ergative associated with intentional actions. But that does not mean there needs to be such an association in Nepali, and indeed there does not appear to be.

So to an extent, Hopper and Thompson have collected together many of the features that we find associated with ergativity among the world's languages and provided a straightforward explanation for the common association between, for example, ergativity and perfective aspect. Furthermore, their notion of individuation in objects provides an explanation for many of the facts about differential object mark-

ing in languages which have, like Nepali, an association between object marking and human, animate, and definite referents. In some nominative-accusative languages we find the reverse case to that of the ergative, in which O-marking associated with low transitivity leads to an imperfective interpretation of the event. Hopper and Thompson give this example from Finnish, in which placing the object in partitive case (giving it a less individuated reading) leads to an imperfective interpretation:

- (48) a. *liikemies kirjoitti kirjeen valiokunnalle*
businessman wrote letter.ACC committee-to
'The businessman wrote a letter to the committee.' (Fromm and Sadienemi 1956:120-21)
- b. *liikemies kirjoitti kirjettä valiokunnalle*
businessman wrote letter.PART committee-to
'The businessman was writing a letter to the committee.' (Fromm and Sadienemi 1956:120-21)

It is tempting to consider all the facts about Nepali ergativity we have looked at so far under this rubric. Then ergativity would simply be a marker of transitivity, and every language with ergative marking develops associations between the ergative and some unpredictable assortment of these high transitivity features (or else the nominative develops associations with the low transitivity features). Over time new feature associations might be introduced. Associations may also shift from optional associations to obligatory associations. Perfectivity and transitivity may once have been the hard requirements for ergative case, but now we find those associations weakening such that ergative case is also possible in their absence if multiple other features conspire to make the clause unusually transitive. So an intransitive clause may get ergative marking if the S_t is volitional (as in an unergative). But ultimately it is impossible to predict which of these features specifically will tend to cluster together in any particular language.

But there are some problems with this notion. Hopper and Thompson's Transi-

tivity Hypothesis does not technically make any predictions about ergative marking when it is “optional” as in Butt and Poudel (2007)’s association between ergativity and individual-level predication. And some of these associations in fact make the wrong predictions.

To take the individual-level predication example, the interpretation of a predicate as individual-level is typically atelic, while the stage-level interpretation is not. Compare *I drive buses* as an ongoing occupation with *I will drive a bus today*. If ergative marking in Nepali is associated with individual-level interpretations, then we have a high transitivity feature (ergativity) paired with a low transitivity feature (atelicity).

Secondly, if we are to interpret Li (2007)’s use of the term telicity as equivalent to the way the term is used in Hopper and Thompson, then we have another apparent counterexample to the Transitivity Hypothesis. Li argues that, among Nepali intransitive unergative predicates, there is a split between those which are telic (like *āunu* ‘come’) and those which are atelic (like *nācnu* ‘dance’). The telic verbs (which are a feature of high transitivity) are specifically those which require the nominative (a feature of low transitivity), and it is optional elsewhere.¹¹

Even more problematic than the previous examples is the apparent association between animacy and ergativity, which goes the opposite direction we would expect: inanimate subjects are more likely to be marked with the ergative, even though they are low in agency (a feature of low transitivity). This will be relevant to the discussion of markedness below.

Argument Proto-Roles

The influential formulation of Dowty (1991) provides an approach to many of these same issues from a different perspective, that of the relationship between grammatical

11. One way around this problem would be to reject the notion that telicity is the relevant factor here, but rather that the reason that verbs like *āunu* ‘come’ and *jānu* disallow ergative marking is that they are in fact unaccusative in Nepali. This will be my argument in section .

structure and argument roles. Dowty notes that the traditionally proposed thematic roles (Agent, Theme, Patient, Experiencer Instrumental, etc.) have unclear boundaries and thus a multiplicity of definitions and formulations. Rather than defining a set of discrete thematic roles, Dowty introduces two proto-roles, the **Proto-Agent Role** and the **Proto Patient Role**. He approaches thematic roles from the perspective of argument selection, meaning that roles are defined by the set of entailments on a group of predicates with respect to one of their arguments (Figure 3.4).

Agent Proto-Role	Patient Proto-Role
volitional involvement in the event or state	undergoes change of state
sentience (and/or perception)	incremental theme
causing an event or change of state in another participant	causally affected by another participant
movement (relative to the position of another participant)	stationary relative to the movement of another participant
exists independently of the event named by the verb	does not exist independently of the event, or not at all

Figure 3.4: Dowty (1991)'s Proto-roles

The proto-role is a cluster concept, which means that any particular argument may exhibit some but not all of the features that define the proto-role. A particular argument will have properties of the proto-role to a greater or lesser extent. In this way, Dowty's approach is similar to Hopper and Thompson's conception of Transitivity, because in both cases any particular form will fall along a cline from more to less prototypical. Dowty writes the following in justification of cluster concepts for thematic roles:

Discrete feature decomposition has its proper place in syntax, morphology, and phonology, because these domains are aspects of the ‘coding system’ of language at various levels and therefore in principle discrete. But semantic distinctions like these entailments ultimately derive from distinctions in kinds of events found ‘out there’ in the real world: they

are natural (physical) classifications of event, and/or those classifications that are significant to human life. There is no reason to believe that all such classes must have discrete boundaries. (Dowty 1991: 575)

This leads to a prediction about which arguments will surface as the grammatical subject and grammatical direct object for a particular predicate in a language:

Argument Selection Principle: In predicates with grammatical subject and object, the argument for which the predicate entails the greatest number of Proto-Agent properties will be lexicalized as the subject of the predicate; the argument having the greatest number of Proto-Patient entailments will be lexicalized as the direct object. (Dowty 1991: 576)

As an illustration of Agent and Patient Proto-Role entailments, consider how the following sentence is expressed in English and Nepali:

- (49) *sangita-le tyo citthi haat-le lekh-i*
sangita-ERG that letter hand-INSTR write-PERF.3.SG.F
'Sangita wrote that letter by hand.' [BB]

The arguments of the predicate of (49), *write*, have nearly all of the entailments for Agent and Patient Proto-roles respectively. *Sangita* intentionally causes the event which leads to the change of state of the argument, and she exists independently of this event. 'That letter' is created as a result of the event, undergoing a definite change of state as a result of the action of the other participant. It is also an **Incremental Theme**, a term Dowty introduces to describe the situation in which subparts of the event correspond to subparts of the argument. In other words, if the letter writing process is halfway complete, then the letter is halfway finished. The letter isn't completely written until the event of letter writing is complete. This can be contrasted with "push the cart," which does not have this property.

The verb ‘write’ is a good example of a verb for which the arguments follow most of their respective entailments, and so languages are expected to lexicalize the arguments in a stable and predictable way: with *Sangita* as the grammatical subject and “that letter” as the grammatical direct object. Under Hopper and Thompson’s formulation this clause is highly transitive, but for Dowty the arguments are very clear examples of Agent and Patient Proto-roles. Now contrast this with the following sentences in English and Nepali:

- (50) *ketā-lāi ekdam.ai dukha lāg-yo*
 boy-ACC very.EMPH sad feel-PERF.3.SG
 ‘The boy was very sad.’ [AG]

“The boy” in this construction has an Experiencer theta role. These types of constructions are known to be lexicalized in different ways in different languages, and for having multiple lexicalizations within the same language (compare, for example, the English perception constructions “I like this”/“This pleases me”, “I fear this”/“This frightens me”, etc.).¹² This is explicable in terms of Proto-roles: “sadness” and the “boy” have about an equal claim to Agent and Patient Proto-role: “the boy” is sentient and perceptive, but undergoes a change of state and is causally affected by the event, while “sadness” causes the change in the other participant but is nonvolitional and non-sentient. When there is no clear differentiation between Agent and Patient Proto-role, then either or both may be lexicalized as the subject/object in a given language; some languages like English take “the boy” to be a subject, and other languages like Nepali take “the boy” to be an object.¹³

Many of the terms used for specific theta roles can be defined by which of the Agent or Patient Proto-role properties are entailments (Dowty 1991: 577). Thus:

12. As discussed in (Postal 1970), among others.

13. More precisely, the Nepali construction treats the subject *ketā-lāi* (boy-ACC) as an object in some ways (obligatory accusative marking) and a subject in other ways (the typical word order is Experiencer-Stimulus-Verb, suggesting that the experiencer is a subject given S-O-V word order. The object *dukhā* cannot take ergative case.

- (51) AGENT = [+Causation] or [+Volition]
 EXPERIENCER = [+Sentience, -Volition, -Causation]
 THEME (typically) = [+Change +Incremental-Theme, +Dependent-Existence,
 +Causally-Affected]
 PATIENT = [+Change +Incremental-Theme, +Dependent-Existence, -Causally-
 Affected]
 INSTRUMENT = [+Causation, +Movement, -Volition, -Sentience]

Dowty emphasizes that Proto-roles are separate from grammatical forms, and it is clear that the Nepali grammatical postpositions *-lāi* and *-le* do not straightforwardly correspond to arguments with Patient and Agent Proto-roles. If *-le* simply marked the “Agent Proto-role” of the sentence, then it would not be able to vary with the nominative.

However, these grammatical postpositions may lexicalize specific properties of the Proto-roles. Note that *-lāi* is optional on direct objects but obligatory on indirect objects (as in *ma sitā-lāi kitā dinchu* “I give Sita a book”). Perhaps then *-lāi* marks the object which has the properties associated with being a source (-Movement), which is necessarily the case for indirect objects.

For the ergative, this raises the possibility that *-le* marks the same features whether it is on an instrumental argument or as an ergative. This is because an instrument has a subset of the properties of a subject: it includes causation and movement but not volition and sentience. This is the approach I will pursue in chapter 6 below.

Causal Structure

Dowty directly compares his approach to the Transitivity approach of Hopper and Thompson:

Hopper & Thompson view transitivity as a property that a CLAUSE can possess to a greater or lesser degree, whereas I think the transitivity of a clause can be derived by summing the independently needed agentivity and patientivity counts of the arguments. (Dowty 1991: 599)

These two approaches deal with many of the same morphosyntactic structure and semantic features. Dowty writes that “the meaning of a telic predicate is a homomorphism from its (structured) theme argument denotations into a (structured) domain of events”(Dowty 1991: 567). This means that the denotation of an object may determine the aspect of the clause, as is fundamentally the case for an Incremental Theme. This interdependence of object and verbal aspect, as discussed in Verkuyl (1972), is present in the following examples:

- (52) a. Janeane drank the beer. ('perfective')
b. Janeane drank beer. ('durative')

When the object is definite, the clause is interpreted as perfective; Janeane drank the entire beer. When it is indefinite, the aspect is durative; she drank some unspecified amount of beer. Dowty conceptualizes the difference in terms of the properties of the arguments (Dowty 1991: 567). In (52a) “the beer”, being definite, is interpreted as an incremental theme, with the subparts of the event of drinking being mapped onto the subparts of the amount of beer drunk. Thus in (52a) there is a telic interpretation rather than the atelic, durative interpretation that arises with the indefinite mass noun “beer” in (52b).

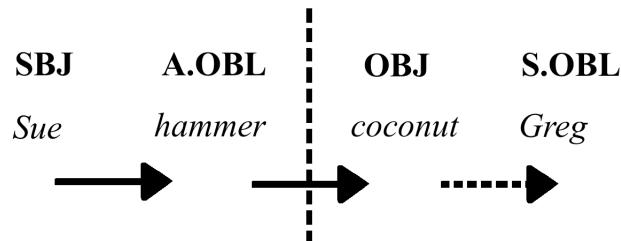
Under Hopper and Thompson’s analysis, (52a) is higher in transitivity than (52b) because “the beer” is a more individuated object than “beer.” One feature of high transitivity (an individuated object) is obligatorily correlated with another feature of high transitivity (telic interpretation of the clause).

These two approaches represent two ways of looking at the same gradient concept

of transitivity, which is expressed in full by the grammatical machinery associated with arguments as well as that of predicates. With regards to Nepali, Anderson (1977) argues that the ergative case in modern Indo-Aryan languages is part of the machinery for representing verbal aspect.¹⁴ This intuition was shared by Verbeke (2011) for Nepali, who argued that ergative marking imparts a sense of completedness to an otherwise imperfective clause.¹⁵

Croft (1991)'s notion of causal structure unifies these two perspectives of argument proto-roles and prototypical transitive events. Croft (1991: 197) takes the fundamental semantic property for determining argument realization to be the causal structure of events, conceived as the transmission of force between participants along a causal chain.

- (53) Sue broke the coconut for Greg with a hammer. (Adapted from Croft 2012: 214)



Each participant in the clause is associated with a particular subevent, and these subevents are linked in a causal chain representing transmission of force from the

14. "Normally we think of verbal categories such as tense and aspect as marked on the verb, and not (partly) in the NP, but this is by no means necessary." (Anderson 1977: 336)

15. This event-argument duality is reminiscent of the wave-particle duality in quantum mechanics:

But what is light really? Is it a wave or a shower of photons? There seems no likelihood for forming a consistent description of the phenomena of light by a choice of only one of the two languages. It seems as though we must use sometimes the one theory and sometimes the other, while at times we may use either. We are faced with a new kind of difficulty. We have two contradictory pictures of reality; separately neither of them fully explains the phenomena of light, but together they do. (Einstein and Infeld 1961: 262-263)

initiation of an event to its endpoint. Thus Sue acts on the hammer by grasping it, the hammer acts on the coconut by hitting it, the coconut is affected by the event in changing its state from broken to unbroken, and the event ultimately benefits Greg. The verb directly profiles the portion of the event that begins when Sue initiates it, the hammer (as an instrumental) enacts the event, and it ends with the effect on the coconut (represented by the solid arrow). The oblique argument “for Greg” profiles the further effect that the event has on Greg (represented by the dashed arrow).

The rules for argument realization relate to the position of arguments along this causal chain. As with Dowty, the manifestation of a particular argument in a particular role relates to gradient properties of two macro-roles (Subject and Object). Croft’s rules have the advantage of also making predictions for **antecedent obliques**, which are non-core arguments implicated in transmitting the force, and **subsequent obliques**, which are implicated in receiving it.

(54) Argument Linking Rules (Croft 2012: 207)

- a. The verbal profile is delimited by Subject and Object (if any)
- b. Subject is antecedent to the Object in the causal chain:

$$\text{SBJ} \rightarrow \text{OBJ}$$
- c. An Antecedent Oblique is antecedent to the Object in the causal chain; a Subsequent Oblique is subsequent to the Object in the causal chain:

$$\text{A.OBL} \rightarrow \text{OBJ} \rightarrow \text{S.OBL}$$
- d. Incorporated arguments are between Subject and Object in the causal chain:

$$\text{SBJ} \rightarrow \text{INCORP} \rightarrow \text{OBJ}$$

This conception of argument realization in terms of causal structure is useful for our purposes because it decomposes the properties of arguments as they relate to a prototypical transitive event. In particular, a prototypical subject is an entity

which instigates, enacts, and completes an event which has an effect on an object. An instrument is an antecedent oblique which shares some (but not all) of these properties: it is involved in enacting (and possibly completing) the event, but not in instigating or controlling it.

This forms the basis of my theory of ergative marking in Nepali: the *-le* marker has the same meaning as an instrumental or as an ergative, and that meaning can be described as an **Effector** of the event. Under this formulation, a prototypical subject is both an Effector and an Instigator, but an instrument is just an Effector. It profiles the enacting of the event but not its instigation. I discuss this more thoroughly in section 6.1.

This usage of the term Effector is somewhat similar to its usage in Role and Reference Grammar, but the difference is that it does not necessarily represent a rarified semantic role (Van Valin and Wilkins 1996). Van Valin and Wilkins (1996) emphasize the distinction between properties of an Agent and properties of an Effector, and argues that Agent properties are pragmatic in nature. Furthermore, Holisky (1987) has a similar but inverse analysis for ergative/nominative alternations in Tsotva-Tush. Ergative marking in this language correlates particularly with properties of the Agent (volitionality, instigation), but *not* the Effector properties.

Croft (2012) further refines the causal chain by introducing aspectual structure to each subevent. The result is a three-dimensional representation of causal-aspectual structure along the causal chain. I employ a simplified version of Croft's causal structure in discussing the relationship between ergative marking and aspectual properties of the event in chapter 6.4.

This unified conception of transitivity and argument realization is an important component of the analysis in chapter 6. A second important component is the mechanism of semantic markedness.

3.2.5 Markedness

The concept of **markedness** was developed under the structuralist theory of the Prague School, and the concept has taken on a wide and varied set of meanings and functions across functionalist, generative, and cognitive theories of linguistics (Battistella 1996, Horn 2001, De Lacy 2006). The terminology was initially applied to phonology in Trubetzkoy (1931), and to semantics by Jakobson (1932) [Jakobson (2011a)]. Markedness is an important theoretical underpinning of early work on ergativity in Dixon (1972, 1979) and Silverstein (1976).

The fundamental observation of markedness is that oppositions in language and cognition tend to be asymmetrical. Opposing values are not of equal status. Roman Jakobson describes the terms in the following way:

One of two mutually opposite categories is “marked” while the other is “unmarked.” The general meaning of a marked category states the presence of a certain (whether positive or negative) property A; the general meaning of the corresponding unmarked category states nothing about the presence of A, and is used chiefly, but not exclusively, to indicate the absence of A. (Jakobson 1957: 47)

Gender marking is a commonly cited example of semantic markedness. In a semantic opposition between masculine and feminine, the feminine is typically marked and the masculine unmarked. Thus for many animal pairs like *fox/vixen*, *lion/lioness*, *horse/mare*, *wolf/she-wolf*, the masculine form is the general term for the species as a whole. It is the form that is used when the gender of the animal is unknown or indefinite. It does not explicitly signal that the referent is masculine.

The feminine form by contrast can only refer to a feminine referent. Thus the feminine form has a more complex, narrower meaning. It is often the case that the

marked meaning is associated with a marked morphological form: for English, we find this with feminine affixes like *-ess*. In semantic markedness, a marked morphological form often varies with the absence of marking, or zero sign (Jakobson 1939, [Jakobson (2011b)]).

In Jakobson's terms, *vixen* is marked relative to *fox* because it signals the property [feminine]. But *fox* does not explicitly signal the opposing property [masculine]. Rather, it is unmarked (non-signalization of [feminine]). This is the **general meaning** (*Gesamtbedeutung*) of the unmarked term. Jakobson also notes that there is usually, but not always, a **narrow meaning** (*Sonderbedeutung*) of the unmarked term: signaling the absence of the given property. In this case, because it is not feminine, *fox* is interpreted as masculine. Waugh (1982) refers to these as the “zero-interpretation” and the “minus-interpretation” respectively.¹⁶

Jakobson was also the first to apply markedness theory to case in his analysis of the Russian case system (Jakobson 1936, [Jakobson (2011a)] Jakobson 1957, Chvany 1984). He considered the Nominative case in Russian to be unmarked relative to the Accusative. The relevant property in the Nominative/Accusative opposition is [directionality], in which the Accusative denotes the entity at which the action is directed, and the Nominative is unspecified. In a second opposition, the Nominative case is unmarked relative to the Instrumental. Here the relevant property is whether or not the denoted entity is peripheral to the action. Jakobson expressed the eight Russian cases by whether they were marked or unmarked with respect to three features [marginal], [quantification], and [directionality].

He argued that for ergative languages the reverse is true, with the marked subject

16. The property signaled by the marked term is an entailment, whereas there is some ambiguity in whether the minus-interpretation of the unmarked term entails the opposing property or merely implicates it. The distinction is between a grammaticalized opposition (e.g. ergative marking in perfective Nepali clauses) and a variable pragmatic opposition (ergative marking in imperfective Nepali clauses). This will be discussed in detail in the Discussion section.

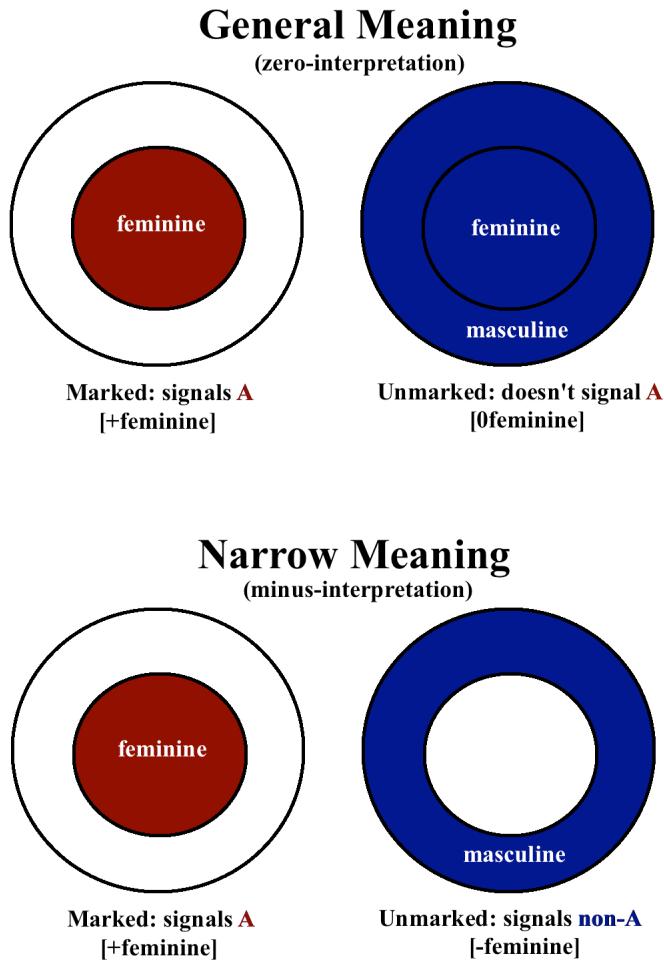


Figure 3.5: General vs. Narrow Meaning of Unmarked; based on Andrews (1990: 154)

indicating a property of affecting another entity with an action.¹⁷ For these languages the object is unmarked.

Markedness is frequently correlated with distributional frequency. The unmarked form is said to be more frequent within a language and/or within a particular gram-

17. The term ‘ergative’ was not yet in common usage. He refers to “languages (for example Basque and the Northern Caucasian languages) in which the aforementioned most prominent function of the N, that is, that of the subject of a transitive action, becomes the only function of that case.”(Jakobson 2011a: 71) As discussed in the preliminary section, early Western descriptions of alignment patterns often described the overall construction as passive, and referred to the case marking on S_t as ‘instrumental’ or ‘agentive.’

matical domain, and to be typologically more frequent. It is correlated with implicational universals, and in generative linguistics the unmarked is considered to be the default form and as such is linked with childhood language acquisition with the unmarked form being acquired earliest. There is considerable disagreement over the extent to which these multiple aspects of markedness should be considered definitional or epiphenomenal (Battistella 1996, Andrews 1990, Croft 2003) or even whether markedness is a useful concept to describe these phenomena (Haspelmath 2006).

Battistella (1996) defines markedness simply in terms of an opposition involving a distinctive property, in which markedness imposes an asymmetry by signaling the presence of that property for one end of the opposition, and not signaling the presence of that property at the other end. All of the other aspects of markedness arise from this basic asymmetry. Battistella summarizes those aspects which have figured into various theories of markedness (Figure 3.6).

Signalization of A	Nonsignalization of A
specific meaning	general meaning
conceptual complexity	conceptual simplicity
narrowly defined	broadly defined
syncretized	nonsyncretized
subset	superset
figure	ground
abnormal	normal
nonprototypical	prototypical
less frequent	more frequent
implying	implied
low valued	high valued
nonneutralizable	neutralizable
nonoptimal	optimal
overt expression	zero expression

Figure 3.6: Aspects of Marked and Unmarked Forms (Battistella 1996: 71)

Theoretical descriptions of split ergativity make use of three extensions of markedness theory. The first of these is the **Markedness Hierarchy**. While markedness

theory was originally formulated only in terms of binary oppositions, many grammatical phenomena can be conceived along a scale from unmarked to marked. Croft (2003) lists a few these hierarchies:

- (1) Number: SINGULAR < PLURAL < DUAL < TRIAL/PAUCAL
- (2) Grammatical Relations: SUBJECT < DIRECT OBJECT < INDIRECT OBJECT < OBLIQUE
- (3) Agency/Nominal: 1PRO/2PRO < 3PRO < PROPER NAMES < HUMAN S < NON-HUMAN ANIMATE < INANIMATE

The value on the left is unmarked, and markedness increases incrementally towards the right. The third of these examples will be particularly relevant to split-ergative languages in which the split is conditioned by the semantics of the NP. Croft notes that this extension to include hierarchies requires markedness to be a relative property rather than a binary one (Croft 2003: 111). However, a hierarchy can also be represented as a succession of multiple binary features, as in Silverstein (1976)'s original formulation of the Nominal hierarchy, which will be the topic of the next section.

The second extension of markedness theory that is relevant to our purpose is the concept of the **Prototype**. A prototype category is characterized by a collection of properties, but none of these properties alone is necessary or sufficient to distinguish a member of this category. Rather, there are core members of this category which have all or most of the relevant properties, and peripheral members which share fewer of these properties. The boundaries of a prototype are fuzzy or variable, but there nevertheless tend to be definitively nonprototypical entities which are unequivocally excluded (Croft 2003: 125).

We have already seen two examples of prototypes in the Transitivity section of this chapter. Hopper and Thompson (1980) define Transitivity as a prototype concept. By conceptualizing transitivity in terms of a collection of properties, they allow for

individual clauses to fall along a scale of transitivity. Dowty (1991)'s conception of Proto-roles, in which the Agent and Patient proto-roles are defined by a series of entailments assigned by the verb, is also a prototype conceptualization. For peripheral cases, in which a particular participant has an approximately equal number of Agent and Patient entailments, one finds crosslinguistic variation and alternate strategies for grammaticalizing the participant as subject or object within a single language. Thus any particular participant falls along a spectrum of Agent-ness or Patient-ness depending upon how closely they share the collection of relevant properties.

In the domain of participant case marking, it is necessary to consider the **prototypical subject** cluster concept and **prototypical object** cluster concept. As discussed below, there is debate over properties in these categories, particularly with respect to ANIMACY and DEFINITENESS.

The third relevant concept is the much-disputed role of **Markedness Reversal** in the description of marking patterns (Andrews 1990: 147). Markedness reversal is “the phenomenon whereby a marked context reverses the normal markedness values of the terms of an opposition.”(Shapiro 1983: 93) In certain contexts, the given markedness values are reversed.

Recall the example of gender marking in animal names. Note that in English, the terms for many common farm animals appear to be exceptions to the generalization that the feminine of the pair is marked and the masculine is unmarked.¹⁸ For *cow/bull*, *chicken/hen*, and *goose/gander* it is the feminine of the pair which is the term generalized to the species as a whole. This can be explained by context. The female of the species has a broader range of uses on a farm, being used for milk, eggs, and meat. Here it is the masculine property which is marked in relation to the unmarked feminine. Markedness is relative to what is expected in a given context, and this is mediated by social norms. Thus Waugh (1982) notes that [feminine] is the

18. This distinction was pointed out to me by Larry Horn (p.c.).

marked property for *doctor* but [masculine] is the marked property for *nurse*.¹⁹

Proponents of markedness reversal argue that these reversals also take place according to grammatical context. This is particularly relevant to Differential Object Marking and its relation to Differential Subject Marking, for which what is marked for a subject is the polar opposite to what is marked for an object.

The assumption is that there is a prototypical subject argument and a prototypical object argument, and case-marking is used to discriminate marked examples.

So an object will be marked when it has properties that are not object-like, and a subject will be marked when it has properties that are not subject-like. When marking is conditioned by the semantics of the argument noun phrase, splits in marking will fall along the Nominal hierarchy, which is depicted in the figure below.

The Nominal Hierarchy

The **Nominal Hierarchy** is a markedness hierarchy relevant to the distribution of ergative marking (on subjects) and accusative marking (on objects) in split-ergative

19. "If we take another opposition in English - namely that exemplified by *woman~man* and *she~he* - the protests of feminists against language use are especially aimed at the 'slipperiness' of *man*, *he*, and at the swing from their use as the representative of the category (*man is a thinking animal* or *he who hesitates is lost* = zero-interpretation), to their use for the non-woman part of the species (*one man and two women came to see you* or *he's a nice person* = minus interpretation), to their use in the plus-interpretation (*everyone in New York State is entitled to an abortion if he wants it*), to unclear uses that may be interpreted in either way but for cultural reasons tend to carry the minus-interpretation (*chairman*, in the context of an academic department). Furthermore, the swing from zero-interpretation to minus interpretation may take place in the midst of one sentence: cf, *man is entitled to life, liberty, and the pursuit of happiness - and to marry the woman of his choice.*"(Waugh 1982: 305) In *The Second Sex*, Simone de Beauvoir critiques the conceptualization of gender in society in what is essentially the terminology of markedness: "The categories 'masculine' and 'feminine' appear as symmetrical in a formal way on town hall records or identification papers. The relation of the two sexes is not that of two electrical poles: the man represents both the positive and the neuter...Woman is the negative, to such a point that any determination is imputed to her as a limitation, without reciprocity...alterity is the fundamental category of human thought. No group ever defines itself as One without immediately setting up the Other opposite itself...This is the fundamental characteristic of woman: she is the Other at the heart of a whole whose two components are necessary to each other."(De Beauvoir and Parshley 1953: 5-9) De Beauvoir was influenced by the anthropologist Lévi-Strauss' application of markedness theory to human cultures. I take up this discussion of gendered language and markedness in the Discussions chapter in the section on Gradient Markedness in English Gender Marking.

or split-accusative languages (as well as inverse marking languages).²⁰ NPs that fall along the left side of the scale are considered to be more natural agents, and NPs that fall along the right side of the scale are more natural patients.

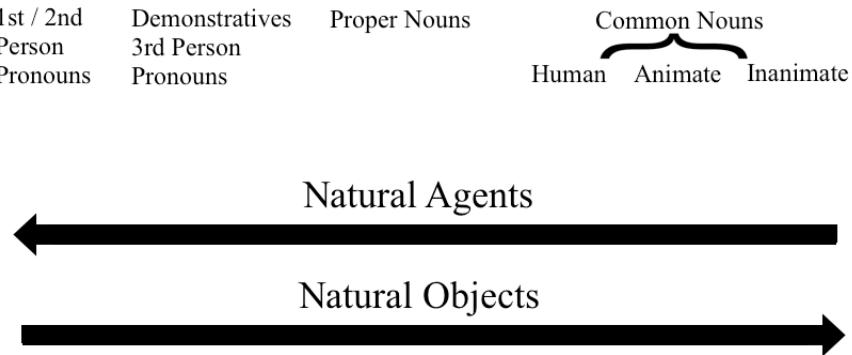


Figure 3.7: The Nominal Hierarchy as represented in Dixon (1994: 85)

The two-way scale represents a markedness reversal: first and second person pronouns are unmarked as subjects but marked as objects. Inanimate common nouns are unmarked as objects but marked as subjects. The scale is composed of multiple interacting semantic features that define a prototypical subject and object.²¹

The idea is that ergative marking will be found on marked (i.e. less natural) subjects, and accusative marking on marked objects. So if a language has an ergative split based on the semantics of the NP, ergative marking will fall along the right side of the split. Dixon takes the Pama-Nyungan language Kuku-Yalanji for an example.

20. This is also frequently referred to as the “Silverstein Hierarchy,” after (Silverstein 1976). Silverstein refers to the Agentive and Patientive Hierarchies depending on the marking, Dixon (1994) calls it the Nominal Hierarchy, and Du Bois (1987) calls it the Potentiality of Agency Scale.

21. Animacy and Definiteness are particularly relevant. Silverstein’s binary system of features included [1st person], [2nd person], [proper noun], [animate], and [plural]. Silverstein puts the second person at the top of his hierarchy. Dixon provides evidence that the first person pronouns are treated as the most natural subjects in a majority of languages, although many Algonquian languages and the Tibeto-Burman language Jyarong apparently rank second person pronouns higher (Dixon 1994: 90).

In this language the ergative case marker is found on the S_t for all NPs except local pronouns. Local pronouns, also commonly referred to as Speech Act Participants (SAP), are first and second person pronouns. The O is case-marked only when it is a first or second pronoun, and so the same split between local arguments and every other argument conditions ergative marking and accusative marking. Because first and second pronouns are very unprototypical objects, accusative marking will always fall on the left side of the split.

If there is both an accusative split and an ergative split in a language, it may be that the two case-marking systems are completely independent of each other. The marking of S_t and O are then essentially independent parameters and it is thus quite possible to have both ergative and accusative marking in the same clause. Another possibility is that case marking of one participant is affected by the semantics of the second participant. For example, the ergative may be marked only when the O outranks the S_t in “subject-ness” according to the Nominal hierarchy (as in the English sentence *The boulder struck me*). This goes back to the theory of a discriminative function of ergative case, under which it is needed to sort out unlikely subjects. So a full investigation into the conditioning factors of ergative case marking in a split system must take into account the possibility of an interaction between the semantics of the O and the S_t .

Marathi is an Indo-Aryan language that features both an ergative and accusative split. The accusative split resembles the accusative split in Hindi and Nepali: objects are obligatorily case-marked if they are animate and definite and cannot be case-marked when they are inanimate. There is an ergative split by perfective aspect:

- (55) a. *to chimṇī bagha-to*
he.NOM sparrow see-PRES.M.SG
'He sees a sparrow.'
- b. *tyā-ne chimṇī baghit-lī*
he.OBL-ERG sparrow see-PERF.M.SG

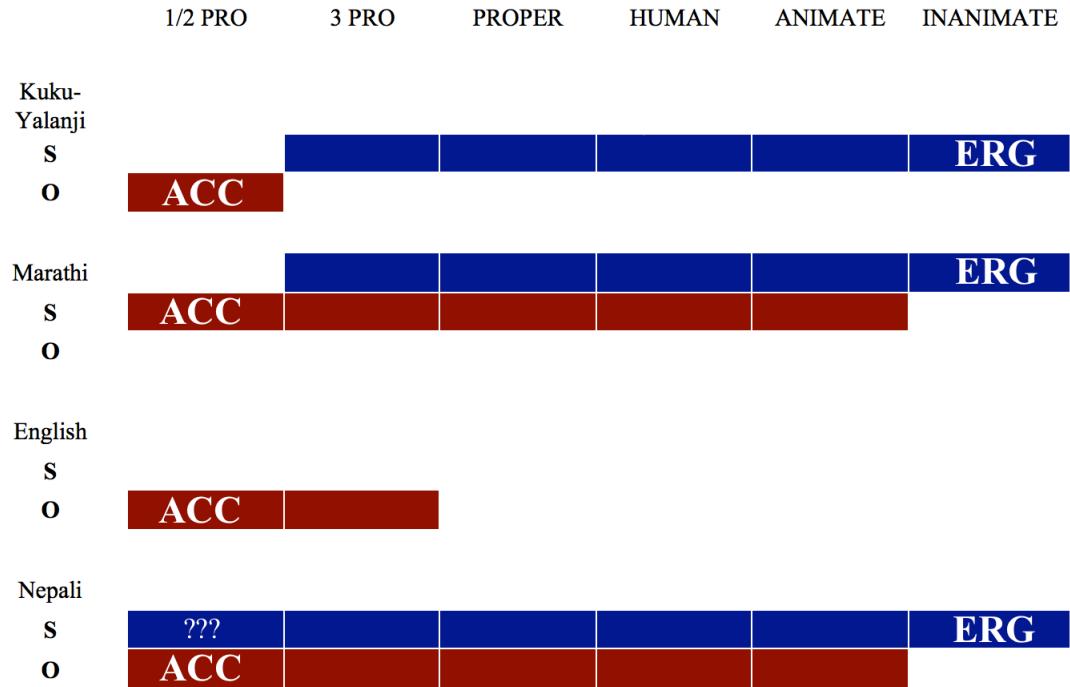


Figure 3.8: Ergative and accusative splits in four languages

'He saw a sparrow.'

There is a second ergative split between local and nonlocal pronouns.²² First and second pronouns are not overtly marked as ergative in opposition to the nominative:

- (56) a. *mī chimnī bagha-to*
 I.NOM.M sparrow see-PRES.M.SG
 'I see a sparrow.'
- b. *mī chimnī baghit-l̥*
 I.NOM.M sparrow see-PRES.M.SG
 'I saw a sparrow.'

The nominal hierarchy is also relevant to languages that do not have split-ergative alignment. Morphologically distinctive accusative/dative marking in English is lim-

22. Note that this split is only relevant to the morphological form of the pronoun. Another way of describing this is that there is a syncretism between nominative and ergative case for first and second pronouns. However, as we shall see in the next section, there is independent evidence for covert ergative case that is nonetheless unmarked on the pronoun itself.

ited to the pronominal domain: *I/me*, *we/us*, *she/her*, *he/him*, and archaically *thou/thee* and *you/ye*. Furthermore, there is an animacy distinction within the pronominal domain: the inanimate third person *it* does not have a separate accusative form. Nominative/ergative distinctions tend to collapse in the plural form of pronouns (Silverstein 1976, Dixon 1994: 92), which will be relevant to the discussion of ergative patterning in Indo-Aryan languages in the next section.²³

The Nepali alignment is superficially quite similar to that of Marathi. The accusative split is similar: although accusative marking is somewhat optional on inanimate common nouns, it is obligatory on first and second person pronouns. There is also a split conditioned by perfectivity. But ergative marking is possible for all types of NPs, and there are distinctive ergative forms for all pronouns.²⁴

On the other hand, we have seen some evidence that the Nominal Hierarchy is relevant to ergative marking in Nepali. Li (2007) argues that inanimate subjects are obligatorily marked with the ergative even in the imperfective domain. Even if this is not always the case, it appears to be a strong tendency. Furthermore, some of my consultants have expressed the intuition that the ergative is less common on first person pronouns. Perhaps the Nominal Hierarchy is not expressed in Nepali in terms of categorical splits but rather in terms of increasing likelihood of usage: this would predict that ergative-marked first and second person pronouns would be proportionally less common than their nominative counterparts, and ergative-marked inanimates would be proportionally more common than their nominative counterparts.

In our corpus analysis in the next chapter, we will need to look for answers to the following questions about Nepali ergative case marking:

23. In any case *they/them* does constitute an outright exception to these generalizations because there is a distinct accusative form even though the referent may be both plural and inanimate. This is compared to the singular inanimate pronoun *it*, which has only one form.

24. Many of them even have separate oblique forms that are the reflexes of inflection in previous stages of the language: two examples are *ma/maile* (as opposed to *ma-le*) in the first person singular, *u/us-le* (as opposed to *u-le*) for the third person singular mid-honorific.

- (1) Do we find nominative/ergative alternations on either side of the Nominal Hierarchy?
- (2) What are their relative proportions in nominative/ergative alternations and do they align with the predictions of the Nominal Hierarchy?
- (3) Is ergative marking on S_t affected by the semantics of O, either absolutely (if O is high on the scale) or relatively (if O is higher than S_t on the scale)?

A Formal Implementation of the Nominal Hierarchy

Optimality Theory (OT) is a theoretical model in which observed linguistic forms arise from the interaction of conflicting constraints. OT has been widely applied to the field of phonology since its formulation by Prince and Smolensky (1993). Aissen (2003) introduced a formal implementation of Differential Object Marking in syntactic OT which has been applied to ergative marking as well.

Markedness in Aissen's model is formulated as a competition between two conflicting constraints: iconicity (that a marked meaning should be represented by a marked form) and economy (that case-marking should be avoided if it is unnecessary). So in an accusative marker is not needed in an unmarked context (say, an object with inanimate reference). A marked context (say, a first person pronoun) is iconically represented with a marked form. The economy condition is implemented with a constraint $*\text{STRUC}_c$ which penalizes any form with an overtly marked case. The iconicity condition is implemented with a constraint ($*0_c$), which penalizes the absence of a case feature, conjoined with other constraints which represent markedness features in a scale.

The Nominal Hierarchy represents a prototype Subject and prototype Object, and the two clustered features relevant to its expression in Aissen's formulation are **Definiteness** and **Animacy**. These two features are represented by separate scales because languages rank animacy and definiteness in different ways. Under Aissen's

formulation, the prototypical object is inanimate and indefinite, and an object is most likely to be marked when it deviates from these properties.

These are the two scales:

- (1) Animacy scale: HUMAN > ANIMATE > INANIMATE
- (2) Definiteness scale: PRONOUN > PROPER NOUN > DEFINITE > INDEFINITE SPECIFIC > NON-SPECIFIC

These scales are implemented as constraint hierarchies, which are a series of constraints which are always ordered in the same way with respect to each other. The ordering of the subject constraints (SU) is precisely the reverse of the ordering of the object constraints (OBJ). These are the subject and object animacy scales:

- (1) *SU/INAN >> *SU/ANIM >> *SU/HUM
- (2) *OBJ/HUM >> *OBJ/INAN >> *OBJ/ANIM

And this is the subject definiteness scale and the object definiteness scale:

- (1) *SU/NSPEC >> *SU/SPEC >> *SU/DEF >> *SU/PN >> *SU/PRO
- (2) *OBJ/PRO >> *OBJ/PN >> *OBJ/DEF >> *OBJ/SPEC >> *OBJ/NSPEC

In the overall constraint ranking, the position of the economy condition $*\text{STRUC}_c$ determines the ergative and accusative splits. If, for example, $*\text{STRUC}_c$ is ranked between $*\text{SU}/\text{NSPEC}$ and $*\text{SU}/\text{SPEC}$, you have a system where ergative marking is obligatory on inanimate subjects and disallowed elsewhere.

Many languages with DOM use both animacy and definiteness to determine case marking. However, languages differ in the relative ranking between them. The generalization for Hindi, as we have seen, is that case marking is obligatory for NPs with human reference (including pronouns and proper nouns), disallowed on nondefinite inanimates, and optional elsewhere.

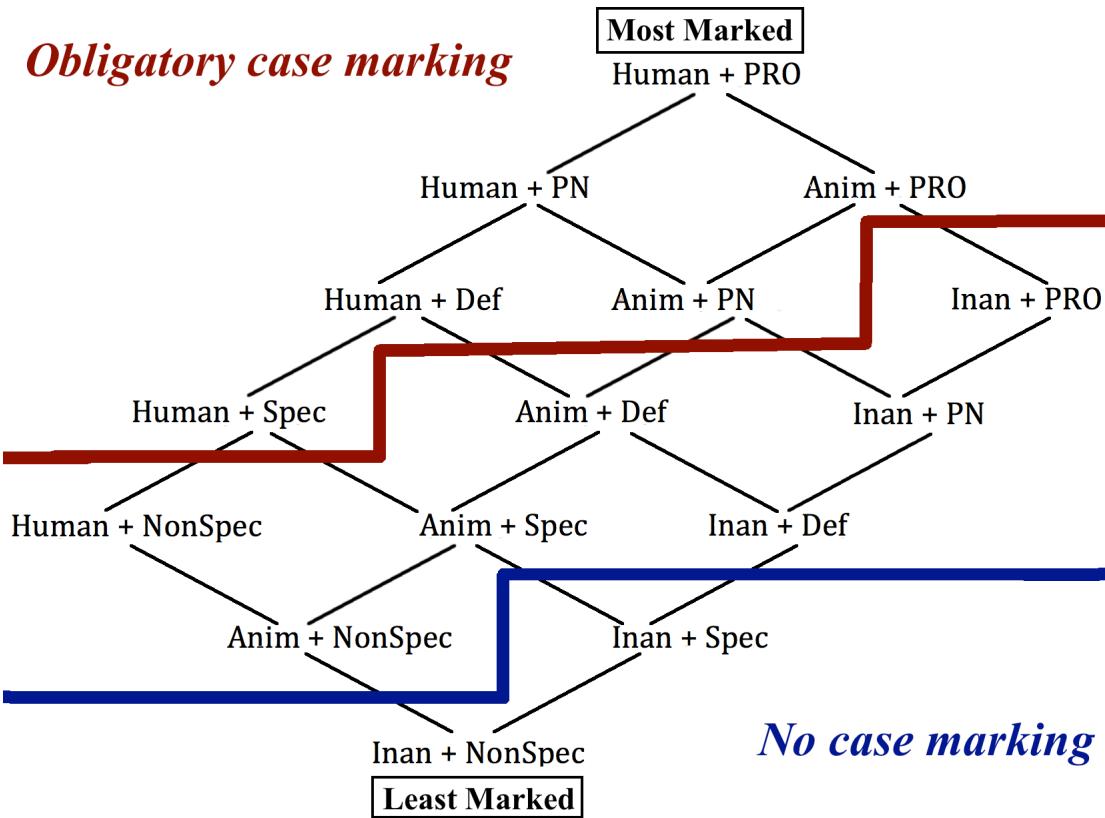


Figure 3.9: Hindi Differential Object Marking; based on (Aissen 2003: 459)

Aissen conjoins each of the constraints in the respective animacy and definiteness scales to create a partial ordering of constraints that depicts in a two dimensional space where splits can occur. Figure (3.9) depicts this space for Hindi object marking. The red and blue lines depict where in these suite of constraints the economy condition is placed to allow for regions of required marking and required non-marking. The space in between is where case marking is optional in Hindi. Aissen notes that where case marking is optional, its presence is determined by other factors like telicity or topicality (Aissen 2003: fn.24). But within the model, optionality is operationalized through optional reranking of the constraint $*\text{STRUC}_c$.

Aissen's model has a number of attractive features. It can faithfully describe the interaction of definiteness and animacy in most DOM languages. It can also

straightforwardly model language change in terms of shifting of constraint rankings over time. It makes predictions about the types of systems which could exist according to principles of markedness and the proposed markedness hierarchies. And it defines subregions of optionality between obligatory and disallowed case marking.

In principle, Aissen's implementation accounts for ergative marking in an entirely analogous way. The partial ordering depicted in Figure (3.9) is simply turned upside down such that human pronouns are the least marked subjects and inanimate nonspecific common nouns are the least marked subjects.

However, Næss (2004) notes that Differential Subject Marking (or Differential Agent Marking) is not precisely the reverse of DOM: animacy and definiteness are not the only relevant features, and they do not interact in the same way.

This is apparent from Deo and Sharma (2006), which implements Aissen's model for ergative marking in modern Indo-Aryan languages and formalizes the various ergative patterns found in these languages in terms of OT constraints. Deo and Sharma compare Hindi, Nepali, Gujarati, Marathi, Punjabi, and Bengali (Bangla), and their model shows how the varied systems in these languages could arise from the Middle Indo-Aryan pattern.

To do this, they employ a definiteness scale that ranks first and second pronouns over third person pronouns to capture the fact that languages like Marathi do not have an ergative case distinction in the first and second person. All of these languages have a perfective/imperfective split, so they stipulate a constraint against ergative case in imperfective clauses. Because languages like Gujarati have lost the distinction between ergative and nominative plural pronouns, they also include a plurality constraint. They also include constraints to distinguish overt case-marking of the ergative from inherent case. They have a separate group of constraints to model the verbal agreement system in each of these languages: agreement may be with the subject, the object, or neither (default marking for languages in which agreement is

blocked when the argument is case-marked). For Nepali, their model captures the typologically unusual fact that ergative case marking on the subject does not block subject agreement (as it does for every other Indo-Aryan language with ergative morphology). However, they do not discuss the variable ergativity in the imperfective of Nepali or the effect of animacy on ergative marking.

Thus to capture the basic facts for modern Indo-Aryan languages it is necessary to include and combine constraints that are based on multiple semantic features. Deo and Sharma (2006) is a good illustration of how the interaction of multiple features defines a complex typology of ergative marking. As noted by Du Bois (1987), Li (2007), and McGregor (2010), a full theory of ergative-marking must take into account interactions with the semantics of the verb as well as potential interactions with the semantics of the O.²⁵

In other words, the conceptualization of the subject prototype solely in terms of animacy and definiteness is inadequate. Furthermore, the subject prototype is not simply a reversal of the object prototype, and proper treatment of ergative marking must include a more multifaceted prototype like Dowty's Proto-roles or Hopper and Thompson's Transitivity, which would presumably be more difficult to adapt into an OT framework.

Markedness Extensions: Prototypes, Reversals, and Hierarchies

The traditional conceptualization of the Nominal hierarchy relies on markedness reversal. In the context of a first person pronoun, ergative case is unmarked; in the context of an inanimate common noun, ergative case is marked. The S_t prototype is animate and definite, and the O prototype is inanimate and indefinite.

25. Another issue, pointed out to me by Larry Horn, is the fact that both objects and subjects may be elided in Nepali (as well as Hindi and Marathi). So when a core argument is overt, it is already marked in opposition to a null term. Thus there could be considered a three-way opposition between elided, nominative, and accusative-marked objects.

If we take these simply as expressions of animacy and definiteness, Nepali ergativity does not entirely accord with the predicted results. On the one hand, the ergative marker is more common with inanimate subjects (following the prediction of the animacy scale) and less common with first person pronouns (following the prediction of the definiteness scale). However, where it varies with the nominative, the prediction is that the ergative form would be more found more commonly on indefinite referents. It should have low prominence. In fact, as will be discussed in the section on Discourse Prominence, the marker seems to correlate with prominence, topicality, and definite interpretations of quantifiers.

Hopper and Thompson's model makes the opposite prediction from Aissen on the prototypical object: being highly individuated, it is animate and definite. Hopper and Thompson's prototypes do not imply markedness reversal. Rather, accusative marking signals high transitivity. So both accounts predict accusative marking on pronouns, but for different reasons.²⁶

Næss (2004) challenges the assumption that the prototypical object is indefinite and inanimate. In her account, the typical direct object is highly affected by the verbal action, and thus tends to be highly individuated (that is, singular, definite, human, pronominal). She argues that marked individuals are more affected by the event, and affected arguments are more likely to be individuated, thus definite and animate. The subjects of transitive verbs are marked as controlling. Her account preserves the markedness reversal between subject and object, but accusative marking marks Affectedness and ergative marking marks Control.

Fauconnier (2011) provides a slightly different picture that is also based upon the

26. "Hopper and Thompson (1980) suggest a different interpretation of DOM, one which is also iconic (see also Magier 1987). In their account, DOM registers a high degree of clausal transitivity on one of the arguments of the clause. Various factors contribute to degree of transitivity, but one of them is individuation of the object... This account and the one proposed here (which is based conceptually on the approaches cited in the text) will lead to similar predictions about DOM... The Hopper and Thompson approach is not related to markedness reversal and therefore does not, as far as I can tell, make predictions about discriminate subject marking systems." (Aissen 2003: fn.4)

conceptual meaning of the ergative and accusative markers. She notes that cross-linguistically many languages simply do not allow inanimates in subject position, and others restrict possible agents to “independent instigators,” meaning that they are the main cause of the event. Languages with ergative splits based on animacy are rare, and ergative marking generally correlates with definiteness rather than indefiniteness. Inanimate subjects either need to be marked because they are unexpected, or avoided entirely. She takes subjects to have two properties: they are Instigators of the event and they are the ultimate Affector of the O.

In Figure (3.10), I summarize the various positions we have seen with regards to the features of a prototypical subject and a prototypical object.

	Prototypical Subject	Prototypical Object
Hopper and Thompson (1980)	volitional, potent	affected, individuated: (proper, human, animate, concrete, singular, count, referential, definite)
Dowty (1991)	volitional, sentient, affecting, moving, independent	changing, affected, stationary, dependent
Dixon (1994)	animate, definite	inanimate, indefinite
Aissen (2003)	animate, definite	inanimate, indefinite
Næss (2004)	controlling	affected
Fauconnier (2011)	affecting, instigating	affected

Figure 3.10: Summary of the Prototypical Subject and Object

There are various ideas about the meaning of the marker itself: Jakobson argues that the accusative marks **directionality of action from the verb**. Hopper and Thompson take both ergative and accusative marking to mark **high transitivity**. The Nominal hierarchy implies that ergative and accusative marking **discriminate atypical subjects and objects**.

Criticisms of Markedness

Hopper and Thompson (1980) and Dixon (1994) are both markedness-based accounts that make use of a prototype. However, in a sense they make opposite predictions about what markedness marks: for Hopper and Thompson a morphological form will mark a *more prototypical* event, while for Dixon a morphological form will mark a

less prototypical event. On the one hand, there is the meaning of the mark itself. On the other, there is the observation that marking is more common on deviations from the prototype. Næss (2004) is an analysis of markedness in DSM and DOM systems. She argues that animate objects are unmarked on the level of the object prototype (because they tend to be affected), but they are marked as members of an accusative case-marking system (for which the unmarked form in a subject-object pair is the subject).

The problem is that a theorist can pick and choose a particular element of the system to argue for or against any hypothesis. We can argue, for example, that ergativity should be more likely with inanimate subjects because the typical subject is animate and the ergative marks deviation from the norm. Or we can argue that ergativity should be more likely with animate subjects, because in an ergative case-marking system the marked member of the subject-object pair is the highly animate, controlling subject. Shannon (1986) summarizes the issue: “Given seemingly arbitrary M[arkedness]-values, M-assimilation, reversal, and complementarity, there seems little that one cannot ‘explain.’”

A second critique of markedness comes from Haspelmath (2006), who argues against the usefulness of markedness as a concept on the grounds that it has been given a number of distinct meanings in the literature (he counts twelve distinct senses). He claims that none of these senses are best understood under the cover term of markedness. In particular, he argues that semantic markedness should be understood in terms of the semantic notions of hyponymy and polysemy, along with conversational implicatures and conventionalization. This must be separated from the notion of naturalness based on textual frequency, which is a separate but important mechanism in the grammar.

Furthermore, he notes that there are varying levels of markedness pairs. To take the example of gendered pairs from above, ‘fox’/‘vixen’ and ‘prince’/‘princess’ are not

equivalent: a vixen is a type of fox, but a princess is not a type of prince. Haspelmath describes a gradient cline of oppositional pairs.

I discuss these issues in depth in chapter 6.2. To preface that discussion, I believe that the way to counter Shannon’s critique is to first consider the semantic meaning entailed by the mark itself, and then to consider what pragmatic motivations exist for using the mark. I would therefore like to discard the notion of markedness as inherently representing deviation from a prototype.

As for Haspelmath’s critique, I argue for a narrow notion of semantic markedness that correlates with the third sense he describes: formal markedness as opposition between a marked form and a zero form. It is this narrow sense of markedness which drives the opposition of two forms along a path of grammaticalization from conversational implicature to semantic entailment, as detailed by Lehmann (1989). Haspelmath’s gradient cline of oppositional pairs is in fact evidence for this particular notion of markedness, which I describe in detail in 6.2.3. It is this form of semantic markedness which drives variable ergativity in Nepali, as it does with any optional case marking system. This is intimately connected with the notion of discourse prominence, which I turn to now.

3.2.6 Discourse Prominence

It is significant that the ergative is frequently described as imparting “emphasis” where it alternates with the nominative (cf. Grierson 1904a, Clark 1963, Masica 1993). This is a strong indication that the usage of the ergative has implications for **information structure**, the way that sentences are organized to provide information in the context of a discourse (Krifka 2008, Féry and Ishihara 2016).

Clark (1963: 279) produces the following dialogue in his reader:

- (57) a. *bāhira ke-ko khalbal?*
outside what-GEN noise

‘What’s the noise about outside?’

- b. **karmi-haru-le** *chānā hāli-rahe-chan*
worker-PL-ERG roof lay-CONT-PERF2.3.PL
‘The workmen are laying the roof.’

In discussing the usage of the ergative over the nominative in this case, Clark writes that when “special emphasis is to be placed on the subject, the Instrumental [that is, Ergative] Case is used, as here. The implication is: *it is the carpenters, they are laying the roof*”(Clark 1963: 279).

Similarly, Hutt and Subedi (1999: 116-118) observe that the ergative may be used “to emphasise the subject of a transitive verb... if the sentence is a question asking who or what is the subject of a transitive verb,” or “if the sentence is a response to a question [...], or focuses in any way upon the subject of the verb.”

- (58) a. **kas-le** *tyo kurā bhan-cha* ?
who.OBL-ERG that thing say-PRES.3.SG ?
‘Who says that?’ (Hutt and Subedi 1999: 117)

- b. **āmā-le** *bhan-nu.huncha ni*
mother-ERG say-PRES.3.SG.HON PRT
‘Mother says so, you know!’ (Hutt and Subedi 1999: 117)

- (59) a. **āja kas-le** *ciyā ban-āū-cha* ?
today who.OBL-ERG tea make-CAUS-PRES.3.SG ?
‘Who will make the tea today?’ (Hutt and Subedi 1999: 117)

- b. **āja bhāi-le** *ciyā ban-āū-cha*
today younger.brother-ERG tea make-CAUS-PRES.3.SG
‘Today younger brother will make the tea.’ (Hutt and Subedi 1999: 117)

Tournadre (1991) reports a somewhat similar use of the ergative in Tibetan. However, here the ergative marking clearly indicates a contrast:

- (60) a. **kha.nub** *dgong.dag slob.khang-la su-s*
day.before.yesterday evening classroom-OBL who-ERG
bsdad-pa.red
stay-AOR.GNOM

‘In the evening of the day before yesterday, *who* stayed in the classroom?’

- b. ***nga-s ma bsdad***
I-ERG NEG stay.AOR.EGO
'I didn't stay.'
- c. ***bkra.shi gcig.po-s bsdad-pa.red***
Tashi alone-ERG stay-AOR.GNOM
'Only *Tashi* stayed.'

Tournadre (1991: 105) concludes that the ergative “has a rhetorical function whose aim is to *underline or ‘highlight’ the agent*” [emphasis in the original]. In fact, we find pragmatic effects like contrastive focus and topic correlating with the usage of the ergative in many of the Tibeto-Burman languages spoken in the Himalayas (DeLancey 2011, Zeisler 2012, Chelliah et al. 2011).

These are explanations based on the idea of **Discourse Prominence**. Clark and Tournadre are ascribing a pragmatic role to the ergative case in the interpretation of the clause. The ergative imparts some level of prominence to the marked S_t , speakers use this mechanism to highlight certain aspects of the discourse, and hearers notice this imposed prominence and interpret it.

Prominence is a common thread through functional-theoretic explanations of optional ergativity. McGregor 2009, McGregor 2010) surveys optional and differential ergative systems around the world. He argues for a semiotic [PROMINENT] feature, which is correlated with the ergative in languages with optional ergative marking, while the nominative correlates with a [BACKGROUNDED] feature (I will discuss this in more detail below).

Furthermore, Hopper and Thompson (1980)’s motivation for the concept of transitivity (which is operationalized as a proto-role) is discourse-based. Discourse is structured by two grounding elements: the Foreground consists of sequential events that provide the basic structure of a narrative. The Background consists of scene-setting statements and evaluative commentary that is extraneous to the overall structure of

the narrative. Sentences with high transitivity are in the foreground. So a marker of high transitivity like the ergative marker is a device for signaling foregrounded information.

Focus, Topic, and the Question Under Discussion

At the level of a particular utterance, different pieces of information are packaged and presented in a structured way. A prominent line of investigation into information structure follows the insights of Mathesius (1975) and the Prague Linguistic Circle, in which utterances are structured by a dichotomy between *topic* and *focus* (*theme/rheme*, *topic/comment*). The **topic** is what the utterance is about, and the **focus** is what information is being said about the topic. Another important distinction in the way that information is presented is **givenness** (Chafe 1976). The focus will generally be new information, while the topic is usually given information, although this is not always the case.

Example (57a) suggests that the function of the ergative is similar to that of an it-cleft in English. I've modified his example below to make the connection between discourse and focalization explicit. It-clefting is a syntactic device for marking focus in the English language, as is intonational stress on the noun phrase.²⁷

- (61) Who is making that noise?
- (62)
 - a. It is [the workers]_F who are making that noise.
 - b. THE WORKERS_F are making that noise.
 - c. #The workers are making THAT NOISE_F.

The question in (61) is about a particular person making a noise. The responses in (62) are answers to this question. The topic of these utterances, what they are about, is the person or persons making the sound that is known in the discourse.

27. Jackendoff (1972) refers to stress associated with focus as the A-accent.

The focus is on the new information which answers the question. (62a) focuses this response with an it-cleft, and (62b) uses a focal accent.

Note that (62c) is generally not a good response to this particular question, although it may be a good response to another question like “Which noise are the workers making?”²⁸ Thus the placement of focus presupposes the implicit question that is being answered. However, conversational participants will tend to accommodate their expectations to the utterances given. The questioner in (61) will understand from (62c) that the answerer has responded to a different question than the one which was asked, and will infer that this new question is relevant to the discourse. Thus speakers use focus pragmatically to guide the discourse by proposing questions and asserting answers to them.

This analysis follows from a question-based model of discourse structure (Roberts 1996, Büring 2003, Velleman and Beaver 2016), under which discourse is guided and shaped by the **Question Under Discussion** (QUD). Every assertion is conceptualized as the answer to an explicit or implicit question given in the discourse. Assertions add new information into the **Common Ground**, the background information that is mutually known to be shared between the participants (Stalnaker 1974). Discourse is guided by hierarchically-embedded questions which represent topics of discourse.

Another line of formalization comes from the theory of alternative semantics of focus from Rooth (1985), under which focus indicates that there are alternatives relevant in the current discourse. The question-based model, in which focus is a pragmatic phenomenon rather than a semantic one, is more relevant to our purposes because it makes explicit the intuition that the utterance form with the ergative is the response to a particular type of question. It also brings together two separate lines of thinking about the notion of topic, in which topic is conceived of as either the

28. It is only a good response with a contrastive topic, as in “The workers are making THAT NOISE, and the boss is making THAT OTHER NOISE.”

entity that the statement is about or the *question* that the statement is answering (McNally 1998). The topic of any sentence which is about a particular referent X can be thought of either as the referent itself or as the answer to a question “What about X ?”. Thus we might expect ergative marking to correlate with either topic or focus.

Abadie (1974) and Bickel (2011) take Example (57a) from Clark to be evidence that the ergative marks focus. However, Abadie (1974) notes that her informants disagree that the element marked by *-le* must be focused or even particularly emphatic. Verbeke (2011) observes that the ergative is often used in “out of the blue” contexts, when there is not clearly an implicit question that involves the subject. Verbeke concludes that “emphasis/focus may be the motivation of the occurrence of *le* when a new [subject] that is different from the previous subject is introduced; in other instances, emphasis/focus is clearly not the decisive factor”(Verbeke 2011: 161).

Furthermore Tournadre’s Tibetan example (60 above) may be a **contrastive topic**. Topics are typically not prominent in an utterance because they depict given information, that is, information which is part of the common ground. In many languages including Nepali, topics may be omitted (in either subject or object position). However, contrastive topics emphasize a contrast:

- (63) Who stayed in the classroom last night?

- (64) a. I_{CT} stayed in town last night. ($TASHI_F$ stayed in the classroom.)
b. [AS FOR ME]_{CT}, I stayed in town last night. ($TASHI_F$ stayed in the classroom.)
c. # I_F stayed in town last night. ($TASHI_F$ stayed in the classroom.)

Contrastive topic may be marked in English by what Jackendoff (1972) calls the English B-contour. Note that the rise-fall pattern of the contrastive contour in (64a) is different from the pitch accent designated by F. The first sentence in (64c) with an A-accent is not a felicitous answer to the question because it addresses a different

question (“Who stayed in town last night?”). The B-contour imparts prominence to a topic, and in doing so it presupposes the answer to a question that is different from the one that was asked. Under the formulation of Büring (2003), this question would be “*Who* stayed *where* last night?” The usage of the contrastive contour presupposes a question with two wh-words rather than one, and it is a strategy for changing the topic under discussion (the QUD).

While it has a slightly different usage, the “As for” construction also has the effect of imparting prominence to a topic in a way that emphasizes contrast. Many languages have markers that associate with topicalized information of some sort, but it is difficult to categorize a unified topic marker category in the world’s languages (Büring 2003). Büring notes that topic markers differ in terms of whether they are obligatory or optional, whether they can serve other functions, and what sorts of referents can be marked. In many languages (as with English “as for”), they must introduce an existing discourse referent as the new topic. In other words, the topicalized entity has to be definite, or at least highly salient in the discourse.²⁹ However, this is not true for topic markers in Tzozil, which can introduce a new discourse referent as the topic.

Another point of variation is whether a topic marker is available on questions. Note the examples in (58) and (59). The ergative is found on both the question and the response. In my elicitations, multiple consultants remarked that if the question has an ergative marker, then the answer has to have an ergative marker. We also find this with Tournadre’s Tibetan example. This is an odd behavior if we think of the ergative as having a contrastive effect similar to the “as for” construction, because once a new topic is introduced there’s no reason to immediately reintroduce it.

In sum, we would like to know whether the usage of the ergative correlates with

29. It would sound odd to begin a conversation with the statement “As for dolphins, I hate them.” The referents of local pronouns (*I* and *you*), on the other hand, are always salient discourse referents.

highly focalized elements or contrastive topics. If it is associated with focus, we expect that the referent of the ergative-marked element

- (1) will usually (but not always) be new to the discourse;
- (2) may be definite or indefinite;
- (3) provides new information about the topic of discussion;

Whereas if it is associated with contrastive topic we might expect that the referent of the ergative-marked element

- (1) is given in the discourse;
- (2) is not the current topic;
- (3) is definite;
- (4) is the topic of discussion, i.e., is a referent X such that the sentence can be considered a relevant partial response to the over-arching question “What about X ? ”

Intonation, Word Order, and Discourse Particles

In English, intonation and pitch convey information about focus and topic. In Nepali, focused elements are not straightforwardly associated with distinctive prosodic contours or accents. In Ladd (1996)’s typology of intonation patterns, Hindi and Bangla are described as lacking lexical stress and lexical accent. Generalizations about these languages are applicable to Nepali as well.³⁰

By contrast, word order in Nepali is highly correlated with information structure. While the default word order for Nepali clauses is SOV, all configurations are possible

30. On the other hand, there is some recent experimental evidence that focus does in fact have an effect on prosodic realization in Hindi (Patil et al. 2008, Féry et al. 2016). But outside of a laboratory environment it is not practical to investigate these subtle effects as they relate to information structure in Nepali.

and have effects on how information is presented. Butt and King (1996) analyze these configurations as the syntactic encoding of discourse functions. Their languages of study are Urdu and Turkish, but their generalizations are applicable to the Nepali language, which behaves quite similarly to Urdu in this respect. Butt and King distinguish new and given information, but they also make a separate distinction based on prominence:

- (1) Topic: [-NEW][+PROMINENT]
- (2) Focus: [+NEW][+PROMINENT]
- (3) Background: [-NEW][-PROMINENT]
- (4) Completive: [+NEW][-PROMINENT]

The first element of a clause is the topic. The subject is clause-initial by default, but when another element is preposed in the utterance this element is topicalized. Here examples from the NNSP sample (discussed in the Methodologies chapter in the Corpus section):

- (65) *tapā̄i-lāi yo kalar dherai māc gar-cha*
 you.HON-ACC this color very match do-PRES.3.SG
 ‘YOU_{TOP} this color matches quite well.’ [V001001004; M9]

The placement for focus is immediately before the verb (the default position of O). If there is more than one focused element, the other focalized elements may remain in-situ, but there is a contrastive interpretation.

- (66) *khānā.sānā sabai bebasthā hāmi gar-ch au*
 food.RED everything manage we do-PRES.3.SG
 ‘Food and everything WE_F will take care of.’ [V001001004; M7]

Backgrounded information, like topic, contains new information. But “while topics are the pointer to the relevant information... to be accessed by the hearer, back-

grounded material only provides more detailed information as to how the new information fits in with the already known information” (Butt and King 1996: 4). So it is new information that is [-PROMINENT], and its placement is after the verb.

- (67) *kurā gar-chu ma*
 conversation do-PRES.1.SG I
 ‘Conversing, [I (am)]_{BG}.’ [V001001001; M3]

Finally, completive information is also non-prominent background information, except that it tends to be referential, that is, discourse old. Completive information may also come before the verb, but the material immediately before the verb is more prominent and focalized. I was unable to find a good example of this type in the corpus. Thus the general structure of Nepali clauses is this:

- (68) [Topic] [Completive] [Focus] V [Backgrounded]

In a default SOV configuration the S is interpreted as a topic, and the O as focus. Another way that Nepali organizes the information structure of elements in a clause is with discourse particles (what Acharya (1991) refers to as “phrasal nuance particles” (Acharya 1991: 143)). In particular, the *cāhī* particle seems to be a marker of contrastive topic. It typically delineates an NP or phrase in the utterance which is preposed, and it clearly correlates with an introduced topic. I have been unable to find detailed descriptions of the particle in the literature, but these examples from the Nepali National Spoken Corpus are illustrative:

- (69) *kalar cāhī yastai thik cha*
 color TOP like.this.EMPH fine COP.3.SG
 ‘As to the color, it is fine like this.’
- (70) *hātti cāhī uhā-le “path-ā-idin-chu”*
 elephant TOP PRO.3.SG.HON-ERG “send-CAUS-BEN-PRES.1.SG”

bhan-nu.bha-eko.cha
say-PRES.PERF1.3.SG.HON
'As to the elephants, he has told us, "I will send (them)."

- (71) *ma cāhī ukālo cāhī hin-na sak-chu orālo*
I TOP uphill TOP walk-NONFIN able-PRES.1.SG downhill
sak-dina
able-PRES.1.SG.NEG

'As for me, as for (hiking) uphill, I can do it. Downhill I can't.'

If correlated with focus, we should expect the referent of a *le*-marked element:

- (1) to be more likely to appear preverbally but not in clause initial position;
- (2) to appear concurrently with or in the same position as *ta* or *po*.

Whereas if correlated with contrastive topic, we expect the referent of a *le*-marked element:

- (1) to be more likely to appear in clause initial position;
- (2) to appear concurrently with or in the same position as *cāhī*.

Categorical Propositions

There is another line of thinking about information structure that is relevant to theories of discourse prominence. This is the theory of Thetic and Categorical Propositions which was originally developed by Marty (1918) as an expansion of Brentano (1874). Kuroda (1972, 1990) apply this theory to linguistic expressions in an analysis of Japanese case markers.

Marty's theory is that there are two general types of propositions, or judgments: the **categorical proposition** and the **thetic proposition**. Only categorical judgments have a logical subject-predicate structure:

[T]he categorical judgment is assumed to consist of two separate acts, one, the act of recognition of that which is to be made the subject, and the

other, the act of affirming or denying what is expressed by the predicate about the subject. With this analysis in mind, the thetic and the categorical judgments are also called the simple and the double judgments.
(Kuroda 1972: 154)

Thetic judgments, on the other hand, impart no particular prominence to one element over others. The subject-predicate grammatical structure does not necessarily correspond to a subject-predicate logical judgment. In an impersonal thetic constructions of English like “It is raining,” the syntactic subject position is taken by a dummy pronoun. The sentence as a whole describes a state of affairs, a judgment about the way that the world is (as opposed to a judgment about the way a particular entity is). It is, in the terms of Ladusaw (2000), an existential assertion about a description.

So a thetic proposition may have a fully realized subject form, but this form is not prominent in the discourse. This means that sentences of English are often ambiguous between thetic and categorical readings when the subject takes and indefinite article.

- (72) *inu ga hasitte iru*
dog SUBJ run PROG
'A DOG is running.' / 'There is a dog running.' (Kuroda 1972: 160)

The Japanese subject marker that follows *inu* “dog” is the subject marker *ga*. Kuroda argues that *ga* marks a thetic proposition. It is a “a direct response to the perceptual intake of an actual situation” (Kuroda 1990: 80). Crucially, it is not a statement about a particular entity. It is a statement about the existence of an ongoing event. The description of this event of running requires an actor, and this actor is designated by “a dog.” The inclusion of a subject is grammatically necessary in English, but it is not the topic, and its reference does not have a strong persistence outside of that particular utterance.

Kuroda contrasts this with a categorical judgment, which is a statement about a particular dog:

- (73) *inu wa hasitte iru*
 dog CAT run PROG
 ‘The dog is running.’ (Kuroda 1972: 160)

Kuroda writes that the speaker’s “interest is primarily directed towards this entity, and the happening of the event referred to...is precisely that he wants to relate the occurrence of the event to this entity”(Kuroda 1972:164). In a thetic proposition, the subject is simply a constituent of an event. But here the subject plays a more prominent role. “It apprehends the dog as a particular entity in the perceived situation” and “it involves the cognitive act of apprehending something as a substance and attributing to it a certain property perceived in a particular situation” (Kuroda 1990: 80).

The Japanese *wa* marker may be found on elements of the sentence as well. Kuroda notes that the logical subject of the utterance (which may differ from its grammatical subject) is the element marked by *wa*:

- (74) *neko wa inu ga oikakete iru*
 cat CAT dog SUBJ chase PROG
 ‘A dog is chasing the cat.’ (Kuroda 1972: 167)
- (75) *niwa de wa inu ga neko o oikakete iru*
 garden LOC CAT dog SUBJ cat ACC chase PROG
 ‘In the garden, the dog is chasing the cat.’ (Kuroda 1972: 168)

This *wa* marker is generally considered to be a prototypical topic marker, and the theory of thetic and categorical propositions may be considered a theory of topicality. Portner and Yabushita (1998) discuss Japanese as it relates to a topics-as-entities formalization of topicality in which topics denote entities that the sentence is “about.” The topic information is part of the common ground (and typically, but not always, definite), and the new information may be understood as the focus. McNally (1998) and Portner and Yabushita (1998) both contrast a topics-as-entities formalization of topic with a question-under-discussion formalization of topic, in which sentence topics

are modeled as questions, in that the marked element fits into the possible question under discussion. The sentence element identified as a topic must be definite and its existence must be presupposed in the mind of its speakers. In comparing these notions of topicality with Kuroda's terms, the notion of a categorical subject is associated with topic.

In some ways Japanese *wa* is more clearly analogous to the Nepali discourse particle *cāhī* because it can attach to different kinds of phrases and is clearly associated with prominence. However, there are some aspects of Kuroda's theory of categorical judgment that make it attractive to the study of the Nepali ergative. It provides an explanation for the habitual and/or generic readings of simple present tense verbs in Nepali (as discussed in the section on stage-level and individual-level predication):

- (76) *kukhurā-le phul pār-cha*
chicken-ERG egg lay-PRES.3.SG

‘Chickens lays eggs.’ (Hutt and Subedi 1999: 116)

- (77) *cālak-le gādi calāu-cha*
driver-ERG car drive-PRES.3.SG

‘The driver drives the vehicles.’ (Butt and Poudel 2007: 7)

Kuroda notes that generic sentences which define an inherent property, as in (76), must be categorical propositions. The usage of the ergative often corresponds with a distinction between a stage-level interpretation (“The driver is driving the vehicle”) and an individual-level interpretation, but this is not always the case. However, my consultants commonly expressed an intuition that the form with the ergative “picks out” the referent from other options. One of the elicitation consultants put it this way:

“It seems that when it is a general statement you can leave out the *-le* and it still makes sense. But having the *-le* just makes it clear as to who is doing the action. You are making the extra statement that it is [the subject] doing the action.” [TD]

Ladusaw (2000), expanding upon Milsark (1974), notes the interaction between proposition type (thetic, categorical), predication type (stage-level, individual-level), and subject type (weak, strong). Strongly construed subjects are highly definite, and weakly construed subjects are indefinite. Milsark's generalization is that weakly construed subjects are not found with individual-level predication.

	Thetic	Categorical
Individual-Level Predicate	No	Yes
Stage-Level Predicate	Yes	Yes
Strong Subject	Yes	Yes
Weak Subject	Yes	No

The upshot is that a categorical proposition may be based upon either an individual-level predicate or a stage-level predicate but must have a strongly construed subject. However, a thetic proposition can only be based upon a stage-level predicate with either a weakly or strongly construed subject. If we hypothesize that the ergative correlates with categorical propositions (that it marks the logical subject of a categorical proposition in the transitive imperfective), then we hypothesize:

- (1) The ergative should be possible with both stage-level and individual-level interpretations, but to correlate with an interpretation in which the subject has a deep connection to the predicate or the predicate defines the subject;
- (2) The ergative should not be possible with weakly construed quantifiers or with indefinite referents;
- (3) The nominative should not be possible on categorical propositions. If we define the subject of a categorical proposition as an aboutness topic, it should not be the possible in the answer to a “What about *X*?” question.

Prominence and Markedness

McGregor (2009, 2010) surveys optional and differential ergative case marking systems and provides an account of ergative marking from a usage-semiotic perspective. McGregor's two features are [**prominent**] and [**backgrounded**]. There is a cluster of features associated with prominence and a cluster of features associated with backgrounding, and in different languages these grammaticalize in different ways such that different features are primarily associated with usage. Figure (3.11) summarizes these features. The features associated with prominence are similar to those expected in a markedness-based account in which prototypical subjects are high in agency.

[Prominent]	[Backgrounded]
Contrastive Focus	Topic
Unexpected	Predictable
High in Agency/Potency	Low in Agency/Potency

Figure 3.11: Summary of Features from McGregor (2010)

Furthermore, McGregor emphasizes that the use or non-use of a marker can itself be a grammatical sign, separate from the marker itself. This observation comes from languages like Gooniyandi, for which McGregor argues that the usage of the ergative marking does not signal anything, but non-use signals low agentivity. While ergative marking is always associated with prominence rather than backgrounding, there are four different ways that a language may operationalize this distinction:

- (1) Usage of the Ergative signifies nothing: [-**prominent**]

Non-Usage of the Ergative signifies nothing: [-**backgrounded**]

- (2) Usage of the Ergative signifies prominence: [+**prominent**]

Non-Usage of the Ergative signifies nothing: [-**backgrounded**]

(3) Usage of the Ergative signifies nothing: [-prominent]

Non-Usage of the Ergative signifies backgrounding: [+backgrounded]

(4) Usage of the Ergative signifies prominence: [+prominent]

Non-Usage of the Ergative signifies backgrounding: [+backgrounded]

In the first case there is a split that is determined entirely by semantic or syntactic factors (there is no optional ergative marking). In the second, the ergative signifies prominence but its non-usage signifies nothing. An example of this would be the Hindi intransitives in which the usage of the ergative signifies that the subject underwent the action purposely (thus, is high in agency). The third example is the Gooniyandi case, in which the ergative signifies nothing but the non-usage of the ergative signifies backgrounding. And McGregor predicts a fourth case in which both usage and non-usage are informative.

This is compatible with Hopper and Thompson (1980), in which the ergative form may signal high transitivity while the nominative form may signal low transitivity. It is less clearly compatible with Aissen's formulation, in which the absence of a marker does not have meaning in itself. An account based on markedness asymmetry could be generally translated to this theory. It would require that the zero-form be considered the marked form in some languages for which the ergative signifies nothing and its non-usage signifies low-agentivity. Additionally, a markedness account would predict that it is not possible for there to be a language in which both usage and non-usage of the ergative are informative.

I believe that discourse prominence is an inherent feature of any optional case marking system, not just OEM systems. For example, Aissen (2003) notes that in DOM systems where marking is optional, the usage of the accusative marker will correlate with topicality. Whenever there is an opposition between a form and its

absence, the usage of the form will correlate with the marked entity being given a greater prominence in the discourse. The marked element will be more likely to be apprehended as the logical subject of a categorical proposition, and it will tend to correlate with definiteness and topicality.

Chapter 4

Observations

इन्द्रका अगाडी स्वर्गको बयान ।

नेपाली उखान-टुक्का

Describing heaven to Indra, Ruler
of the Sky

Nepali proverb

In the last chapter I discussed the various explanations of the Nepali pattern of ergative marking where its distribution alternates with the nominative. I also discussed the broader literature on optional ergativity and how it touches on the theoretical issues of markedness, transitivity, and discourse prominence. There is substantial overlap in these concepts, and taken together they make a number of predictions about the distribution of the ergative. These predictions can be conceptualized as a set of potential associations between ergative marking and various morphosyntactic, semantic, and pragmatic features.

The goal of this chapter is to thoroughly investigate these feature associations using the tools described in the Methodology section: directed elicitations, grammaticality judgment survey results, and the analysis NNSP sample. For the conclusions

that have already been drawn about Nepali in the literature, this investigation will provide further evidence for these associations. From the broader literature on optional ergativity and transitivity there are a number of additional potential associations that are relevant to Nepali, and I will examine each of these in turn.

The feature associations broadly fall into four categories: those that involve the interpretation of the **Event**, those that involve the interpretation of the **Subject** (both S_t and S_i), those that involve the interpretation of the **Object** (O), and those associations that correlate with features of the **Discourse**.

I find multiple associations between ergative marking and semantic/pragmatic features, almost all of which follow the predictions of Hopper and Thompson (1980)'s Transitivity Hypothesis. However, with few exceptions these associations are not categorical but rather represent general tendencies. They are features of pragmatically-conditioned case marking and all of them ultimately derive from two sources:

- (1) The meaning of *-le* as the marker of an effector role
- (2) The prominence accorded to a case-marked subject in an optional case marking system

I find the following generalizations to be categorical:

- (1) Ergative marking is required in perfective transitive main clauses and varies with the nominative in all other transitive clauses.
- (2) Ergative marking is excluded from all copular clauses and intransitive clauses (although some intransitive clauses can exhibit ergativity if they are specifically construed as transitive, and some transitive clauses may lack it if they are construed as intransitive).

These are not pragmatically conditioned associations, but rather represent a grammaticalized correlation between ergative marking and specific properties of the clause.

4.1 Ergative marking and the interpretation of the Event

Ergative marking is required in the domain of perfective transitive main clauses, varies with the nominative in other transitive clauses, and is disallowed elsewhere. It correlates somewhat with individual-level predication, but more closely with the notion of characterizing predicates. Differing judgments on whether the ergative is associated with habitual readings are due to two conflicting interpretations of the ergative: as an effector of an event with an individuated object, or as a marker of a categorical proposition with a characterizing interpretation of the event.

4.1.1 Ergativity and the Nepali Verb Form

Before investigating whether there is an association between ergative case-marking and the interpretation of event structure, it is important to precisely delineate where ergativity is obligatory, where it is variable, and where it is disallowed. Here I present the results of elicitation, survey, and corpus data as it relates to the presence or absence of *-le* in each of the various verb forms of the Nepali language. I restrict my investigation to transitive clauses because the distribution of *-le* in intransitive clauses is a separate issue which I will discuss in the following subsection.

The overall results of the Kathmandu survey and NNSP analysis are presented in Figure (4.1). In the survey results, the average judgment for sentences in these verb forms (rated on a scale from 1-5) is presented for nominative and ergative forms of the same sentences. In the survey results I present the overall percentage of ergative (vs. nominative) marking on the S_t . In the following subsections I will discuss each of these results individually. The main takeaway from these figures is that ergative marking is heavily dispreferred/disallowed in the perfective verb forms and variable elsewhere. In some cases (such as with the hypothetical future and definite future verb forms)

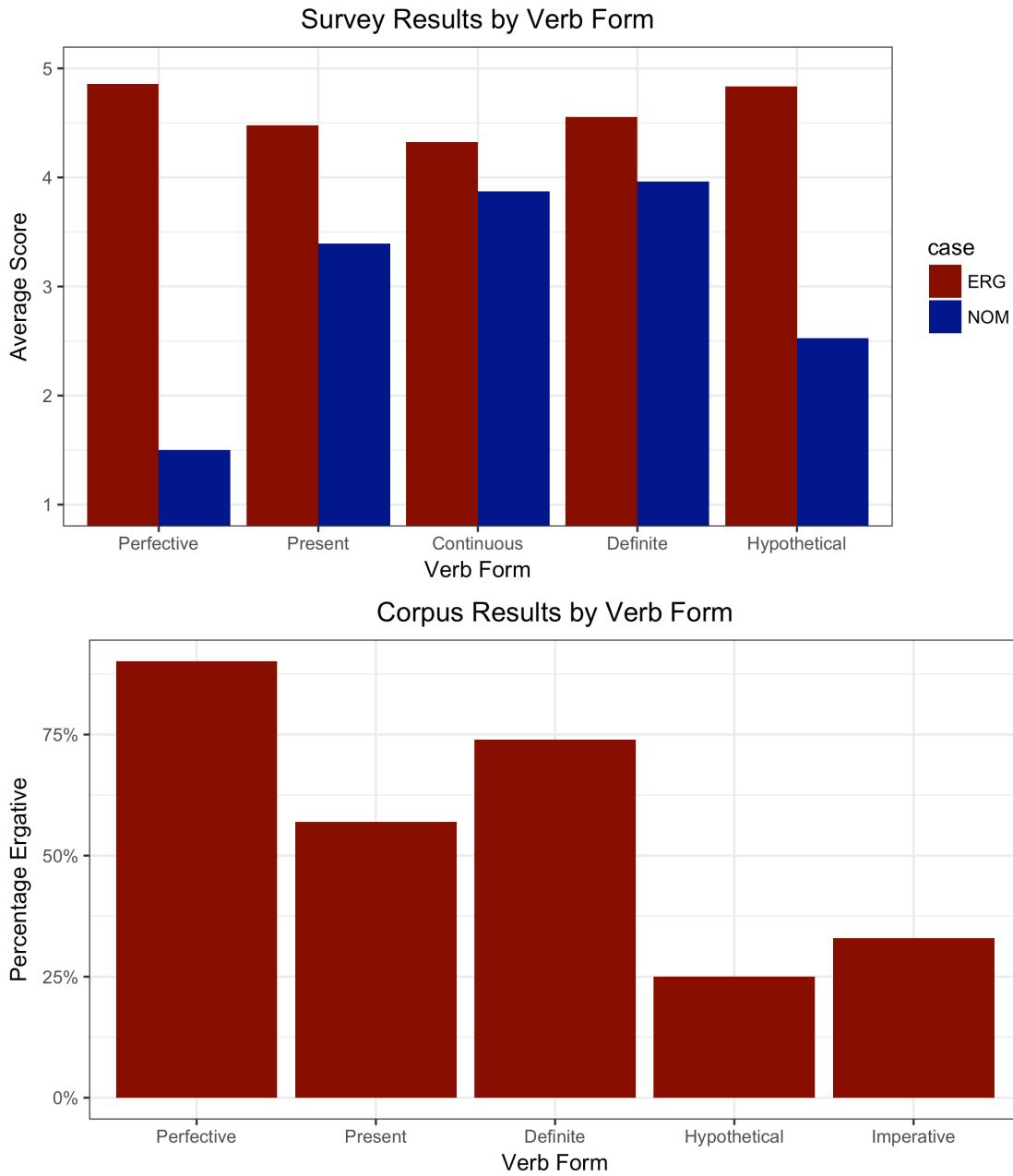


Figure 4.1: Ergative Marking Results by Verb Form

there is a marked preference for the ergative over the nominative, but in all of these imperfective verb forms there is variation between ergative and nominative-marked transitive subjects.

The Perfective Domain

We expect ergative case marking to be obligatory for transitive clauses with perfective verbal morphology. The perfective domain consists of verb forms which contain *-e-* or *-yo*: PERF, PRES.PERF, PST.PERF, PRES.MIR:¹

- (78) a. *rām-le āap khaa-yo*
ram-ERG mango eat-PERF.3.SG
'Ram ate the mango.' [PK]
- b. *maile doctor-lāi bhet-eko chu*
I.ERG doctor-ACC meet-PRES.PERF.1.SG
'I have met the doctor.' [PK]
- c. *tei ta maile khoj-eko thiē*
there PRT I.ERG search-PST.PERF.1.SG
'I had looked for you there.' [V001002005; M24]
- d. *us-le kām gar-echa*
PRO.OBL-ERG work do-PRES.MIR.3.SG
'Apparently, he has done the work.' [BB]

Every one of the elicitation consultants was unwavering in the judgment that the nominative form is ungrammatical in sentences with these verb forms. Considering that the expression of ergativity is sensitive to discourse context in many other grammatical domains, we might expect a nominative subject to be construable in the right situation, but this does not appear to be the case. When respondents are asked about whether they can construct a context in which they or someone else might be able to use sentences like the above with a bare subject, the only possibility that comes to mind is that the sentence comes from a nonnative speaker of Nepali.

The Kathmandu Survey did not have a section dedicated to testing perfective verb forms because the focus of the survey was variability in the imperfective domain.

1. As to the past mirative *-e-thyo*, there were no examples in the corpus and multiple respondents expressed the belief that it was simply a shortened form of the past perfect. While it may be a separate case for some speakers (cf. Michailovsky 1996, Peterson 2000), I'm not including it here due to lack of evidence.

However, there was a single question (B4) for which the response clauses were in a perfective tense. From the results in Figure (4.2) it is clear that the ergative form is heavily preferred and the nominative form is heavily dispreferred, although the percentage of favorable responses (meaning, 4 or 5) for the nominative form was somewhat high (7.1%) considering the categorical judgments of all nine of my elicitation consultants.

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
ERG (n=28)	4.86	96.4%	0.0%
NOM (n=28)	1.50	7.1%	85.7%

Figure 4.2: Survey Results: ERG/NOM with Perfective Transitive Clauses

In the NNSP sample analysis I coded 558 clauses as perfective, of which 195 were non-modal transitive main clauses. The S_t argument was elided for 53% of these clauses, and of the remaining 91 clauses, **the S_t was in the ergative case 90.1% of the time.** While this result also follows the expected pattern in that the ergative is heavily preferred, the nominative bare form was found nearly 10% of the time, which is a relatively high percentage given the categorical judgments of every elicitation consultant.

One explanation for this unusual percentage is dialectal variation in the language of the survey respondents and corpus speakers. SB in particular gave his impression that this usage of the nominative in transitive perfectives is a feature of Nepali spoken in the Terai (southern plains) south of Kathmandu, particularly by native speakers of the Tharu language.

Among the survey respondents, the only Tharu speaker was also the only respondent to rate the nominative form of the transitive perfective sentence a ‘5,’ which is suggestive but anecdotal. In the NNSP, individual speakers ranged from 0-33% usage of the nominative form in the perfective. The speakers who were recorded in the Terai (interview V001002005; speakers M00024, M00025, M00026, and M00027) used the

nominative similarly if not somewhat less frequently (8-13%) than those interviews with speakers from elsewhere (0-33%). With the corpus, it may be the case that the ergative form is more likely to be omitted in rapid speech, or reduced to the level of not being audible to the transcribers.

Exceptions aside, we should consider these perfective forms to represent a fairly categorical split in ergative marking. The nominative form is dispreferred for each of these verb forms in the transitive perfective domain.

The Present Imperfective Domain

In contrast with the perfective domain, there is a very clear alternation between nominative and ergative case on the S_t for those imperfective tenses which are built upon the simple present. We find both nominative and ergative forms in the corpus:

- (79) a. *kasari - ma udāharan̄ din-chu*
how - I.NOM example give-PRES.1.SG
'How - I'll give you an example ...' [V001002003; M13]
- b. *ani pheri 'lagāu' bhan-chau timi-le*
and again 'wear.IMP' say-PRES.2.SG you-ERG
'And then once again 'wear it!' you'll say.' [V001001004; M7]

And elicitation respondents generally find both forms to be possible:

- (80) a. *sabai birālo-haru/birālo-haru-le māsu khān-chan*
all cat-PL-ERG/cat-PL-ERG meat eat-PRES.3.PL
'All cats eat meat.' [ST]
- b. *prakāsh/prakāsh-le cigarette piun-cha (tyasari*
prakash.NOM/prakash-ERG cigarette drink-PRES.3.SG (therefore
khok-irah-eko cha)
cough-PROG.PRES.3.SG)
'Prakash smokes cigarettes (and that's why he has been coughing).' [PK]²

2. The verbs typically used for the inhalation of tobacco are *khānu* 'to eat/consume' and *piunu* 'to drink.' Schmidt (1993) notes in the dictionary entries for these terms that *khānu* is the general term for eating, smoking and drinking; *piunu* is more commonly associated with the consumption of liquor.

A few of these examples come from elicitations in which I was specifically interested in whether an alternation was possible, and for these I have written the two possibilities. For the moment I will leave aside the question of whether the ergative correlates with individual-level predication or a habitual or generic reading (this will be discussed later in this section) but just point out that both the nominative and the ergative may be found in clauses with the PRES form.

This alternation exists with the present continuous (CONT) and archaic present (ARCH.PRES) forms as well. These forms are the same except that there is an aspectual morpheme which attaches to the stem (*-dai* and *dā* respectively):³

- (81) a. *john/john-le kitāb-ko antim pāna lekh-dai-cha*
john.NOM/john-ERG book-GEN last page write-CONT-PRES.3.SG
'John is writing the last page of (his) book.' [PK]
- b. *prāya.jasto bacā-haru bakunda khel-dai-chan*
most child-PL football play-CONT-PRES.3.PL
'Most of the children are playing football.' [RM]
- c. *u/usle kām gar-dā-cha*
PRO.NOM/PRO.ERG work do-ARCH.PRES.3.SG
'(S)he does work.' [SB]
- d. *pāle-le dhyān di-era her-dā-cha*
guard-ERG attention give-CONJ watch-ARCH.PRES.3.SG
'The guard watches (the bank) attentively.' [BA]

All elicitation consultants found both the ergative and the nominative to be grammatical in transitive clauses with the present imperfective verb form. This is evident from the Kathmandu survey, in which the majority of the questions were in the present or present continuous verb forms. These results are summarized in Figure (4.3) and Figure (4.4).

3. The archaic present form is not found in the corpus and elicitation consultants consider it to be a mark of written language or formal speech.

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
ERG (n=662)	4.48	87.6%	8.0%
NOM (n=660)	3.39	55.5%	28.5%

Figure 4.3: Survey Results: ERG/NOM with Simple Present Transitive Clauses

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
ERG (n=413)	4.30	82.8%	9.9%
NOM (n=413)	3.92	67.5%	15.5%

Figure 4.4: Survey Results: ERG/NOM with Present Continuous Transitive Clauses

Note that the average score is high and a majority of the respondents like both the ergative and the nominative form, which suggests that both are acceptable responses in many cases. In the simple present there is a notable preference for the ergative form, although the majority of speakers find the nominative form to be acceptable as well.⁴

In the NNSP sample I coded 691 verbs in the present imperfective tenses, of which 116 were non-modal transitive main clause verbs. The S_t was elided for 53% of these verbs. Of the remaining 55 clauses, all of them were in the simple present tense, and **the ergative was found on the S_t 58% of the time.** Both the ergative and nominative are frequently found in these verb forms built upon the simple present.

The Past Habitual Imperfective

For the past imperfective (habitual) verb form, the extent to which the nominative varies with the ergative is unclear. I asked PK, TD, BB, and SB about the usage of this form, and while PK absolutely disliked the ergative form in the first person, they all accepted both nominative and ergative forms depending on the context.

- (82) a. *ma cricket khel-the*
 I cricket play-PST.HAB.1.SG

4. The main purpose of the survey was to test the usage of the ergative in certain contexts, such as with generic or habitual readings, which may have contributed to the preference for the ergative here.

‘I used to play cricket.’ [PK]

- b. *cor/cor-le* *scarf cor-āu-thyo*
thief.NOM/thief-ERG scarf steal-CAUSE-PST.HAB.3.SG
‘Thieves would steal the scarf.’ [TD]
- c. *u/usle* *ruk̥h rop-ne* *kām gar-thyo*
PRO.NOM/PRO.ERG tree grow-NONFIN work do-PST.HAB.3.SG
‘S(he) used to do work planting trees.’ [SB]
- d. *eutā.eutā lāltin bāl-thyo*
single.single lantern light-PST.HAB.3.SG
‘(They) used to light one lantern at a time.’ [V001002005; M26]

Unfortunately, there is no data from the survey or corpus analysis to back up these judgments. I did not include this verb form in the Kathmandu Survey, and it is relatively rare in the NNSP: I coded 50 utterances as past habitual (including the *thyo* form of the copula), of which only 5 were non-modal transitive main clauses, and the subject was elided in every single case (one of these examples is 82d above).

The rarity of the past habitual in transitive clauses is itself noteworthy. Of the fifty examples of this verb form, only 6 were transitive, while 22 were intransitive and 22 were copular. This may be a result of the likelihood for a past habitual to encode backgrounded, scene-setting information.

From the judgments of elicitation consultants I tentatively conclude that there is an ergative/nominative alternation with this verb form as well, although more research on this topic would be helpful.

The Definite Future

The definite future is a periphrastic verb form which consists of the *-ne* nonfinite marker and a present tense copula (typically this form is cited with the *cha* copula, but as we see from (83d) it is also possible with *ho*, as are all periphrastic verb forms). The Definite Future is counted as an imperfective verb form as it does not contain

the *-e-* perfective marker, but multiple elicitation respondents (PK, ST, and TD) expressed the intuition that the ergative is required in all transitive clauses marked with this form that were discussed, while others (BB, SB) found it to be optional. Abadie (1974) describes the ergative as optional here, but my general impression is that it is heavily preferred. In Verbeke and de Cuypere's corpus analysis, they found the ergative marker on 74% of the subjects in definite future clauses (Verbeke and De Cuypere 2015: 11).

- (83)
- a. *bholi socrates-le bikh khā-ne cha*
tomorrow socrates-ERG poison eat-DEF.FUT.3.SG
'Socrates will drink poison tomorrow.' [TD]
 - b. *maile doctor-lāi bhet-ne chu*
I.ERG doctor-ACC meet-DEF.FUT.1.SG
'I will meet the doctor.' [PK]
 - c. *khamda-haru-le hotel ban-aau-ne chan*
worker-PL-ERG hotel make-CAUS-DEF.FUT.3.PL
'The workers will build the hotel.' [PK]
 - d. *ma talab wahā-bāṭa khā-ne ho*
I.NOM salary there-ABL eat-DEF.FUT.1.SG
'I will get my salary (regularly) from there.' [V001002005; M7]

The Kathmandu Survey contained a single question trial on the definite future verb form (E4), but the responses are instructive (Figure 4.5). The average score for both the ergative nominative form are quite high, but the ergative form is heavily preferred. Similarly, all but one respondent rated the ergative form 4 or 5, compared to the three quarters who rated the nominative form 4 or 5. This tells us that both forms are acceptable to most speakers, but the ergative form is heavily preferred.

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
ERG (n=27)	4.56	92.6%	3.7%
NOM (n=27)	3.96	74.1%	11.1%

Figure 4.5: Survey Results: ERG/NOM with Definite Future Transitive Clauses

In the NNSP sample there were 115 examples of the definite future, of which only 15 were in transitive non-modal clauses, and the subject was elided in all but one example (the nominative form in 83d above). This single result does not tell us very much, but the grammaticality judgments from elicitation consultants and survey respondents confirm the findings of Abadie (1974) and Verbeke and De Cuypere (2015) that the ergative is more common on the S_t for this verb form, although the nominative form is also possible.

The Hypothetical Future, Optative, and Imperative Mood

On transitive subjects with the hypothetical future and optative inflectional verb forms, as well as in the imperative mood, ergative marking is variable.

- (84) a. *rām/rām-le* *film her-lā*
ram/ram-ERG film see-HYP.FUT.3.SG
'Ram may see a film.' [TD]
- b. *ke gar-āū* *ta hāmi ?*
what do-HYP.FUT.1.PL F we ?
'What should we do?' [V001001001; M3]
- c. *yo yug-ko bholi jun itihās lekhni-le ke*
this era-GEN tomorrow which history writer-ERG what
bhan-lān ?
say-HYP.FUT.3.PL ?
'What will (future) writers of the history of this era say?' [V001002003;
M13]
- d. *u/us-le* *tyo sarpa mār-os*
PRO.NOM/PRO-ERG this snake kill-OPT.3.SG
'I wish he would kill that snake.' [SB; example from Poudel (2008)]
- e. *je hos, suītar cāhī sāno sāij-mā dekh-āu-nus*
whatever COP.OPT, sweater TOP small size-LOC see-CAUS-IMPER.HON
na
PRT
'Whatever, show me a sweater in a smaller size.' [V001001001; M1]

- f. *aba khānā timi-haru-le haami-lāi gar*
now food you-PL-ERG we-DAT do.IMPER
‘Now, you make food for us.’ [V001002005; M24]

The elicitation consultants found both the nominative and ergative to be possible in the optative and hypothetical future. In the Kathmandu survey, there were no question trials with the imperative or optative but there were three question trials with the hypothetical future. The data in Figure (4.6) indicate a pattern similar but more pronounced than that of the definite future: the ergative is almost universally acceptable, while the nominative is dispreferred in about half of the judgments.⁵

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
ERG (n=54)	4.83	96.3%	1.9%
NOM (n=53)	2.53	24.5%	52.8%

Figure 4.6: Survey Results: ERG/NOM with Hypothetical Future Transitive Clauses

In the NNSP sample, I encoded 82 clauses as hypothetical future of which 63 were in non-modal transitive main clauses, but the subject was elided 94% of the time. Of the remaining four sentences **in the hypothetical future, the ergative was used once.** The hypothetical future is most commonly used in the first person plural in a hortative sense or as a polite way to ask for something (*gar-āūlā!* ‘Let’s do it!’). Of the 82 total hypothetical future clauses, 66 (80%) were with the first person plural subject. The difference in results between the corpus and the survey results is likely attributable to this bias towards the first person plural form in usage: the nominative is more common with first person forms and the ergative is more common with the third person forms that were tested in the survey. This relates to the correlation between the usage of the ergative and the Nominal Hierarchy, which I will discuss later in this chapter.

5. Here I tabulated the results from the two transitive examples F2 and G5. F2 was included in the intransitive section, but as I note in the section on Intransitive observations, I believe the verb *hernu* “to watch, to see” in F2 is transitive.

The optative was found rarely in the corpus sample, and exclusively with the *ho* copula as in (84e) (*je hos* ‘whatever it may be’), in what I consider to be fixed expressions.⁶ In any case, there were no transitive examples with subjects to confirm the judgment from respondents that both the ergative and nominative are possible in sentences like (84d).

I encoded 142 clauses as imperative of which 126 were transitive non-modal main clauses. As might be expected with imperatives, the subject was elided 93% of the time, and **the ergative was found on 3 of the remaining 9 clauses in the imperative.**

The particular usages of these three verb forms are such that the subject is rarely present, but there does appear to be some variability in the usage of the ergative and nominative here as well, with the nominative being somewhat more common.

Subordinate Clauses

In both perfective and imperfective subordinate clauses, there is substantial variation between ergative and nominative expression of the transitive S_t . This is at least partially because a main clause nominative subject can control an ergative subject in a subordinate clause (or an ergative may control a nominative), and either may be elided. This makes it particularly difficult to discern precisely which argument is the subject of a subordinate clause, and in the corpus I erred towards including an overt subject even if it may have originated in the main clause. This is a complex topic that I cover more thoroughly in the syntax section, where I will discuss the judgments of elicitation respondents.

In the NNSP sample, I distinguished three different types of subordinate clauses.

The first of these are adverbial clauses in which the verb form is unmarked for

6. I believe that the form is commonly used in blessings and as a more formal form of the imperative (*garnu hos*, ‘I wish it were that you do (that)’, which I would suspect is the origin of the honorific imperative form *garnus*, ‘Do (that)’).

person or tense but rather takes a (typically perfective) adverbializer. For example, from the verb *khā-nu* ‘to eat’, there is *khā-era* ‘eating and ...’, *khā-dā(kheri)* ‘while eating’, *na-khā-ikana* ‘without eating’, *khā-epacchi* ‘after eating’, *khā-epani* ‘despite eating’, etc. Here are some examples from the NNSP sample with overt subjects. In both (85a) and (85b) the first subordinate clause is transitive and perfective, but in the first example the subject is ergative and in the second it is nominative.

- (85) a. *aba* [*hāmi-le tyas-lāi* *pariwartan gar-dākheri*] *arājāktā*
 now [we-ERG it.OBL-ACC change do-WHILE] chaos
 ā-yo
 come-PERF.3.SG

‘Then, while we were changing it, chaos arose.’ [V001002003; M13]

- b. [*khaire-haru* *ke* *yasto* *ghām-mā kapadā.sapadā phukāl-era*]
 [white.person-PL what like.this sun-LOC clothes.RED remove-CONJ]
 ghām-mā bas-iraheko huncha [*jyān sek-ā-era*]
 sun-LOC lie-PROG.3.SG [body cook-CAUS-CONJ]
 ‘In this heat, white people tend to take off all their clothes and lie in the sun, cooking their bodies.’ [V001002005; M26]

Of the 297 examples of clauses of this type, 171 were non-modal transitive clauses. The subject was elided 74% of the time, and of the remaining 48 examples **the ergative was expressed 54% of the time in adverbial clauses**.

The second form of subordinate clauses is nominal relative clauses, which use either the *-eko* perfective participial or the *-ne* nonfinite marker.

- (86) a. *yahā ta pheri* [*indibhijual na-di-ne* *niem*]_{NP}
 here FOC again [individual NEG-give-NONFIN rule]_{NP}
 cha
 COP.PRES.3.SG

‘Here there is a rule against individuals giving (tips to guides).’ lit. ‘Here an individuals-not-giving-rule exists.’ [V001002005; M7]

- b. *uhā-haru-le* [*agillo tim-le* *gar-eko pariwartan*]_{NP-lāi}
 PRO.HON-PLURAL-ERG [first team-ERG do-PERF change]_{NP-ACC}
 mānnetā di-na ...
 acceptance give-NON.FIN ...

‘They give acceptance to the changes made by the first team...’ lit. ‘They give acceptance to the first-team-made-changes...’ [V001002003; M13]

The corpus contained 67 examples of these clauses, of which 27 were transitive non-modal constructions and 11 of these had overt subjects. There was only one example of a perfective participial with an overt subject, and it is ergative. Among the ten non-perfective forms, the ergative was expressed once.

The final type of subordinate clause uses the non-finite markers (-na, -ne, -nu). In this category I also include the usage of -ne as an unmarked verb form in simple main clauses, as in (87c).

- (87) a. [*gaīdā-le tapāi-lāi hān-na*] *ā-eko cha ki*
 [rhino-ERG you.HON-ACC hit-NON.FIN] come-pres.perf.3.sg or
 chaina āja-samma ?
 COP.3.SG.NEG today-UNTIL ?
 ‘Has a rhino ever come to attack you or not?’ [V001002005; M7]
- b. [*yo nayā cātaut maile cāhī nibedan gar-na*]
 [this new idea I.ERG TOP propose do-NON.FIN]
 khoj-eko chu
 search-PRES.PERF.1.SG
 ‘I have sought to propose this new idea.’ [V001002003; M13]
- c. *hāmi-le di-ne, haina ?*
 we-ERG give-NONFIN, COP.3.SG.NEG ?
 ‘We usually give (that much), right?’ [V001001001; M9]
- d. [*pujāri ra shikshak-le us-lāi samāt-na*] *prayās*
 [priest and teacher-ERG PRO.OBL-ACC catch-NON.FIN] try
 gar-e
 do-PERF.3.PL
 ‘The priests and teachers tried to arrest him.’ [BB]

Of the 201 examples of this type, 134 were transitive non-modal constructions, of which 83% had elided subjects, and of the remaining 23 clauses, the ergative was expressed 52% of the time. Figure (4.7) summarizes these findings.

Case	Adverbials (n=48)	Nominals (n=11)	Nonfinite (n=23)
ERG	54%	18%	52%
NOM	46%	72%	48%

Figure 4.7: Corpus Results: ERG/NOM in Subordinate Clauses

In sum, both the ergative and nominative are possible for all the transitive subordinate clauses examined. The nominalized clauses are a special case, because the nominal subject is typically the head of the noun phrase, and thus the nominative is more common here (the main clause may be intransitive or copular).

It is somewhat surprising that the adverbials should pattern similarly to the non-finite clauses, because most of them are straightforwardly perfective and we should expect the ergative to be categorical.⁷

To understand the picture better, it would be necessary to categorize and tabulate each of the adverbial markers as perfective or non-perfective. However, it is clear that even when the adverbial clearly has perfective reference, there is a nominative/ergative alternation. The most common adverbial subordinate marker in the corpus is *-era*, which clearly has perfective reference, and of the 31 instances of this marker used with a transitive verb and an overtly realized subject in the corpus, the ergative marker appeared 42% of the time. Of course, the issue is complicated by the fact that the S_t of the subordinate clause may or may not be coreferential with the main clause, which may be nominative or ergative.

4.1.2 Intransitive Clauses

According to Li (2007), the *-le* marker may be found on the S_i of certain intransitive clauses in Nepali like the examples below.

7. Many contain the perfective *-e-*. A separate piece of evidence for perfective reference here is the behavior of the verb *jānu* ‘to go’, which has a separate perfective allomorph (*jā/gā*): hence *jā-nu*, *jā-na*, *jā-ne*, but *gā-eko*, *gā-era*, *gā-epani*.

- (88) a. *kukur-haru/kukur-haru-le rāti bhak-tyo*
 dog-PL/dog-PL-ERG night bark-PST.HAB.3.SG
 ‘Dogs would bark at night.’ [PK]
- b. *ma/maile nāc-dai-chu*
 I.NOM/I.ERG dance-CONT-PRES.1.SG
 ‘I am dancing.’ [TD]
- c. *mero didi-ko chorī/chorī-le dherai khel-chin*
 my sister-GEN daughter/daughter-ERG much play-PRES.3.F.SG
 ‘My sister’s daughter plays a lot.’ [ST]
- d. *sahuji/sahuji-le jahile.pani khok-nu huncha*
 shopkeeper/shopkeeper-(ERG) always cough-PRES.3.SG.HON
 ‘The shopkeeper is always coughing.’ [ST]

This would appear to undermine the usage of the term “ergative,” because in such a system we expect (S_i and O) to pattern together in opposition to S_t . If the distribution of *-le* is precisely the same on the S_i of intransitive clauses as it is for the S_t of transitive clauses, then the term “subject marker” might be more apt. Another possibility is that Nepali has a Split-S pattern, in which some intransitive subjects take ergative marking and others do not. In an active-inactive alignment unergative intransitives (clauses in which the S_i is semantically an Agent) are marked but unaccusative intransitives (clauses in which the S_i is semantically a Patient) are not. In this case, the term “agent marker” might be more apt. Alternatively, marking could vary based on some semantic property of the referent, such as volitionality or agency. In Hindi, certain intransitive clauses that describe actions like coughing or tripping can be marked if the action is done volitionally. This is an example of Fluid-S marking.

In Nepali, the distribution of *-le* on the S_i is dependent upon the particular predicate in the clause. The overarching generalization is that there is a great deal of variation: particular lexical items exhibit particular distributions, marking is possible in some verb forms but not others, and different speakers have slightly different intuitions. There is potentially more dialectal variation. One of my elicitation re-

spondents was a clear outlier: MG was the only consultant who dispreferred every instance of ergative marking on an intransitive verb. He grew up in a Gurung-speaking community and moved to the western city of Pokhara as a young adult.

There are, however, some categorical generalizations, and, MG aside, all elicitation respondents followed the pattern described by Li (2007):

- (1) The ergative is disallowed with **unaccusative** predicates (e.g. *paglinu* “melt”, *paknu* “cook”, *khasnu* “fall”, *dubinu* “sink”)
- (2) The ergative is never allowed with **telic unergative** predicates (e.g. *āunu* “come”, *jānu* “go”, *pharkinu* “return”)
- (3) The ergative is possible with **atelic unergative** predicates (e.g. *nācnu* “dance”, *gāunu* “sing”, *nuhāunu* “bathe”, *khelnu* “play”)

These generalizations were mostly borne out in the Kathmandu survey (Figure 4.8), with some exceptions to be discussed below. In the NNSP sample, however, there were almost no examples of ergative marking with intransitives, because motion verbs and unaccusatives were much more common than other types of intransitives.

I coded 463 clauses as intransitive, of which 244 were non-modal clauses with overt subjects in any verb form. These included 22 tokens of stative verbs (*basnu* ‘to sit/stay/live’, *dukhnu* ‘to hurt’, and *suhāunu/sudnu* ‘to suit/match/look nice’), 1 token of a semelfactive (*karāunu* ‘to shout’), 44 tokens of aspectual verbs (*lāgnu* ‘to attach/feel’, *rahanu* ‘to stay/appear’, *roknu* ‘to stop’, *saknu* ‘to finish’, *thālnu* ‘to begin’), 21 tokens of atelic motion verbs (*ghumnu* ‘to visit’, *hidnu* ‘to walk’), 120 tokens of telic motion verbs (*āunu* ‘to come’, *āipugnu* ‘to arrive’, *bāghnu* ‘to escape’, *daudnu* ‘to run’, *jānu* ‘to go’, *pugnu* ‘to reach/arrive’, *niklanu* ‘to come out’, *niskinu* ‘to get out’, *pharkinu* ‘to return’, *uthnu* ‘to get up/arise’) and 47 tokens of other unaccusative verbs (*badnu* ‘to increase’, *katnu* ‘to be deducted’, *milnu* ‘to be arranged’, *motāunu* ‘to get fat’, *parnu* ‘to happen/cost’, *pāunu* ‘to receive’).

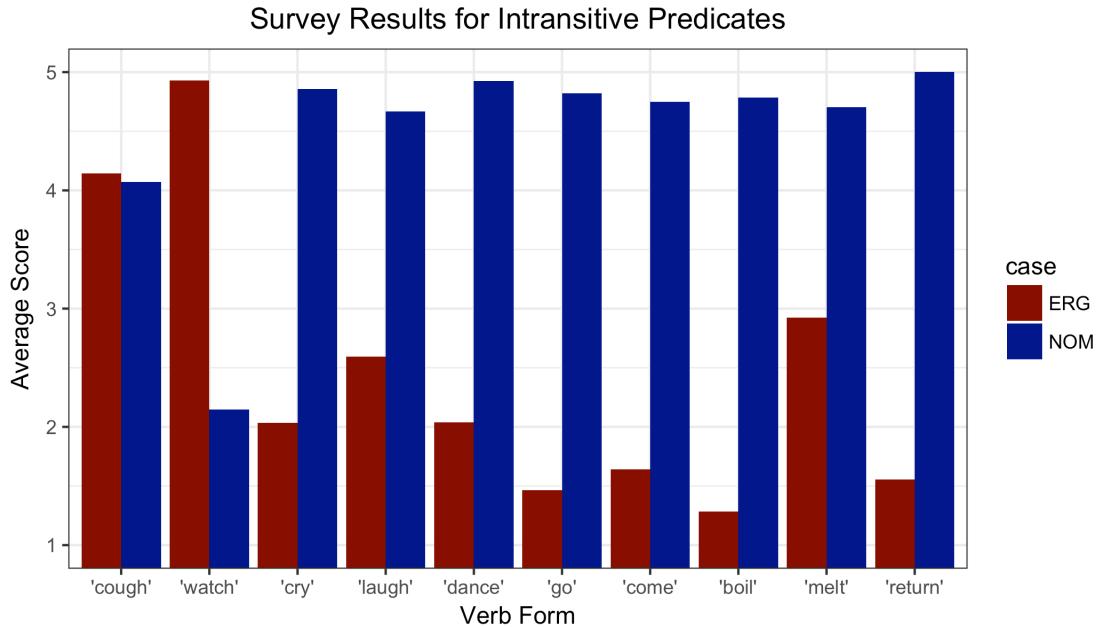


Figure 4.8: Ergative Marking Results for Intransitive Clauses

Out of these 244 clauses, the ergative was present in only ten utterances. Of these ten utterances, seven were cases in which the S_i is coreferential with a transitive subject in another clause. As discussed in the 5.1.1, the ergative is generally allowed in such cases:

- (89) [*tapāī-le* *tyo jasto* “*jangal wāk*” *jā-daakheri*] *bāgh*
 [you.HON-ERG that how “Jungle Walk” go-WHILE] tiger
 dekh-nu bhaeko cha *ki chaina* ?
 see-PRES.PERF.HON.3.SG or COP.PRES.3.SG.NEG ?
 ‘Going on a jungle walk, have you ever seen a tiger?’ [V001002005; M7]

In the next subsections, I will discuss the presence of alternations for unaccusative intransitives, verbs of emission, and telic and atelic motion verbs. Unlike the situation in Hindi, volitionality does not appear to play a role in these alternations. I conclude that every single instance of *-le* on the sole argument of an intransitive clauses should actually be analyzed as ergative marking on the S_t of an underlyingly transitive clause. Thus “ergative” is the best term to describe this marker in Nepali.

Unaccusative Predicates

All elicitation respondents agreed that the ergative could not mark the S_i in any unaccusative intransitive clauses. Note that this prohibition on the ergative holds whether the S_i referent is animate or inanimate, and whether the verb form is perfective or imperfective.⁸

- (90) a. *sikka bu̯i-mā gud-yo*
coin.NOM floor-LOC roll-PERF.3.SG
'The coin rolled across the floor.' [ST]
- b. *jahāj ādi-mā dubi-yo*
ship.NOM storm-LOC sink-PERF.3.SG
'The ship sank in the storm.' [BA]
- c. *ma jāni-jāni lād-ē*
I.NOM purpose trip-PERF.1.SG
'I tripped on purpose.' [TD]
- d. *ghām-mā giu pagli-yo*
sun-LOC ghee.NOM melt-PERF.3.SG
'The ghee melted in the sun.' [BA]
- e. *khānā pak-dai-cha*
food cook-CONT-PRES.3.SG
'The food is cooking.' [BB]
- f. *u motā-yo*
PRO get.fat-PERF.3.SG
'(S)he has gotten fat.' [SB]

In the Kathmandu Survey, questions K3 and K4 examined clauses with the unaccusative imperfective verbs *umlanu* “boil” and *paglinu* “melt” (Figure 4.9). Curiously, the former follows the expected pattern while a surprisingly large number of respondents (roughly half) accepted the ergative on “to melt.”⁹

8. Other unaccusative verbs I examined include *khasnu* “fall”, *tātnu* “heat up”, and *khiinu* “wear out”.

9. After consulting with SB about this, I believe this is due to a flaw in the question design. The

In the NNSP sample, the following unaccusative intransitive verbs appeared with overt subjects in any verb form: *bañnu* ‘to increase’, *duñnu* ‘to hurt’, *katnu* ‘to be deducted’, *milnu* ‘to be arranged’, *motāunu* ‘to get fat’, *parnu* ‘to happen/cost’, *pāunu* ‘to receive’, *roknu* ‘to stop’, *saknu* ‘to finish’, *suhāunu/suñnu* ‘to suit/match/look nice’.¹⁰ In total, there were 56 instances of **unaccusative intransitive verbs with overt subjects, of which 0% were marked with the ergative.**

	Average Score	Like (4 or 5)	Dislike (1 or 2)
umlanu “to boil”			
ERG (n=28)	1.29	3.6%	89.3%
NOM (n=28)	4.79	96.4%	0.0%
paglinu “to melt”			
ERG (n=27)	2.93	48.1%	40.7%
NOM (n=27)	4.70	96.3%	3.7%

Figure 4.9: Survey Results: ERG/NOM with Unaccusative Predicates

In some languages there are transitive predicates in which neither the S_t nor the O are semantically agents. For such transitive unaccusatives, we might expect ergativity to be disallowed in Nepali considering that ergative marking appears to be disallowed with unaccusative intransitives.

Some examples of such predicates in English include *This film stars Jonny Lee Miller*, *I need a passport*, and *These lentils need/lack/require salt*. For Nepali, these sentences are typically expressed with copular constructions (*yo film-ko hero Jonny*

respondents who accepted the ergative here were from many different regions of Nepal including Kathmandu, and SB believes that it is unlikely to be a result of dialect variation. The response included an instrumental (*Chaina, tara ghamko kiran-le ghiu-(le) paglielā* “It hasn’t, but the ghee will melt with the sunshine.”) One respondent commented that there was a mistake on this question because both (a) and (b) were identical (when in fact they differed on whether the *ghiū* had an ergative marker). Every respondent with one exception scored the bare form higher or as high.

10. Many of these verbs are reflective of the fact that half of the sample interviews were recorded in a marketplace. Additionally, the most frequent of these verbs, *milnu* and *pārnū*, which occurred 37 times, have additional aspectual properties and may take clausal complements. I have counted *pārnū* as intransitive in the frame *us-laai das rupiyā pār-cha* (it-DAT ten rupees costs), although Schmidt (1993) calls it transitive. The verbs *roknu*, *saknu*, and *pārnū* are also used in various modal constructions.

Lee Miller ho ‘This film’s hero is Jonny Lee Miller’; *dal-mā nun kami cha* ‘In the lentils salt is lacking’) or impersonal/passive constructions ((*ma-lāi*) *pasport cahinchá* ‘A passport is needed (by me)'). The best examples of such predicates I could find actually require ergative marking, but in each case one could make the argument that the S_t is an agent:

- (91) a. *gaidā-haru-le hāti-lāi gher-e*
 rhino-PL-ERG elephant-ACC surround-PERF.3.PL
 ‘The rhinos surrounded the elephant.’ [BB]

b. *jangal-le upatyakā-lāi gher-yo*
 forest-ERG valley-ACC surround-PERF.3.SG
 ‘The forest covered the whole valley.’ [BB]

c. *kamilā-haru-le khānā-lāi dhāk-e*
 ant-PL-ERG food-ACC cover-PERF.3.PL
 ‘The ants covered the food.’ [BB]

Unergative Semelfactives/Verbs of Emission

For the verbs *khoknu* ‘to cough’, *karaunu* ‘to shout’, and *bhaknu* ‘to bark’, elicitation consultants found the ergative to be possible in both perfective and imperfective verb forms.

- (92) *sahuji/sahuji-le* *jahile.pani khok-nu huncha*
 shopkeeper/shopkeeper-(ERG) always cough-INF HON.PRES.3.SG
 ‘The shopkeeper is always coughing.’ [ST]

(93) *neta/neta-le* *karāu-cha*
 politician/politician-(ERG) shout-PRES.3.SG
 ‘The politician shouts.’ [BA]

(94) *kukur-haru/kukur-haru-le* *rāti bhak-chan*
 dog/dog-PL-ERG night bark-PRES.3.PL
 ‘The dogs bark at night.’ [PK]

There was some variation in the precise judgments: TD expressed the intuition that the ergative is obligatory for *bhaknu* ‘to bark’ in in the perfective, and ST expressed

the same intuition for *khoknu* ‘to cough’, but in general consultants found the optionality to extend to the perfective.

For other predicates, the ergative was categorically rejected in the perfective and imperfective.

- (95) ***paale*** *run-cha* / *ro-yo*
guard cry-PRES.3.SG / cry-PERF.3.SG

‘The guard cries/cried.’ [BA]

- (96) ***chorā*** *chād-cha* / *chād-yo*
child vomit-PRES.3.SG / vomit-PERF.3.SG

‘The child vomits/vomited.’ [ST]¹¹

- (97) ***naika*** *hās-che* / *hās-in*
actress laugh-PRES.3.F.SG / laugh-PERF.3.F.SG

‘The actress laughs / laughed.’ [BA]

The generalization appears to be that predicates which express involuntary actions disallow the ergative marker. Judgments differed between consultants with the verb *kāmnu* ‘to tremble,’ with ST rejecting the ergative and BA accepting it.

	Average Score	Like (4 or 5)	Dislike (1 or 2)
<i>khoknu</i> “to cough”			
ERG (n=28)	4.14	67.9%	10.7%
NOM (n=28)	4.07	75.0%	7.1%
<i>runu</i> “to cry”			
ERG (n=28)	4.86	21.4%	67.9%
NOM (n=28)	2.04	96.2%	0.0%
<i>hāsnu</i> “to laugh”			
ERG (n=27)	4.67	33.3%	51.9%
NOM (n=27)	2.59	100.0%	0.0%

Figure 4.10: Survey Results: ERG/NOM with Verbs of Emission

In the Kathmandu survey, questions F1, F3, and F4 tested respondent judgments on the verbs *khoknu* ‘to cough’, *runu* ‘to cry’, and *hāsnu* ‘to laugh’, respectively.

11. SB notes that this term is a little dated, and the (transitive) *bāntā garnu* ‘to do vomit’ is a better translation for ‘vomit.’

The results (Figure 4.10) follow the same pattern. Sentences with both ergative and nominative subjects score high with *khoknu*, while the nominative is heavily preferred and the ergative dispreferred with *runu* and *hāsnu*.

Verbs of this type were not present in the NNSP sample, with the exception of one instance of *karaunu* ‘to shout’ with an overt (nominative) subject.

Ergativity and Volitionality in the Intransitive Domain

In Hindi certain intransitive verbs have a nominative/ergative alternation conditioned by volitionality. If an action is performed on purpose, the subject may be marked in the ergative. We might expect the same alternation to exist in Nepali, where a wider range of intransitive clauses accept ergative marking. However, this does not appear to be the case:

- (98) *ma jāni.jāni lād-ē*
 I.NOM purposefully trip-PERF.1.SG
 ‘I tripped on purpose.’ [TD]

Elicitation respondents categorically rejected the ergative in (98). Even for intransitives which allow an alternation between nominative and ergative, this alternation does not appear to be conditioned by volitionality. In (99), the context of the first example is that a shopkeeper is politely coughing to get the attention of a browsing customer, while in the second example the shopkeeper is coughing uncontrollably because of a cold. Both ergative and nominative are available in both contexts, which indicates that volitionality is not a conditioning factor here.

- (99) a. *sahuji/sahuji-le jāni.jāni khokh-nu bhayo*
 shopkeeper-ERG purposefully cough-INF HON.PERF.3.SG
 ‘The shopkeeper coughed on purpose.’ [BA]
- b. *ciso-ko karan-le sahuji/sahuji-le khokh-nu bhayo*
 cold-GEN reason-INSTR shopkeeper-ERG cough-INF HON.PERF.3.SG
 ‘The shopkeeper coughed because he had a cold.’ [BA]

This is significant because it suggests that the *-le* marker does not come with a presupposition that the subject is acting volitionally. While volitionality might be a factor in whether ergative marking is available on particular lexical items (*khoknu* ‘to cough’ vs. *runu* ‘to cry’), a more general explanation would simply be that verbs like *runu* and *hāsnu* are unaccusatives in Nepali, while *khoknu* is unergative.

Telicity and Motion Verbs

Li notes that ergative marking is disallowed with unergative predicates which are telic, meaning that they are aspectually bounded. The examples given by Li are all verbs of motion which are bounded by a spatial path, and in agreement with Li my elicitation consultants uniformly rejected the ergative in such cases. The results of the Kathmandu survey (Figure 4.11) confirm these intuitions.

- (100) ***bhariyā*** *pokhara-bāta pharki-yo*
porter pokhara-ABL return-PERF.3.SG
'The porter returned from Pokhara.' [ST]
- (101) ***neta*** *lumbini-mā gā-yo*
politician lumbini-LOC go-PERF.3.SG
'The politician went to Lumbini.' [ST]
- (102) ***bhaīsi*** *ā-yo!*
water.buffalo come-PERF.3.SG
'The water buffalo has come!' [AG]

The ergative is also rejected with other verbs that are bounded by a path of motion: *basnu* in the meaning of ‘to sit’, *uthnu* as ‘to stand/to rise’, *ubhinu* ‘to stand up’, *bhāgnu* ‘to escape’, *ciplinu* ‘to slip’, *niklanu* ‘to come out’, *niskinu* ‘to get out’, *pasnu* ‘to enter’, *khasnu* ‘to drop’.

This is contrasted with atelic motion verbs like *hidnu* ‘to walk’, *ghumnu* ‘to visit’,¹² *dagurnu* ‘to run’, and *kudnu* ‘to rush.’ These verbs also describe motions, but they

12. As opposed to the unaccusative ‘to spin.’

	Average Score	Like (4 or 5)	Dislike (1 or 2)
jānu “to go”			
ERG (n=28)	1.46	7.1%	85.7%
NOM (n=28)	4.82	92.9%	0.0%
āunu “to come”			
ERG (n=28)	1.64	7.1%	85.7%
NOM (n=28)	4.79	96.4%	0.0%
pharkinu “to return”			
ERG (n=27)	1.55	7.4%	81.5%
NOM (n=27)	5.00	100.0%	0.0%

Figure 4.11: Survey Results: ERG/NOM with Telic Intransitives

are unbounded activities, and Li notes that the ergative is possible here. However, the judgments of my elicitation consultants differ from Li’s description: they reject the ergative with atelic motion verbs in sentences like (103).

- (103) a. *u din.bhāri hid-yo*
 PRO.NOM day.full walk-PERF.3.SG
 ‘(S)he walked all day.’
- b. *u kud-dai-cha*
 PRO.NOM rush-CONT-PRES.3.SG
 ‘(S)he is rushing.’
- c. *u stupa-wāri.pāri ghum-yo*
 PRO.NOM stupa-around visit-PERF.3.SG
 ‘(S)he circumnambulated at the stupa.’

However, SB and BB note that the ergative becomes possible with such verbs **if they are specifically construed with a bounded path of motion**, as in (104):

- (104) a. *u/us-le dui kilometer hid-yo*
 PRO.NOM/PRO.OBL-ERG two kilometer walk-PERF.3.SG
 ‘(S)he walked two kilometers.’
- b. *u/us-le saya meter kud-dai-cha*
 PRO.NOM/PRO.OBL-ERG hundred meter rush-CONT-PRES.3.SG
 ‘(S)he is sprinting a hundred meters.’

- c. *u/us-le* *stupa-wāri.pāri das paṭak ghum-yo*
 PRO.NOM/PRO.OBL-ERG stupa-around ten time visit-PERF.3.SG
 '(S)he circumnambulated ten turns at the stupa.'

In fact, SB finds this to be true with the generally stative *vācnu* ‘to survive/live’, which otherwise resists ergative marking:

- (105) *u/us-le* *jiwān rāmrari vāc-yo*
 PRO.NOM/PRO.OBL-ERG life well live-PERF.3.SG
 ‘They lived their life well.’

It would appear that ergative marking is possible with at least some atelic verbs of motion (and states) only when they are construed as transitive (e.g., “walk” vs. “walk two kilometers”; “live” vs. “live your life”). The effect is to provide a bounded path of motion. However, this presents a puzzle. Why should ergative marking be disallowed with lexically-encoded telicity (in verbs like *pharkinu* ‘to return’) but allowed when generally atelic lexical verbs are construed as bounded?

Motion verbs cross-linguistically behave in a way that straddles the boundary between unaccusative and unergative (Levin et al. 1995, Narasimhan et al. 1996). Li, following Perlmutter, includes ‘to go’ and ‘to come’ in a list of unaccusative verbs, but also refers to them as taking agentive subjects (in contrast with more clearly unaccusative predicates like *bhatkinu* ‘to collapse’).

It is likely that verbs like *āunu* ‘to come’, *jānu* ‘to go’, and *pugnu* ‘to reach’ are unaccusative, because they are quite often found with non-agentive subjects. For example, when water is being poured into a cup at a meal, the cup’s recipient might say *Pāni pugyo*. “That’s enough water” (lit. “The water has arrived/reached”). When electricity is restored after a blackout, one hears the cry “*Bātti āyo!*” “The lights have come!” When giving directions, the subject of these verbs is often a location, e.g. *Dui kilometer pacchi pul āuncha* “After two kilometers a bridge comes.”¹³ These are all

13. These examples come from personal experiences in Nepal.

examples in which the literal English translation is infelicitous because “go” and “come” typically require a volitional subject. So the issue with most of these telic motion verbs is not that they are telic but rather that they are (lexically) unaccusative whether the subject is agentive or not.

Unergative Atelic Predicates and “Do” Light Verbs

Elicitation respondents generally find the ergative to be possible on verbs which describe activities. The exception below is (106d), for which RM strongly preferred the nominative form, but in other situations *khelnu* ‘to play’ may take an ergative marker.¹⁴

- (106) a. *ma/mailə naac-dai-chu*
I.NOM/I.ERG dance-CONT-PRES.1.SG
'I am dancing.' [TD]
- b. *pāle/pāle-le dhyān di-era her-dā-cha*
guard/guard-ERG attention give-CONJ watch-ARCH.PRES.3.SG
'The guard always watches carefully.' [BA]
- c. *prayek bihāna sahuji/sahuji-le nuhān-cha*
every morning shopkeeper/shopkeeper-ERG bathe-PRES.3.SG
'The shopkeeper bathes every morning.' [BA]
- d. *sab bacā-haru bahira khel-irah-e-chan*
all child-PL outside play-PROG-PERF-PRES.3.PL
'All the children have been playing outside in the courtyard.' [RM]

In the Kathmandu survey, on the other hand, the ergative was strongly preferred in an example with *hernu* ‘to watch’, while it was strongly dispreferred with *nācnu* ‘to dance.’

14. Other predicates in which I observed a nominative/ergative alternation include *gāunu* ‘to sing’, *dhyān garnu* ‘to focus’, *jagada garnu* ‘to fight’, *relax garnu* ‘to relax’, *sunnu* ‘to hear’, and *pāuḍnu/pāuḍi khelnu* ‘to float/to swim’.

	Average Score	Like (4 or 5)	Dislike (1 or 2)
hernu “to watch”			
ERG (n=28)	4.93	100.0%	0.0%
NOM (n=28)	2.15	14.8%	63.0%
nācnu “to dance”			
ERG (n=28)	2.03	19.2%	6.9%
NOM (n=28)	4.92	100.0%	0.0%

Figure 4.12: Survey Results: ERG/NOM with Unergative Atelic Verbs

In this domain there is more individual variation in speaker judgments. For example, TD expressed the intuition that the ergative is required in the perfective for *nācnu* ‘to dance.’ Thus, it has the same pattern as a transitive verb:

- (107) *hijo mero *bhāi/bhā-le naac-yo*
 yesterday my *little.brother/little.brother-ERG dance-PERF.3.SG
 ‘Yesterday my little brother danced.’

ST and BA did not share this intuition, finding the bare form to be acceptable in perfective verb forms.

Some of these verbs that I have categorized as intransitive are probably underlyingly transitive. The object may be elided particularly if it is indefinite, so a transitive verb may have only one overt argument. This is almost certainly the case with *khelnu* ‘to play,’ which for most if not all speakers requires an ergative marker when it is overtly transitive:

- (108) *maile hijo bakunda khel-eko thiē*
 I.ERG yesterday football play-pst.perf.1.sg
 ‘I was playing football yesterday.’ [RM]

This is likely also the case with *hernu* ‘to watch’ and *sunnu* ‘to listen’. The argument of these predicates commonly takes an ergative case marker in the perfective whether or not there is an overt object. However, for some speakers *khelnu* can also be construed of as intransitive. BA noted, with some surprise, that the nominative form was perfectly acceptable:

- (109) *prahāri khel-yo.*
 police.officer play-PERF.3.SG
 ‘The police officers played.’ [BA]

This is the source of the observed variation between speakers. Verbs which are transitive may be construed as intransitive in certain situations, and this allows for a nominative form of the S_i in the perfective.

In fact, nearly all of the activity verbs in which there is a nominative/ergative alternation may be considered transitive: for *nuhānu* ‘to bathe’ there is an implied reflexive, while the object of *gānu* ‘to sing’ and *nācnu* ‘to dance’ is a song and a dance. The underlying object for verbs of emission is that which is emitted: the spit in *thuknu*, the noise of laughter in *hāsnu*, or the cry in *karañnu*.

In Nepali there is a productive process of verb derivation which consists of an uninflected noun and the verb *garnu* ‘to do.’ Some examples include *māya garnu* ‘to love’, *dhyān garnu* ‘to focus/meditate’, *jagada garnu* ‘to fight’, *biswās garnu* ‘to believe’, *prayog garnu* ‘to try’, *kam garnu* ‘to bargain’, *nibedhan garnu* ‘to propose’, *gaph garnu* ‘to gossip’, and *chalphal garnu* ‘to discuss.’ It is also common in loanword derivations, for example *calculate garnu* ‘to calculate’ (found in the corpus) and *relax garnu* ‘to relax.’ The incorporated noun cannot take case marking or inflection, and if the derived predicate is transitive, the overall behavior of the S_t and O is as it is in any ditransitive verb. However, if the derived predicate is intransitive (as in *dhyān garnu* or *relax garnu*), ergative marking is still required in the perfective. So syntactically these light verb constructions are transitive.

However, there are some unergative activities for which ergative marking is completely disallowed. Li notes that *sutnu* “to sleep” and *udnu* “to fly” do not take ergative marking, and I add to this list *luknu* “to hide.” My elicitation consultants were in agreement that ergative marking is not possible in these clauses. These are also all verbs in which there is not plausibly an underlying object.¹⁵ Taken together,

15. The verb *udnu* ‘to fly’ can only refer to an entity like a bird flying under its own power. The

the natural conclusion is that most if not all ergative marking on intransitives is in fact ergative marking on underlyingly transitive clauses.

Ergativity and Transitivity

If we consider all clauses with ergative/nominative alternations to be underlyingly transitive, then we do not need to appeal to telicity or agentivity to explain the split between so-called intransitives which have ergative/nominative alternations and those which do not:

- (1) Unaccusative intransitives do not ever take ergative marking, because *-le* seems to be restricted to underlying external subjects.
- (2) Telic motion verbs like *jānu* ‘go’ and *pugnu* ‘reach/arrive’ are unaccusative.
- (3) Activity verbs like *khelnu* ‘to play’ and *gāunu* ‘to sing’ are typically transitive, but for some speakers they can be construed as intransitive in the right circumstances.
- (4) Unergative motion verbs like *hidnu* ‘to walk’ are typically intransitive, but for some speakers they can be construed as transitive in the right circumstances.

This analysis stipulates that all transitive clauses in which we find an ergative case marker on the subject contain two participants, an S_t and an O, whether or not they are overtly realized. However, this is not the case for copular clause, because they do not describe an event but rather describe a state that holds for a particular entity:

4.1.3 Ergativity in Copular Clauses

Ergativity is completely disallowed in all copular clauses of the language. All elicitation respondents uniformly reject ergative marking in any type of copular clause:

verb *udāunu* ‘to fly’ with the causative morpheme is used for operating a flying machine like a plane.

	Average Score	Like (4 or 5)	Dislike (1 or 2)
J1			
ERG (n=28)	1.18	0.0%	92.9%
NOM (n=28)	5.00	100.0%	0.0%
J2			
ERG (n=26)	1.31	0.0%	92.3%
NOM (n=27)	4.37	85.2%	11.1%
J3			
ERG (n=28)	1.54	10.7%	85.7%
NOM (n=28)	4.93	92.9%	3.6%
J4			
ERG (n=28)	1.54	10.7%	85.7%
NOM (n=28)	5.00	100.0%	0.0%
J5			
ERG (n=27)	1.22	0.0%	96.3%
NOM (n=27)	4.93	100.0%	0.0%

Figure 4.13: Survey Results: ERG/NOM with Copulas

- (110) *sunita chalākh cha*
 Sunita-ERG clever COP.PRES.3.SG
 ‘Sunita is clever.’ [AG]
- (111) *ma kabir ho*
 I poet COP.PRES.1.SG
 ‘I am a poet.’ [BA]
- (112) *tyo kholā thulo huncha, haina ?*
 that river big COP.FUT.3.SG, COP.3.SG.NEG ?
 ‘That river becomes big (swells), right?’ [V001002005; M24]
- (113) *baghutā ekdam.ai kushi bhayo*
 frog very.EMPH happy COP.PERF.3.SG
 ‘The frog became very happy.’ [AG]
- (114) *yo-bhandā kati garmi thyo butwal-tira*
 this-COMP how hot COP.PST.3.SG Butwal-toward
 ‘Compared to this, how hot it was near Butwal.’ [V001002005; M24]
- (115) [*hāmi ghum-eko chaū*], [*dui-tin jānā bha-era*],
 [we visit-PRES.PERF.1.PL], [two-three person COP-CONJ],

jangalai tetti-kai, hātti.sātti na-li-ikana
jungle.EMPH that-GEN.EMPH, elephant.RED NEG-take-WITHOUT

We wandered, being two or three people, like that in the jungle, without taking an elephant or anything. [V001002005; M7]

Section J of the Kathmandu Survey (J1-J5) examined copular clauses. The results for each question are in Figure (4.13). The ergative is markedly dispreferred. In the NNSP sample I coded 703 clauses as copular, of which 481 (68%) had overt subjects. Out of these 481 copular clauses, only one clause had an ergative marker on the argument.

Brief observations on the dual copula system

There are two present-tense copulas *cha* and *ho*, which Butt and Poudel (2007) characterize as **stage-level** and **individual-level** copular predicates respectively. In many cases they do seem to distinguish between temporary stages and inherent, lasting properties:

- (116) a. *ma birāmi chu/#ho.*
I sick COP1.1.SG/COP2
'I am sick (currently).' [RM]
- b. *ma jahile birāmi ho.*
I always sick COP2
'I am always sick.' [RM]
- (117) a. *mero ghar amerikā ho/#cha.*
my home America COP2/COP.3.SG
'My home is America.' [RM]

The grammars of Acharya (1991) and Schmidt (1993) note that *cha* and *ho* are present tense forms of the verb *hunu* 'to be.' The meaning of *cha* is described as **existential** and the meaning of *ho* as **identificational** (Acharya 1991: 154-155). While *cha* inflects for person, gender, number, and honorificity, *ho* typically only inflects for

number.¹⁶

The *cha* form of the copula is used to express possession (both alienable and inalienable). It is also typically the form used with adjectival copular clauses and locational copular clauses.

(118) Possession (alienable and inalienable)

- a. *mero gadi cha*
my car COP.PRES.3.SG
'I have a car.' [RM]
- b. *merā dui jānā bhāi chan*
my.PL two CT younger.brother COP1.PRES.3.PL
'I have two younger brothers.' [RM]

(119) Adjectival copulas (both stage-level and individual level)

- a. *yo kukur birāmi cha*
this dog sick PRES.3.SG
'This dog is sick.' (Stage-level) [RM]
- b. *yo camij nilo cha*
this shirt blue PRES.3.SG
'This shirt is blue.' (Individual-level) [RM]

(120) Location: (both animate and inanimate)

- a. *prakāsh bhānsā.khotā-mā cha.*
prakash kitchen-LOC COP1.3.SG
'Prakash is in the kitchen.' (Animate) [RM]
- b. *pāras france-mā cha.*
paris france-LOC COP1.3.SG
'Paris is in France.' (Inanimate) [RM]

From these examples it is clear that *cha* can be used with individual-level predictions.

It is in fact the default for copulas describing properties both temporary and lasting.

16. However, the plural *hun* may also be used as an honorific, and the 1st person singular is pronounced either *ho* or *hū*. Acharya (1991) does give a full inflectional paradigm for *ho*, but my impression from conversation and corpus analysis is that these inflections are quite rare in spoken Nepali.

The *ho* form of the copula is the only copula available for nominal copular clauses:

- (121) *mero ghar america ho / #cha*
my home america COP2 / #COP1.PRES.3.SG
'My home is America.' [RM]
- (122) *tapāi̯ doctor hun / #hunu-huncha*
you doctor COP2.HON / #COP1.PRES.2.SG.HON
'You are a doctor.' [RM]

However, in addition to being the required copula with nominal copular clauses, *ho* is also possible in every environment in which *cha* is used: adjectival copulas, possession, and locative copulas. While *ho* can be used anywhere, *cha* is somewhat restricted.

(123) Possession (alienable and inalienable)

- a. *mero gadī ho*
my car COP2
'I have a car.' [RM]
- b. *merā dui jānā bhāi hun*
my.PL two CT younger.brother COP2
'I have two younger brothers.' [RM]

(124) Adjectival copulas (stage-level and individual level)

- a. *yo kukur birāmi ho*
this dog sick COP2
'This dog is sick.' [RM]
- b. *yo camij nilo ho*
this shirt blue COP2
'This shirt is blue.' [RM]

(125) Location: (both animate and inanimate)

- a. *prakāsh bhānsā.khotā-mā ho.*
prakash kitchen-LOC COP2
'Prakash is in the kitchen.' [RM]
- b. *pāras france-mā ho.*
paris france-LOC COP2

‘Paris is in France.’ [RM]

A better way to characterize this distinction is as an instantiation of a contrast between **characterizing** and **particular** claims about the subject. A characterizing claim holds over a disparate and larger interval than a particular claim, and *ho* would appear to be restricted to characterizing claims, while *cha* can be either. Deo (2017) takes this distinction to characterize the double copula system of Marathi, and we might expect similar distinctions to be found across Indo-Aryan languages. A somewhat similar distinction is proposed by Roy (2013) as a more precise formulation of stage/individual contrasts in Irish, French, and Spanish.

More work needs to be done to adequately characterize the nuances of semantic meaning and pragmatic usages of the two copulas.¹⁷ I bring the distinction up here to emphasize that even in the copular domain the difference is not simply a stage-level/individual-level distinction.

4.1.4 Summary of the domains of ergative marking

Ergative marking is restricted to transitive clauses and some unergative intransitive clauses which are arguably transitive underlyingly. It is obligatory in clauses with perfective reference and varies with the nominative elsewhere. Motion verbs, which are telic and possibly unaccusative in Nepali, pattern with the unaccusatives. This is summarized in Figure (4.14), in which light regions represent obligatory marking, dark regions represent disallowed marking, and grey regions represent a possible alternation between marked and unmarked subjects.

I provided an example for each cell of the chart. For transitive clauses, the verb is *garnu* ‘to do.’ For unergative intransitives, the verb is *khelnu* ‘to play.’ For unac-

17. This is a dissertation-length topic in itself. See Sánchez-Alonso (2018) for a treatment of the semantic and pragmatic features of the dual copula system in Spanish varieties, and Harris (2019) on properties of the copular system in African-American English.

cusative intransitives, it is *aaunu* ‘to come.’

TRANSITIVE		INTRANSITIVE		COPULAR
IMPERFECTIVE	Unergative		Unaccusative	
	MAIN			
	MAIN			
	SUBORDINATE			
	<i>X-(le) garcha</i>	<i>X-(le) khelcha</i>	<i>X aauncha</i>	<i>X cha</i>
	<i>X-(le) garne</i>	<i>X-(le) khelne</i>	<i>X aaune</i>	<i>X bhanne</i>
	<i>X-le garyo</i>	<i>X-(le) khelyo</i>	<i>X aayo</i>	<i>X bhayo</i>
	<i>X-le garera</i>	<i>X-(le) khelera</i>	<i>X aaera</i>	<i>X bhaera</i>

Figure 4.14: Summary of the Domains of Ergativity in Nepali Clauses

For imperfective verb forms, I gave a main clause example of the verb in simple form and a subordinate clause with the *-ne* marker. In the perfective domain, I used the perfective form and the conjunctive *-era*. Note that in the perfective subordinate clause cell I have marked *-le* as alternating, although whether or not alternation is possible may be entirely determined by the properties of the main clause.

For the rest of this section, I restrict my focus to the gray regions of nominative/ergative alternation to investigate whether properties of the event structure correlate with the expression of the ergative. In particular, I am interested in the top left cell representing the alternation in the transitive imperfective verb forms of main clauses.

4.1.5 Individual-Level Predication

Depending upon the context, verbs in the simple present form may refer to events which are **ongoing**, **future-oriented** or **habitual**. The following examples are taken

from a list of the fifty-four main clause transitive sentences with present imperfective reference in the NNSP sample.

- (126) a. *her-erai kurā gar-chu ma*
see-CONJ.EMPH conversation do-PRES.1.SG I
'Having seen (you), I'm talking to you.' (Ongoing) [V001001001; M3]
- b. *ma pheri ghar-mā gā-era phon gar-chu*
I again home-LOC go-CONJ phone do-PRES.1.SG
'I will go back home and then call you.' (Future) [V001002003; M13]
- c. *tyas-le ek ghanṭā tin liṭar khā-idin-cha*
it.OBL-ERG one hour three liter eat-BEN-PRES.3.SG
'It (the generator) consumes three liters (of petrol) an hour.' (Habitual)
[V001002005; M7]

The **individual-level predication** theory presumes that the ergative will be found with the habitual interpretation of the simple present, because the event is a lasting property of the subject rather than a single spatio-temporally constrained event. This is true for the examples presented above. The ergative marker is present on the subject in (126c), the sentence with a habitual reading, and it is absent in the examples that describe ongoing and future-oriented events.

On the one hand, this theory does seem to capture a common intuition that the form with the ergative describes an event that is some way more central to the identity of the subject. Furthermore, Verbeke and De Cuyper (2015) conclude that individual-level predication is significantly correlated with ergative-marking in their corpus analysis.

However, we saw from the section 4.1.1 that the ergative/nominative alternation exists for imperfective verb forms which are unambiguous. For example, the alternation exists for the continuous form, which is unambiguously stage-level. So even if this is an explanation in the simple present, this is an explanation for a limited portion of the imperfective domain. This is reminiscent of the argument that the

ergative discriminates participants when there is a possibility of confusing them: it might explain ergative marking in those particular cases, but it cannot be the entire story.

A bigger issue is that while ergative marking can be associated particularly with habituality in one interpretation, the *nominative* can be associated with habituality under a different interpretation.

Judgments on Ergative Marking and Habitual Interpretations

I presented each of my elicitation consultants with some version of one of the Nepali minimal pairs in (127a), (127b), and (127c).

- (127) a. *mero kākā-(le) rājdhāni-mā gādi cil-āu-nu huncha*
my uncle-(ERG) capitol-LOC car drive-CAUS-PRES.3.SG.HON
'My uncle is driving/will drive/drives cars in the capitol.' [BA]
- b. *prakāsh-(le) curot khān-cha*
Prakash-(ERG) cigarette eat-PRES.3.SG.NEG
'Prakash is smoking/will smoke/doesn't smoke cigarettes.' [PK]
- c. *sunita-(le) kitāb lekh-cha*
sunita-(ERG) book write-PRES.3.SG
'Sunita is writing/ will write/ writes a book.' [TD]

Many consultants immediately noted a difference in interpretation with the ergative version that had something to do with habituality. MG, UK, and ST noted that the interpretation for (127a) with the ergative is that this is the man's occupation, whereas without it the interpretation is that he will perform the action later. Similarly, in (127c) implies that Sunita is a poet.

PK's intuition for (127b) was slightly different: a habitual interpretation is likely in both cases, but the unmarked form suggests that smoking is just a general tendency, while the ergative marked form suggests that smoking is his habit, i.e. that *Prakāsh curot khāne maanche ho* ('Prakash is a cigarette smoker.') It is something more central

to his character, and as such the statement is likely to be a negative value judgment about his character. It is an interpretation of the predicate as characterizing. ST, on the other hand, had the intuition that the reading was ongoing in both cases, but with the unmarked form Prakash is currently smoking by himself, and with the marked form he is under some external pressure to smoke, perhaps from a group of friends. In this case the subject is specifically not the initiator of the event, but is rather the effector of the event.

Other consultants did not find habituality to be an differing factor in these sentences. Surprisingly, TD, UK, BB, and SB had the opposite intuition: it is the bare form of the transitive subject that suggests a habitual interpretation, while the ergative-marked subject is amenable to a future or ongoing interpretation. SB observed that the ergative suggests a particular instance of the event, while the bare form suggests an action:

- (128) a. *u kām gar-cha*
 PRO.NOM work do-PRES.3.SG
 '(S)he does work.' [SB]
 - b. *us-le kām gar-cha*
 PRO.OBL-ERG work do-PRES.3.SG
 '(S)he does a job.' [SB]
- (129) a. *u māsu khān-daina*
 PRO meat eat-PRES.3.SG.NEG
 '(S)he eats meat (i.e., is not a vegetarian).' [SB]
 - b. *us-le yo māsu khān-daina*
 PRO.OBL-ERG this meat eat-PRES.3.SG
 '(S)he will not eat this meat.' [SB]

The interpretation of the object is relevant here. Note that *kām* can be glossed in English as ‘job’ or ‘work,’ and when the subject is nominative, the interpretation of the object is indefinite, that is, the subject does work generally, while with the ergative the interpretation is (nonspecific) definite: doing a particular work. In (129b) the

presence of an unambiguously definite object (*yo māsu*, ‘this meat’) correlates with ergative (SB finds the nominative form here to be awkward), while the bare form correlates with a general activity of meat-eating.

These observations are in line with the predictions of the Transitivity Hypothesis, because there is a positive correlation between one marker of high transitivity, the presence of the ergative, and another marker of high transitivity, individuation of the object and/or boundedness of the event. The interpretation of *kām* as ‘work’ is closer to an intransitive interpretation “(S)he works” precisely because the object is less individuated. Conversely, the absence of the ergative correlates with an interpretation that can be interpreted as intransitive. The opposite judgment, that the ergative correlates with a habitual interpretation, is more problematic for the Transitivity Hypothesis.

This variation does not appear to correlate with the speaker being from a particular geographical region or background. It is not the case that different speakers have internalized a different grammatical rule. Rather, these varying intuitions indicate that the phenomenon under consideration is based on a pragmatic inference rather than a semantic correlation. Without a particular context, there are many possible interpretations of these situations, and while it may be the case that simple present clauses with an ergative subject are somewhat more likely to be interpreted as individual-level, it is quite possible for the ergative to be present without an individual-level interpretation, and also possible for the ergative to be absent without a stage-level interpretation. Consider the following:

- (130) a. Context: My friend and I are being served food. My friend is not a vegetarian generally, but is avoiding meat at the moment. I need to advise my hosts not to serve my friend goat curry.
- b. *u/us-le māsu ahile khān-daina tara sākāhāri*
PRO/PRO.OBL-ERG meat now eat-PRES.3.SG.NEG but vegetarian

hoīna
COP.3.SG.NEG

‘(S)he won’t eat meat now, but (s)he’s not a vegetarian.’ [BB]

If the usage of the ergative presupposes an individual-level interpretation of the event, then the version of (130b) with an ergative should be infelicitous in this context. In fact, the sentence should be a contradiction, given the assumption that “not eating meat” and “being a vegetarian” are the same thing. Most speakers who have an intuition *-le* about habitual readings nevertheless accept both versions of (130b) as acceptable sentences.¹⁸ This tells us that the usage of *-le* is not associated with a presupposition.

In Chapter 6 I argue that the interpretation that accords with the Transitivity Hypothesis is due to the meaning of the marker itself as an effector of the event. Emphasizing the subject as an effector is associated with an event in which the effect and completion of the event is profiled rather than its initiation. The other interpretations in which the ergative correlates with a characterizing predicate are due to an interpretation in which the marked element is given prominence as the subject of a categorical proposition.

In the Kathmandu survey, I attempted to investigate a correlation between ergative marking and habitual readings by setting up three question pairs in which the responses were identical and the questions set up an individual or stage-level context:

C1

Q: *tapāi-ko kākā ke garnu huncha?* What does your uncle do?

A: *mero kākā-(le) rājdhāni-mā gādi calāunu huncha.* My uncle drives a car in the capitol.

18. SB was an exception, finding the bare form to be a contradiction.

C2

Q: *tapāī-ko kākā bholi ke garnu huncha?* What will your uncle do tomorrow?

A: *mero kākā-(le) rājdhāni-mā gādi calāunu huncha.* My uncle will drive a car in the capitol.

C3

Q: *surya sākāhāri ho ki?* Surya is a vegetarian, right?

A: *hoina, surya-(le) māsu khāncha.* No, Surya eats meat.

C4

Q: *surya timilāi tarkāri thapi diū?* Surya, should I give you more vegetables?

A: *hoina, surya-(le) māsu khāncha.* No, Surya will eat meat.

C5

Q: *prabhu upanyāskār ho ki?* Prabhu is a novelist, right?

A: *ho, prabhu-(le) māyā prem sambandhi upanyās lekhcha.* Yes, Prabhu writes romance novels.

C6

Q: *ki prabhu-le pheri upanyās lekhcha?* Is Prabhu going to write another novel?

A: *ho, prabhu-(le) māyā prem sambandhi upanyās lekhcha.* Yes, Prabhu will write a romance novel.

As I mentioned in 2.3.2, there are some issues with spelling and unnatural phrasing with some of these questions (particularly with C3 & C4), so the results of the survey must be viewed with some caution. However, if the question itself is poorly phrased then this should be reflected in a low response for both the nominative and the

ergative judgments overall, and in general this is not what we find.

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
C1 (habitual)			
ERG (n=27)	4.41	88.9%	7.4%
NOM (n=27)	4.30	85.2%	3.7%
C2 (future)			
ERG (n=28)	4.04	75.0%	17.9%
NOM (n=27)	4.00	70.4%	14.8%
C3 (habitual)			
ERG (n=28)	4.43	89.3%	7.1%
NOM (n=28)	3.93	71.4%	14.3%
C4 (future)			
ERG (n=28)	4.25	85.7%	14.3%
NOM (n=28)	4.00	78.6%	10.7%
C5 (habitual)			
ERG (n=26)	4.38	88.5%	4.8%
NOM (n=26)	4.04	69.2%	7.7%
C6 (future)			
ERG (n=28)	4.50	85.7%	0.0%
NOM (n=28)	3.54	57.1%	17.9%

Figure 4.15: Survey Results: ERG/NOM with Simple Present Transitive Clauses

The results from Figures (4.15) and (4.16) do not align with the theory that ergative marking correlates with a habitual reading, nor with the theory that nominative marking correlates with a habitual reading. While some pairs were rated higher than others overall, in most cases both the ergative and nominative were rated highly in both situations. This indicates that all responses are generally acceptable to most speakers. The overall trend is that the ergative is simply judged slightly higher overall for both individual-level and stage-level readings of the given sentences. The largest disparity is in C6.¹⁹

We might expect there to be individual variation in whether the ergative or the

19. This may be because of the “*Le* in the question requires a *-le* in the answer” rule; in general I tried to avoid putting an ergative in the question, but in C6 the subject in the question is marked ergative, as opposed to nominative in the other questions. Yet even when the question subject is nominative there is still an overall bias towards the ergative in each case.

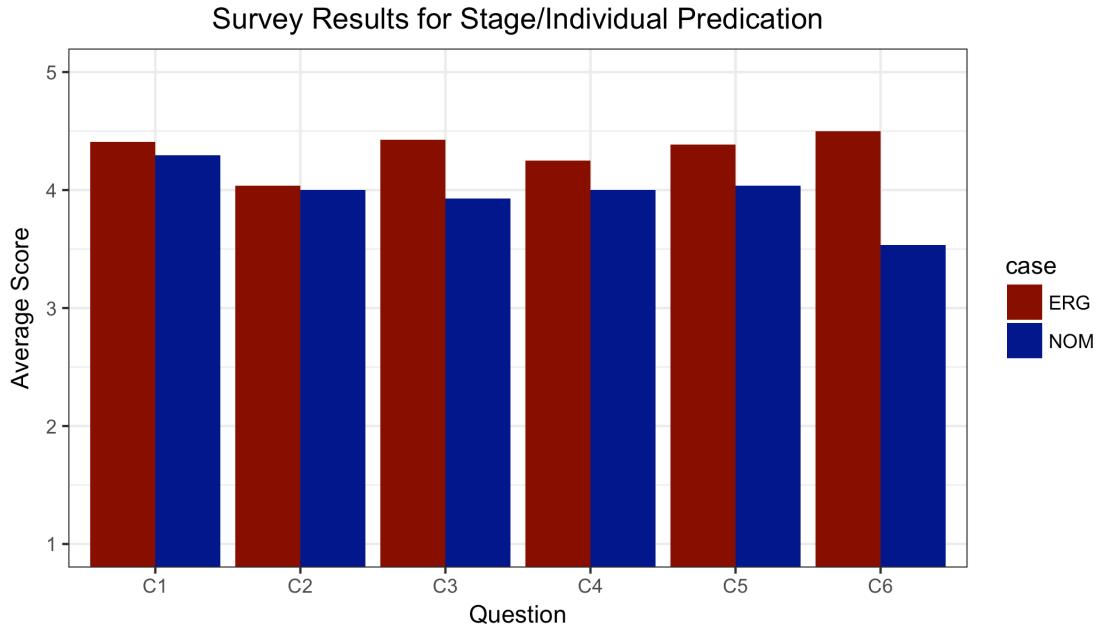


Figure 4.16: Ergative Marking Results in Survey Section C

Question	ERG on Habitual NOM on Future	NOM on Habitual ERG on Future	ERG always Preferred	NOM always Preferred	All Equally Preferred	Other Patterns
C1 & C2	4	2	9	6	2	5
C3 & C4	1	1	13	3	2	8
C5 & C6	1	1	13	3	3	7

Figure 4.17: Survey Results: Individual Strategies for Ergative Marking in Section C

nominative is correlated with the habitual reading. It is possible that this aggregate of responses could hide a number of distinct individual patterns. Figure (4.17) is a breakdown of individual strategies. For each question, I looked at each individual's responses and noted whether (a) the ergative judgment was higher on the habitual reading and lower on the future reading; (b) the ergative judgment was higher on the habitual reading and lower on the future reading; (c) the ergative judgment was higher for both readings; (d) the ergative was lower for both readings; (e) the ergative and nominative judgments were identical in both cases; or (f) there was some other pattern (e.g., the respondent didn't answer a question, or the ergative and nominative were judged the same in one reading but not another). The most common recognizable pattern was that the ergative was preferred in each case, followed by the nominative always being preferred. The least common patterns were the ones in which ergative is

rated higher in one reading and lower in another. And no respondent followed exactly the same strategy for each of the three questions. The Kathmandu survey clearly did not find a correlation between individual-level predication and ergative marking.

In the NNSP sample, there were fifty-four transitive main clauses with predicates in the simple present verb form. For each of these, I noted whether the referenced events in the context of the discourse would be construed as habitual, ongoing, or future-oriented. While determining this is not always straightforward, there were some unambiguous cases of habitual readings with both nominative and ergative subjects:

- (131) a. *bhitra.bhitra tyo gaīdā āl-mā khel-cha, haina*
 inside.RED there rhinoceros lake-LOC play-PRES.3.SG, COP.3.SG.NEG
 ? *tyo tyo khān-eko thyā.*
 ? there there dig-PST.PERF.3.PL

'Way inside there the rhinos play, right? We had dug right there.' [V001002005; M7]²⁰

- b. *jhusilkirā-le tyo hāt-mā chun-cha*
 caterpillar-ERG there hand-LOC sting-PRES.3.SG

'Caterpillars sting you there on the hand.' [V001002005; M7]

- c. *din-dainan turist-haru-le tips.sips jangal*
 give-PRES.3.SG.NEG tourist-PL-ERG tips.RED jungle
 jā-ne-haru-lāi
 go-NON.FIN-PL-DAT

'The tourists do not give tips or anything to those who go into the jungle.'

[V001002005; M8]

And there are some unambiguous cases of future-oriented readings with both nominative and ergative subjects:

- (132) *tyāg ta ajhai tapāi-le khoj-nu huncha bhane, ma*
 tag FOC still you.HON-ERG search-PRES.3.SG.HON COND, I

20. Arguably this is an intransitive construal of *khelnu* which would be an alternate explanation for the lack of ergative marking on the subject.

lyā-idin-chu

bring-BEN-PRES.1.SG

‘If you’re still looking for the tag, I’ll bring it for you.’ [V001001001; M3]

- (133) *kasari - ma udāharan din-chu*

how - I.NOM example give-PRES.1.SG

‘How - I’ll give you an example ...’ [V001002003; M13]

- (134) *sāno bha-epani lāun-dina maile ta yo ta*

small COP.PERF-EVEN wear-PRES.3.SG.NEG I.ERG FOC this FOC

‘Even if it is smaller, I will not wear this.’ [V001001004; M9]

- (135) *yo lā-nu par-yo bhane, tei lān-chu hai*

this take-NON.FIN need-PERF.3.SG COND, that.EMPH take-PRES.1.SG PRT

maile

I.ERG

‘If I have to take this one, I’ll that one as well.’ [V001001004; M7]

The tabulation of these results in Figure (4.18) indicates that ergative marking is slightly more common with habitual readings, though it is not excluded from future and ongoing readings, and in fact the ergative form is in the majority for each case.²¹

Case	Future-oriented	Habitual	Ongoing	Stage-Level	Individual-Level
NOM	n=12	n=5	n=1	n=13	n=5
ERG	n=18	n=20	n=2	n=20	n=10
ERG Percentage	60.0%	66.7%	66.7%	60.6%	66.7%

Figure 4.18: Corpus Results: Individual Strategies for Ergative Marking in Section C

As I discuss in section 4.2.3, the ergative is more common when the subject is interpreted as a kind (as in 131b), and the predicate in such cases will be individual-level and have a habitual construal.

Overall the correlation between individual-level predication and ergative marking is not strong, and is certainly not categorical. Rather, the presence of the ergative

21. I consider the ongoing and future-oriented readings to be stage-level and the habitual readings to be individual-level. I have excluded sentences with the ability modal (*-na saknu* and the one example in the CONT verb form, leaving 48 sentences. I included the three sentences with verbs which may be intransitive in the given context: *khelnu* ‘to play’ and *polnu* ‘to burn’ (e.g., ‘this jacket burns (my skin)’. If we remove these examples the rate of ergative-marking with habitual readings rises to 83%.

is marked, and this has a pragmatic effect that is interpreted by the speaker. When opposing interpretations are possible, it is due to multiple possible interpretations of the clause as either a categorical proposition or as a prototypically transitive event.

4.1.6 Transitivity and the Event

Six of the properties listed in Hopper and Thompson (1980)'s transitivity prototype relate to the interpretation of the event: Participants, Kinesis, Aspect, Punctuality, Affirmation, and Mode. These were discussed in the Transitivity subsection of the Theory section of this work. In considering potential correlations between ergative marking and aspects of the predicated event, we would expect a correlation between ergativity and highly transitivity with regards to these properties.

For **Participants**, we have seen that transitivity plays a role in the expression of ergative marking, and in fact I have argued that there must be a subject and (implied) object for ergative marking to be possible. Conversely, when verbs like *khelnu* ‘to play’ are construed as intransitive, it becomes possible for a subject to be nominative in the perfective domain. This accords with the Transitivity Hypothesis.

For **Kinesis**, the prediction is that ergative marking should be found on eventive predicates as opposed to stative predicates, and indeed we find ergative marking to be disallowed with stative intransitives.

A brief observation regarding Kinesis: TD and BB at different times expressed the intuition that the ergative sounds better in clauses with verbs which are more “action-oriented.” For example, BB preferred the ergative with the verbs *kinnu* ‘to buy’ and *khānu* ‘to eat’ over verbs which had less tangible effects on the object like *bhannu* ‘to say.’ This could suggest that kinesis plays a gradient role in the expression of ergativity, but my evidence here is anecdotal. I leave the possibility for future research.

For **Aspect**, the Transitivity Hypothesis predicts that if a correlation exists, erga-

Feature	Predicted Correlation
Participants	Yes
Kinesis	-
Aspect	Yes
Punctuality	No
Affirmation	-
Mode	Yes

Figure 4.19: Ergative Correlations and the Event

tive marking will be correlated with telic construals of events or perfective verb morphology. I have argued that ergative marking is obligatory in the perfective domain and alternates with the nominative in the imperfective. This follows the predictions of the Transitivity Hypothesis. Furthermore, I have argued against Li's notion that telic intransitives resist ergative marking while atelic intransitives allow it, and instead argue that this is a generalization about unaccusativity rather than verbal aspect.

For **Punctuality**, the prediction of the Transitivity Hypothesis is that a potential correlation between ergative marking and will favor punctual, instantaneous events over durative ones. Here the apparent correlation between ergative marking and habitual readings of the simple present goes against the prediction of the Transitivity hypothesis, and I argue that this is due to the interpretation of the clause as a categorical proposition.

For **Affirmation** and **Mode**, the prediction of the Transitivity Hypothesis is that ergative marking will favor affirmative events rather than negative events, and will favor realis over irrealis modes. The intuition is that an event is higher on the transitivity scale if it exists in the real world. As to the former, I have found no evidence that ergativity is dispreferred on clauses with negation, but I have not examined the question directly. As to the latter, we can compare three imperfective verb forms which may describe future-oriented events: the Simple Present, the Definite Future, and the Hypothetical Future (the Definite Future being the only one that cannot also describe events which are not future-oriented). The intuitions of consultants and

the corpus sample results indicate that the definite future is much more likely to be marked than the hypothetical, with the simple present in between, which is reasonable if we think of them as representing different degrees of certainty that an event will occur.

4.2 Ergative marking and the Subject

Li (2007) argues that the animate/inanimate distinction constitutes a second ergative split. The general observation that inanimate subjects must take ergative case is found in Verma (1976) and Pokharel (1998) as well, but Verbeke (2011) brings up exceptions and notes that this is not a categorical division but rather a strong tendency. Animacy is a strong predictor of ergative marking in Verbeke and De Cuyper (2015). Pokharel suggests that human referents are somewhat less likely to receive argument marking than non-human referents, raising the possibility that the expression of ergativity falls along the Nominal Hierarchy in a gradient manner. A split by animacy is in line with the predictions of the Nominal Hierarchy, for which ergative marking should be preferred on the right end of the hierarchy. NP-based splits are found in other Indo-Aryan languages like Marathi, but these are typically in the pronominal domain.²²

I find that ergative marking is more common on inanimate subjects, but this represents a tendency rather than a categorical split. I find ergative marking to be least common on first person pronouns. Ergative marking is not gradiently scaled along the entire Nominal hierarchy. Rather, ergative marking is sensitive to two factors in two domains: in the pronominal domain, there is a clear distinction between Speaker (first person pronouns) and all other pronouns. In the domain of common nouns, it is sensitive to animacy, with ergative marking increasingly more common from human-denoting common nouns to nonhuman-denoting common nouns to inanimate nouns

22. According to Li, ergative splits according to animacy are typologically rare, and this is supported by Fauconnier (2011).

(in section 4.3 I argue that this follows a pattern of marking on unexpected subjects based on the frequency of overt subjects of each type).

Hopper and Thompson (1980)'s only feature that relates to the interpretation of the subject is agency. Ergative marking should correlate with high agency, which is mentioned as a potential factor for Nepali by Verma (1976). Properties of Dowty (1991)'s Agent Proto-role include being volitional, sentient, causing the given event, moving in relation to the object, and existing independently. Næss (2004) takes the relevant properties to be Controlling and Unaffected by the event, while Fauconnier (2011) emphasizes the properties of being an Affector and Instigator of the event.

I find that ergative marking in Nepali is **not correlated with properties related to Agency** (volitionality, sentience, controlling, instigating). However, it is correlated with properties related to causation and completion of the event. These are related to the meaning of the *-le* marker as an effector. Secondly, it is associated with kind readings of the subject, definiteness, and strong construals of quantifier. These relate to the ergative as a marker of discourse prominence (as discussed particularly in section 4.5).

4.2.1 The Nominal Hierarchy

For ergative languages with a split conditioned by the semantic nature of the NP, the locus of the split is commonly within the pronominals. For example, Dyirbal and Yidiny are Pama-nyungan languages in which the ergative marker is available on all nouns and pronouns except the first and second pronouns (Dixon 1994: 83). It is somewhat rarer for there to be a distinction between first and second pronouns, and while the general rule seems to be for first person pronouns to be further to the extreme left (e.g., the Nad eb language has ergative marking on second person but not first person), there are some exceptions to this, such as in Ojibwe and Cheyenne. For these Algonquian languages, the second person seems to be on the furthest end of

the hierarchy (Dixon 1994: 90). Within the Indo-Aryan language family, the Marathi language has a split in the ergative morphology. As in Dyirbal, the ergative marker is not available on the first and second person pronominal forms as it is elsewhere in the language.

PK and BB expressed the intuition that the ergative sounds somewhat more natural with third person pronouns than first person pronouns. In one elicited scenario, BB and his friend are at a clothing shop trying on shirts. The shopkeeper asks him if he wants to buy anything, and BB shakes his head no and says,

- (136) *tara rām-le (cahī) yo shirt kin-cha*
but ram-ERG (TOP) this shirt buy-PRES.3.SG
'But Ram will buy this shirt.' [BB]

The presence of the contrastive topic marker *cahī* is similar to the effect of contrastive topic intonation in English ("but RAM will buy this shirt.") The ergative is perfectly natural here and the sentence may sound a little odd without it. In an alternate context, the shopkeeper asks Ram if he wants to buy anything and he shakes his head no, and then the speaker says,

- (137) *tara ma (cahī) yo shirt kin-chu*
but I (TOP) this shirt buy-PRES.1.SG
'But I will buy this shirt.' [BB]

The speaker finds the ergative here to be somewhat stilted, although the only difference in the context is that the subject is 1st person rather than 3rd person. This intuition, as I show below, is supported by the NNSP corpus. As with animacy, person appears to have an affect on the expression of ergative marking in a gradient rather than an absolute way. In contrast to Dyirbal, it is the first person pronouns that seem to pattern differently from the second and third pronouns (rather than the first and second pronouns patterning together).

In the Kathmandu Survey, there was one section (A) in which the question responses were subjects with inanimate referents. For the other trials, I attempted to vary the types of subjects that were used in responses across the other sections. Figure (4.20) and (4.21) show average judgments in transitive clauses with subjects which were overt pronouns (always third person), proper names, other human nouns (e.g. *kākā* ‘uncle’, *pāle* ‘guard’), other animate nouns (*bāgh* ‘tiger’, *carā* ‘bird’) and inanimate nouns (*hāwā* ‘wind’, *asina* ‘hail’). While the ergative is judged higher in each case, there is a clear trend towards dispreference of the nominative and preference for the ergative as animacy decreases. The pronominal domain will be considered in the discussion of corpus results. Note the overall trend in the nominal domain: the judgments for Proper Nouns and Human nouns are almost identical, there is a marked drop in preference for nominative forms on animate nouns, and another marked drop on inanimate nouns.

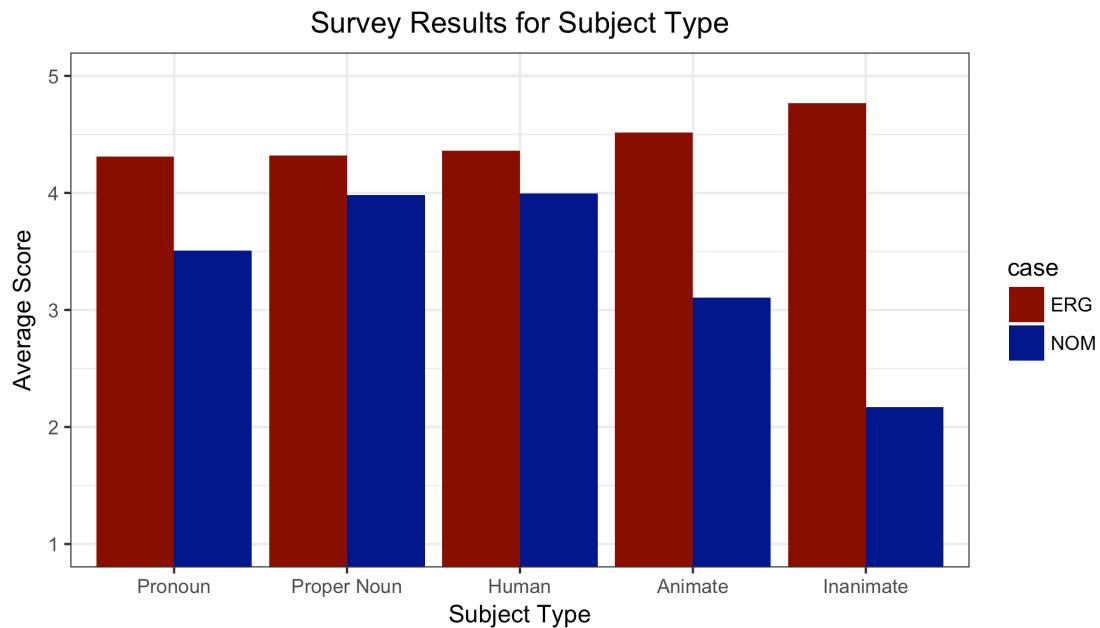


Figure 4.20: Survey Judgments for Subject Type

We find a similar gradient pattern in the NNSP sample, although there is a difference between the pronominal and nominal domain. Of the 119 transitive present

Case	Average Score	Like (4 or 5)	Dislike (1 or 2)
Pronouns			
ERG (n=112)	4.31	83.0%	11.6%
NOM (n=112)	3.51	56.3%	26.8%
Proper Names			
ERG (n=249)	4.32	83.1%	9.6%
NOM (n=249)	3.98	72.7%	10.8%
Human Nouns			
ERG (n=465)	4.36	85.2%	08.8%
NOM (n=464)	3.99	73.1%	12.9%
Animate Nouns			
ERG (n=165)	4.5	87.9%	9.7%
NOM (n=164)	3.1	45.1%	36.6%
Inanimate Nouns			
ERG (n=137)	4.77	94.9%	1.5%
NOM (n=137)	2.16	18.2%	65.0%

Figure 4.21: Survey Results: ERG/NOM by Subject Animacy

tense non-modal main clauses in the sample, 56 have overt subjects. I also expanded the sample to include all transitive imperfective clauses, either main clause or subordinate. I included conditional constructions but excluded all other constructions. This expanded sample represents all (transitive) domains under analysis, and consisted of 508 clauses, of which 119 have overt subjects.

In both cases, we find a particular pattern in the pronominal domain and a separate but parallel pattern in the nominal domain. First person pronouns are nominative the majority of the time, while the ergative form is in the majority for second and third person pronouns. From the larger sample, it appears that ergativity is equally as common on second and third person pronouns. In the nominal domain, ergative marking is more common on inanimate arguments than animate arguments, although there are a few examples of inanimate subjects in the nominative which suggests that ergativity is not required on inanimate subjects.

There are no categorical splits, nor does the distribution of ergative marking gradually follow the entire Nominal Hierarchy. Rather, animacy seems to be a property

negatively correlated with ergative marking in the nominal domain, while first person pronouns are negatively correlated with ergative marking in the pronominal domain. A simpler explanation for this pattern lies in the frequency of usage: first person pronouns are the most frequent and are least commonly ergative-marked. Common nouns are typically animate, and animate nouns are less likely to be marked than inanimate ones.

	1PRO	2PRO	3PRO	ANIMATE	INANIMATE
Transitive Present	37.5% (n=24)	100.0% (n=8)	85.7% (n=7)	55.6% (n=9)	75.0% (n=4)
Transitive Imperfective	38.8% (n=49)	72.2% (n=18)	78.6% (n=14)	48.0% (n=25)	60.0% (n=5)

Figure 4.22: Percentage of Ergative Marking on Overt Transitive Subjects

Inanimate Common Nouns

Because the instrumental and ergative *-le* are homophonous, it can be difficult to ascertain whether a noun with inanimate reference is an ergative or an instrumental. In Nepali, overt nouns with inanimate reference will always trigger third person agreement, and do not have a distinction in gender. The only verbal agreement alternation is the optional agreement in number. Li gives the following example of a sentence with an inanimate subject:

- (138) *dhungā-haru-le jhāl phut-ā-dai-chan*
 stone-PL-ERG window break-CAUS-CONT-SCIMP.3.PL
 ‘The stones are breaking the window.’(Li 2007: 1467)

This sentence is ambiguous between the reading “The stones are breaking the window” and “(They) are breaking the window with stones,” in which the verb agrees with an omitted plural subject. The only way to disambiguate the two readings is to provide a context in which it is clear that only one person is throwing stones (or, less plausibly, that the stones are moving of their own accord).

In Nepali, many verbs like “break” are unaccusative in their root form. The causative suffix *-ā* makes the form transitive. In many cases a more natural way to express the situation would be with an unaccusative sentence like the following:

- (139) *dhunga-haru-le jhāl phut-dai-cha*
 stone-PL-ERG window break-CAUS-CONT-IMPF.3.SG
 ‘The window is breaking because of the stones.’

Fauconnier (2011) provides evidence that this is a common situation crosslinguis-tically. Many languages restrict the S_t to animate referents, and this is one reason it is typologically rare for ergative splits to be based on animacy. Some languages allow inanimate subjects, but the subjects must be “independent instigators.” So, for example, “The gun shot the man” is grammatical but only if the gun went off by itself. As far as I have been able to tell this is not the case in Nepali.

Examining the NNSP sample, Figure (4.23) is a breakdown of arguments with inanimate reference in copular, transitive, and intransitive clauses. Here I have ex-cluded passive constructions because the subjects are more object-like, but I included other clauses and constructions. In the “All Animate” row I have included all argu-ments with inanimate reference whether or not they were overtly represented in the clause, and in “Overt Inanimate” I have included both animate common nouns and demonstratives/pronouns with inanimate reference. The least common position for arguments with inanimate reference is in S_t by far, but it does not appear to be the case that they are completely disallowed. Nor are inanimate transitive subjects more likely to be elided than inanimate transitive objects.

	(S_t)	(O)	(S_i)	
All Inanimate	85	813	220	630
Overt Inanimate	46	443	133	436
Percentage Elided	45.9%	45.5%	39.5%	30.8%

Figure 4.23: Arguments with Inanimate Reference in Different Argument Positions

The following sentences from elicitation sessions are those which I believe to be

relatively natural instances of inanimate subjects. All consultants strongly prefer the ergative. The first example (140) was the only one in which a speaker expressed the intuition that the nominative was possible.

- (140) *yo ukān-(le) ke bhan-cha*
 this expression-(ERG) what say-PRES.3.SG
 ‘What does this expression mean?’ [TD]
- (141) *traffic-ko durghadana-le harek din mānche mār-dai-chan*
 traffic-GEN accident-ERG every day person kill-CONT-PRES.3.SG
 ‘Traffic accidents kill people every day.’ [PK]
- (142) *bādi-le pul-lāi bag-āun-dai-cha*
 flood-ERG bridge-ACC sweep-CAUS-CONT-PRES.3.SG
 ‘The flood is sweeping away the bridge.’ [TD]
- (143) *hāwā-le scarf ud-āun-cha*
 air-ERG scarf fly-CAUS-PRES.3.SG
 ‘The wind takes the scarf.’ (The wind makes-fly the scarf) [TD]
- (144) *das barsha-mā yo rukh-le pahād-ko drishya cop-cha*
 ten year-LOC tree-PL-ERG mountain-GEN view cover-PRES.3.SG
 ‘In ten years, this tree will block the view of the mountains.’ [TD]

In the NNSP sample there are a small number of transitive present tense main clause utterances in the NNSP sample which had overt inanimate subjects. Note the nominative form on the subject in (147) and (148), which are counterexamples to Li’s generalization that inanimate subjects must be ergative-marked. It may be possible to conceive of a construal in which the argument marked by *-le* is an instrumental in (145) and (149), but this is less plausible in the other examples.

- (145) *byāun-cha ke siks.hundred-le ta*
 manage-PRES.3.SG what six.hundred-ERG FOC
 ‘Six hundred (rupees) takes care of it.’ [V001001001; M3]²³

23. The speakers are bargaining using English numbers.

- (146) *garmi-mā katrāijko-le pol-can ta ma-lāi ke*
 summer-LOC corduroy-ERG burn-PRES.3.PL FOC I-ACC what
 ‘In the summer, corduroy (fabric) burns me.’ [V001001004; F4]
- (147) *tara pol-daina dāi yo*
 but burn-PRES.3.SG.NEG older.brother this
 ‘But, brother, this (corduroy) doesn’t burn?’ [V001001004; F4]
- (148) *tapāi-lāi yo kalar dherai myāc gar-cha yo bhandā*
 you.HON-ACC this color much match do-PRES.3.SG this COMP
 ‘This color matches you a lot better than this one.’ [V001001004; M7]
- (149) *kehi kharcai hūdaina sabai kampani-le*
 some expense.EMPH COP.FUT.3.SG.NEG all company-ERG
behor-ihāl-cha
 bear-PUT-PRES.3.SG
 ‘There are no expenses. The company bears everything.’ [V001002005; M7]

4.2.2 Agency and Volitionality

In Hopper and Thompson (1980)’s conception of transitivity, there are two components that relate to the interpretation of the subject. The first of these is **Agency**. The referent of the S_t has a high level of potency; it is highly capable of carrying out the action. The second of these is **Volitionality**, which relates particularly to the intentions of the referent. While Dowty (1991)’s conception of an Agent Proto-Role does not include a component that corresponds to the *potential* to affect the O (but only the extent to which it moves in relation to and affects the O), it does have a component of volitionality. Relating more specifically to OEM, Næss (2004) defines a crucial property of the typical S_t as a “controller,” and Fauconnier (2011) defines it as an “affector.”

As discussed in the subsection on intransitive clauses, the S_i in certain unergative intransitives of Hindi participate in an ergative/nominative alternation that is associated with whether or not the act is volitional. However, this does not seem to be the case in Nepali.

Nor does the ergative appear to correlate with a high degree of agency. In the following minimal pair, no elicitation respondents found the ergative to emphasize that the speaker is engaging in the act of their own accord. In fact, ST expressed the opposite intuition; the ergative is associated with a context in which the speaker was coerced into smoking by someone else. PK's intuition is that the ergative form with a habitual reading is associated with a context in which the speaker has an addiction, which also indicates lower potentiality.

- (150) *ma/maile curot khān-chu*
 I.NOM/I.ERG cigarette eat-PRES.3.SG
 'I am smoking a cigarette.' / 'I smoke cigarettes.' [ST/PK]

Furthermore, I do not find agency nor volitionality to be at play in Nepali dative/ergative modal alternations. Again, if anything the usage of the ergative correlates with *less* agency. The ergative form correlates with a stronger obligation.

- (151) a. *pratyek din rām-le eutā syāu khā-nu par-cha*
 every day ram-ERG single apple eat-INF need-PRES.3.SG
 'Ram *must* eat an apple every day.' [AG]
- b. *pratyek din rām eutā syāu khā-nu par-cha*
 every day ram single apple eat-INF need-PRES.3.SG
 'Ram *should* eat an apple every day.' [AG]

However, agency may be at play in nominative/ergative alternations in other dialects of Nepali, particularly in communities where Nepali is an L2 for many speakers. Ahearn (2001a) analyzed the emerging practice of letter writing among youths in the Magar village of Junigau in Gandaki²⁴, which she related to a shift in the way young people conceive of relationships, identity, and agency. She described an emerging usage of the ergative marker:

24. Check this.

In standard Nepali, ‘le’ is required for the past tense, but in the dialect spoken in Junigau it is often used for other tenses as well. When a person appends this particle to a pronoun such as ‘I’ when speaking of the present or future, the effect is an emphasis on agency. I have translated such statements by italicizing the pronoun in question. (Ahearn 2001a: 41)

Ahearn’s observations indicate that the expression of the ergative *is* correlated with agency at least in the Nepali spoken by turn-of-the-century adolescents of Junigau. Glossing this usage by italicizing the pronoun suggests that the subject is marked with a particular prominence that emphasizes agency.²⁵ It is an open question whether this particular usage is found in other dialects of Nepali which allow ergativity outside of the perfective. I found no evidence of it with any of my elicitation respondents nor in my corpus and survey analysis.

4.2.3 Kind Readings of the Subject

Following the influential analysis of English bare plurals in Carlson (1977b), certain noun phrases may be interpreted as the proper name of a kind, i.e., a reference to a type rather than to a specific individual entity. A kind reading is particularly associated with generic statements like “cats drink milk.” Such readings require a bare plural subject in English, while in Nepali the plural marker is possible in such statements but not required. Butt and Poudel (2007) note a correspondence between -*le* and kind readings of the referent, providing the following example:

- (152) *rāute-le jangel-ko kandamul khān-chan*
Raute-ERG forest-GEN wild.edibles eat-PRES.3.PL

25. Considering that there is potential influence from Magar, it would be useful to know whether the Magar language participates in nominative/ergative alternations correlated with an emphasis on agency, or whether this is entirely an innovation of young Nepali speakers. Madhav Pokharel (p.c.) particularly associates the usage of the ergative in nonperfective clauses with the dialects of the *Magarrat*, the historical territory of the Magars.

‘The Rautes (ethnic group) eat the wild edibles of the forest.’ (Butt and Poudel 2007: 8)

All elicitation consultants do generally prefer the ergative on kind readings, although they do not appear to be absolutely necessary for a generic interpretation:

- (153) a. *kukur-haru-le āap khān-chan*
 dog-PL-ERG mango eat-PRES.3.PL
 ‘Dogs eat mangoes.’ [AG]
- b. *sabai birālo-haru-le māsu khā-e*
 all cat-PL-ERG meat eat-PERF.3.PL
 ‘All cats eat meat.’ [ST]
- c. *atānka-le harek din mānche mār-dai-chan*
 terrorist-ERG every day person kill-CONT-PRES.3.SG
 ‘Terrorists are killing people every day.’ [PK]
- d. *cyā tyo ta paket na-bha-eko ta keṭi-haru-le*
 ugh those FOC pocket NEG-COP-PERF FOC girl-PL-ERG
 lāun-chan kere
 wear-PRES.3.PL EVID
 ‘Ugh, those (jackets) without pockets, women wear those.’ [V001001004;
 F4]
- e. *jhusilkirā-le tyo hāt-mā chun-cha*
 caterpillar-ERG there hand-LOC sting-PRES.3.SG
 ‘Caterpillars sting you there on the hand.’ [V001002005; M7]
- f. *tala ā-yo, ani tyāhā-bāṭa khaire, utā-ko*
 below come-PERF.3.SG, and there-ABL white.person, opposite.side-GEN
 khaire-haru, sab katā.katā bhāg-isak-yo,
 white.person-PL, all what.direction.RED escape-COMPL-PERF.3.SG,
 tyo kukharā-ko baccā cil-le cop-cha ni ?
 that chicken-GEN child eagle-ERG charge-PRES.3.SG and ?
 tyastai bhayo
 like.that.EMPH COP.PERF.3.SG
 ‘Then (the rhino) came down below, and from there the tourists, the
 tourists on the opposite bank, they all fled in different directions. **You**

know how eagles swoop down on chicks (and they scatter)? It was like that.’ [V001002005; M26]

Of the 20 present tense non-modal transitive clauses with habitual readings, the three examples given above are the ones that most clearly represent generic statements about kinds, although there is some ambiguity in the interpretation of some subject noun phrases as kinds or bare existentials. The ergative is present in all of these relatively straightforward cases.

In the Kathmandu survey, Section G contained question responses containing subjects with kind readings: results generally support this tendency:

(G1)

Q: *Pradhān māntri-ko kārya ke ho?* What is the job of a/the prime minister?

A: *Pradhān māntri-(le) desh calāunu huncha.* Prime ministers run the country.

(G2)

Q: *maile kina padnu parcha?* Why do I have to study?

A: *Raamrā bidhyārthi-(le) pratyek din pad-chan.* Good students study every day.

(G3)

Q: *Gaiḍā-haru din bhara ke garchan?* What do rhinos do all day?

A: *Uni-haru-(le) ghās khānchan.* They eat grass.

(G4)

Q: *Pilot-haru kati ucaī-mā hawāijahāj udāū chan?* How high do pilots fly their planes?

A: *Pilot-haru-(le) dherai jasto bāhra hajār meter-mā hawāijahāj udāū chan.* Pilots typically fly their planes at twelve thousand meters.

(G5)

Q: *Bagh-le kati belā janāwar-lāi sikār garlā?* What time will tigers hunt animals?

A: *Bagh-(le) rāti-mā janāwar-lāi sikā garlā.* Tigers will usually hunt animals at night.

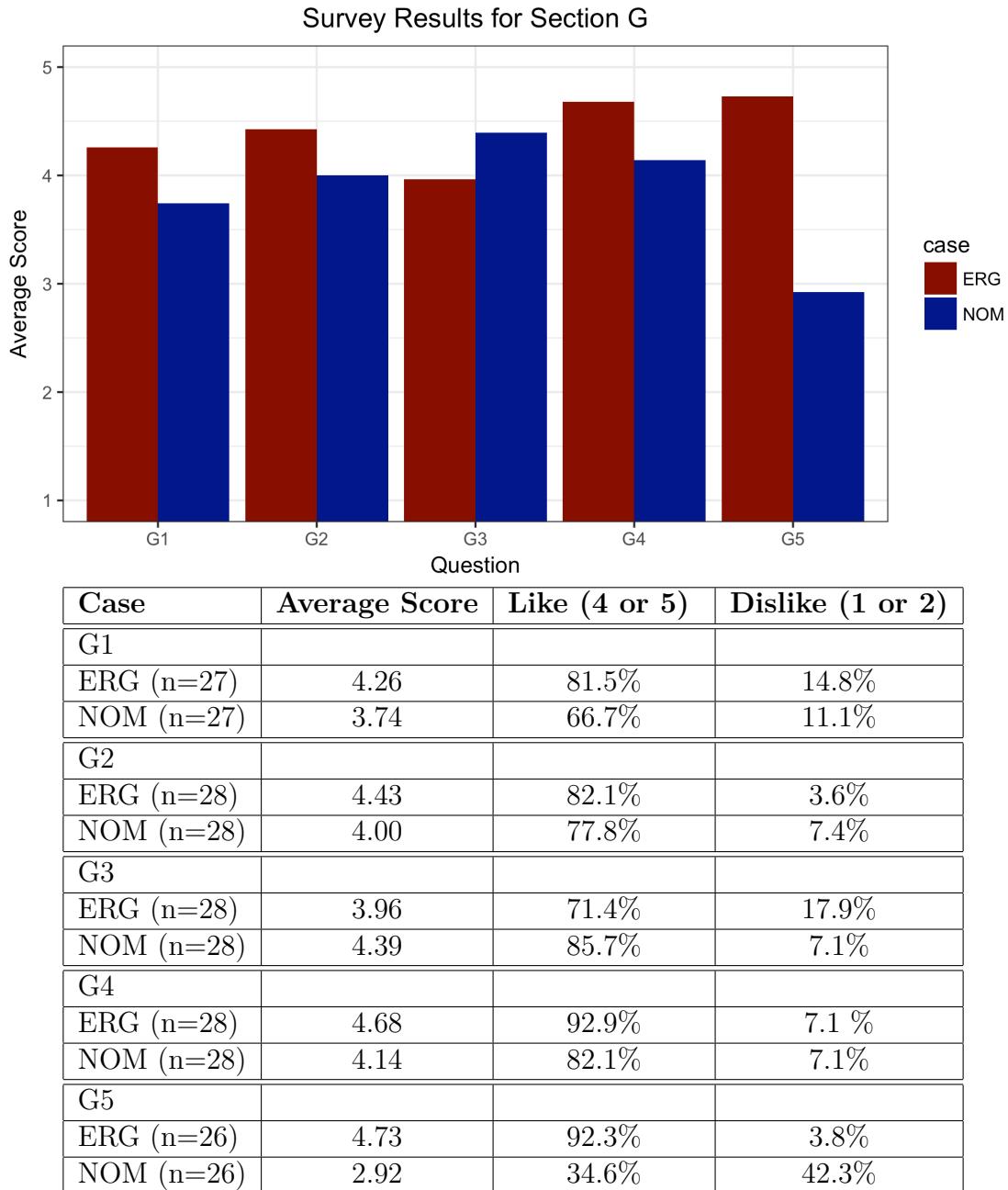


Figure 4.24: Survey Results for Questions with Kind Readings (Section G)

The responses are in (Figure 4.24). The average response is generally positive for

both versions, but the ergative is somewhat preferred in each case, as is generally the case throughout the data set. The exception of (G3), which is an outlier in two ways: the question is given with the transitive subject in the nominative form, and it is the only response in which the subject is realized as a pronominal rather than a noun. The nominative is particularly dispreferred in (G5), which is likely due to the verb form being in the hypothetical future.

4.2.4 Sets and Strong Construals of Quantifiers

Some elicitation respondents (particularly PK) expressed the intuition that usage of the ergative has the effect of “pointing out” the marked entity from a group of possible entities. Thus the following sentences might be a natural response to a question about which of a group of employees is a smoker:

- (154) *#hari/hari-le curot khān-cha*
 Hari/Hari-ERG cigarette eat-PRES.3.SG
 ‘Hari smokes cigarettes.’ [PK]

For PK the nominative form sounds odd here, although other speakers accepted it in this context. This general preference for the ergative is stronger when the context set is explicitly stated in the sentence itself:

- (155) *yahā madhye #hari/hari-le cahī curot khān-cha*
 here among hari/hari-ERG TOP cigarette eat-PRES.3.SG
 ‘Among (those) here, Hari smokes cigarettes.’ [ST]
- (156) *cidiya.kanā-mā bhay-ekā saya watā cituwa-madhye saicālis watā*
 zoo-LOC be-PERF.OBL hundred CT leopard-among forty.seven CT
#cituwa-haru/cituwa-haru-le aushadi khān-dai-chan
 leopard-PL/leopard-PL-ERG medicine eat-PROG-PRES.3.PL
 ‘Of the zoo’s one hundred panthers, forty-seven panthers are taking medicine.’
 [BB]
- (157) *das watā-mā āṭh watā #birālo/birālo-le mācā khān-chan*
 ten CT-LOC eight CT cat/cat-ERG fish eat-PRES.3.PL

‘Eight cats out of the ten cats (that I own) eat fish.’ [BA]

Similarly, the ergative is preferred on strong construals of quantifiers. In (158) the ambiguous quantifier *kohi* ‘some’ has a weak interpretation, whereas in (159) it has a strong interpretation; we are referring specifically to a group of Nepali students who go to the university.²⁶

- (158) ***kohi.kohi/#kohi.kohi-le*** *bakunda khel-dai-chan*
some.RED/some.RED.PL-ERG football play-CONT-PRES.3.PL
‘Some (Nepalis) are playing football.’ [RM]
- (159) **#*kohi.kohi/kohi.kohi-le*** *bakunda khel-dai-chan*
some.RED.PL/some.RED.PL-ERG football play-CONT-PRES.3.PL
‘Some (of the Nepalis) are playing football.’ [RM]

There is an alternation between two quantifiers *dherai* and *dheraijaso*, which would both be translated into English as ‘many’ or ‘most.’ While *dherai* can be construed as weak or strong, *dheraijaso* is unambiguously strong. All elicitation respondents had a very strong intuition that the ergative is required here.²⁷

- (160) ****dheraijaso/dheraijaso-le*** *bakunda khel-dai-chan*
many/many-ERG football play-CONT-PRES.3.PL
‘Many (of the Nepalis) are playing football.’ [RM]

In each of these cases, the ergative form correlates with a definite interpretation of the subject in which the speaker is drawing from a set which is explicitly stated or presupposed in the discourse. For a more detailed explanation of this phenomenon, see Lindemann (2016).

26. Quantifier reduplication marks plurality, hence *kohi* ‘some (singular)’ *kohi.kohi* ‘some (plural)’.

27. One explanation for this is that the quantifiers are unselective (Lewis 1975), and the ergative is marking the subject to disambiguate it from what would otherwise be an interpretation where *dheraijaso* quantifies over the event rather than the subject: “(Nepalis) play football some of the time.” In any case, it is striking that the speaker judgments here were more categorical than with any of the other ergative/nominative alternations with quantifiers.

4.2.5 Summary

There are no categorical splits based upon the semantics of the subject noun phrase. Rather, ergative marking is gradually less common on animate subjects, and it is least common on first person pronouns. Rather than a smooth transition along the Nominal hierarchy, ergative marking appears to be sensitive to different properties in the pronominal domain than it does in the nominal domain. In section 4.3 I argue that this is ultimately derived from the tendency for ergative-marking to be found on unexpected (less frequently-occurring) subject types, which supports a markedness-based analysis.

We do not find a correlation with agency or volitionality as in Hindi. Nor do we find correlations with the subject being a controller or an instigator. This suggests a division between those aspects of transitive subjecthood that relate to the instigation of an event and those that relate to its causation and completion.

Feature	Predicted Correlation
Dowty (1991)	
Volitionality	No
Sentience	No
Causation	Yes
Movement	Yes
Næss (2004)	
Controlling	No
Unaffected	Yes
Fauconnier (2011)	
Affector	No
Instigator	Yes

Figure 4.25: Ergative Correlations and the Subject

Finally, ergativity is correlated with kind readings of the subject, definiteness, and strongly-presupposed sets. These properties arise from the interpretation of the ergative-marked element as the logical subject of a categorical proposition.

4.3 Argument Realization and Case Frequency in the NNSP Sample

In this subsection I discuss the frequency of different types of arguments types in S_i , O, and S_t position, as well as the properties of datives and instrumentals. The results support the general conclusion that ergative marking is more common on subjects of types which are rare (those with inanimate reference) and less common subject types which are the most common (overt first person pronouns).

In the NNSP sample I analyzed argument realization by compiling the data on participants in each utterance (excluding passive, modal, nominal subordinate, and conditional constructions). For the core arguments in transitive, intransitive, and copular clauses, I recorded whether the argument was overt or elided, and whether or not it was case-marked. Additionally, I recorded dative and instrumental arguments when they were present. This data set consists of 2,823 arguments.

I recorded 829 arguments in the S_t position of transitive clauses; the results are in Figure (4.26). In this position, 65.6% of the arguments refer to speech act participants (SAP). However, second person pronouns are more likely to be elided. First person pronouns are by far the most frequent type of overt subject form in S_t position, consisting of 46.6% of all overt arguments. Inanimate subjects in S_t are quite rare: only 6.3% of subjects have inanimate reference. Arguments are also more likely to be elided in this position: the argument is omitted 72.0% of the time.²⁸ Of the 232 overtly realized transitive subjects, the subjects were in the ergative case 65.9% of the time and (unmarked) nominative the rest of the time.

28. This accords with the morphological pattern described by Du Bois (1987) as the **Avoid Lexical A Constraint**. Arguments in this position tend to be omitted. In fact, overtly-realized lexical items (excluding pronouns) comprise only 6.2% of the total in this sample, a total remarkably close to the 6% reported by Du Bois in his Sacapultec sample. Du Bois argues that this derives from a pragmatic constraint: the **Given A Constraint** describes the tendency for arguments in this position to be given rather than new information.

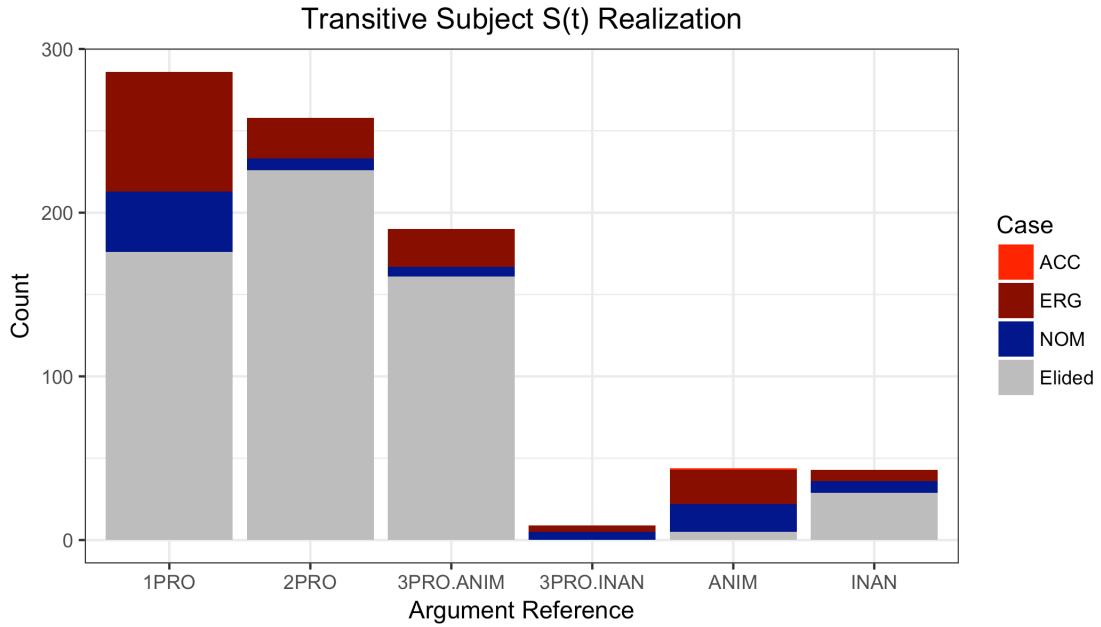


Figure 4.26: Transitive Subject Types in the NNSP Sample

The results for the 808 O arguments are given in Figure (4.27). The first graph displays every object type, while the second is a breakdown specifically of those few arguments that have animate reference. The main generalization is that O arguments tend to have inanimate reference (88.8%). Only 2.1% of the arguments are expressed as overt pronouns. Arguments are omitted less frequently: 44.1% of the objects were elided. There is a strong tendency for accusative marking on pronouns (88.2%) and for inanimate arguments to be nominative (95.8%), but these are not categorical distinctions. It is apparently possible for an inanimate subject to have accusative marking, and it is possible for a pronoun to go without.

The results for the 450 S_i arguments in intransitive clauses are given in Figure (4.28). The overall generalization is that there is no type of argument that is prolific in this position. Arguments with animate reference make up 55.1% of the total. The arguments are elided 52.9% of the time. The ergative, as expected, is overall very uncommon, being present on just 4.7% of the overtly realized arguments. The ergative is not found on any S_i arguments with inanimate reference.

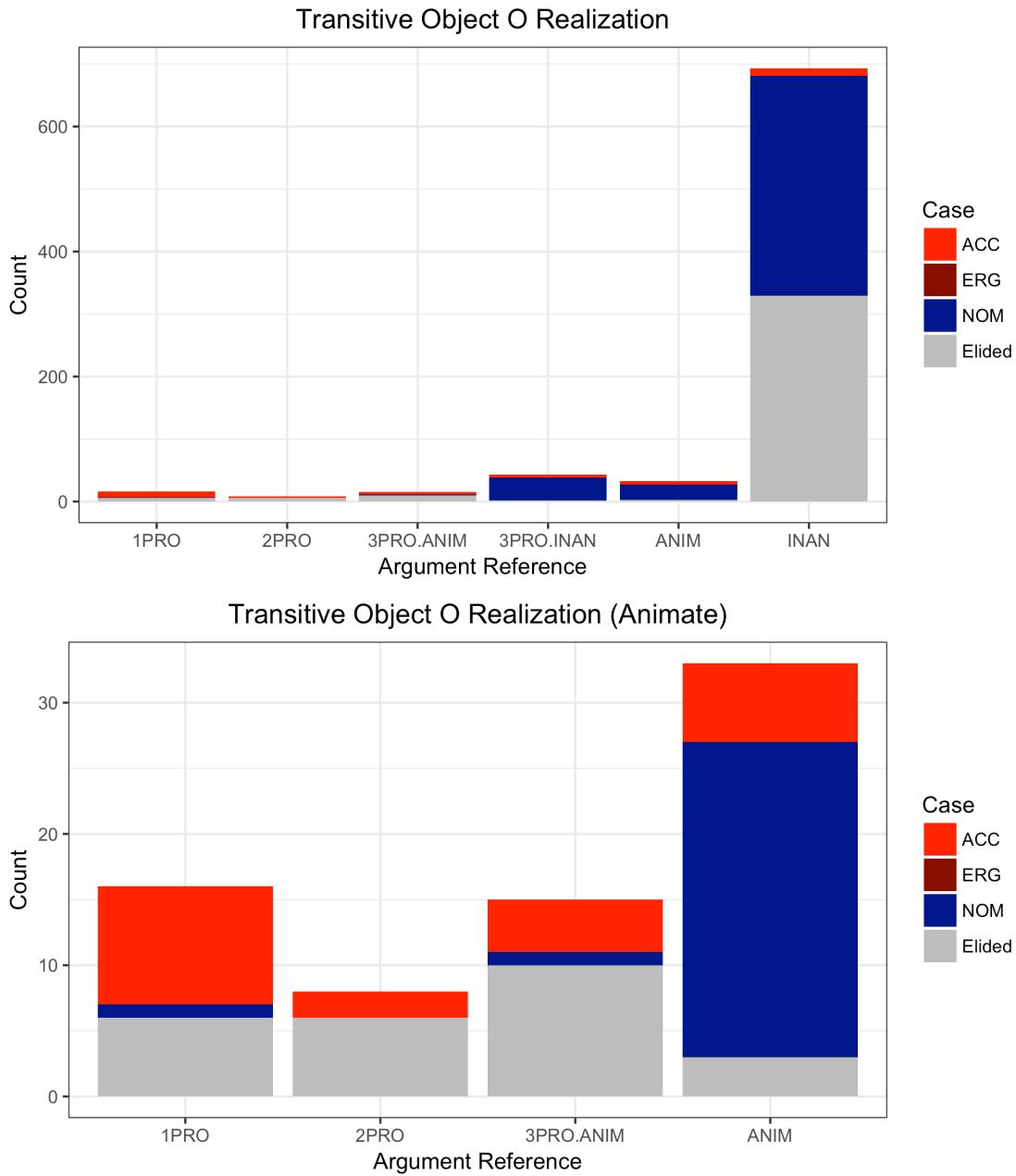


Figure 4.27: Transitive Object Types in the NNSP Sample

The results for the 665 copular subjects are given in Figure (4.29). As with the O position, the majority of arguments in this domain have inanimate reference (89.9%). Also like the O position, relatively few arguments are omitted (35.4%). There are no examples of ergative or accusative marking in the copular domain.

There are relatively few examples of oblique arguments marked in the corpus. The

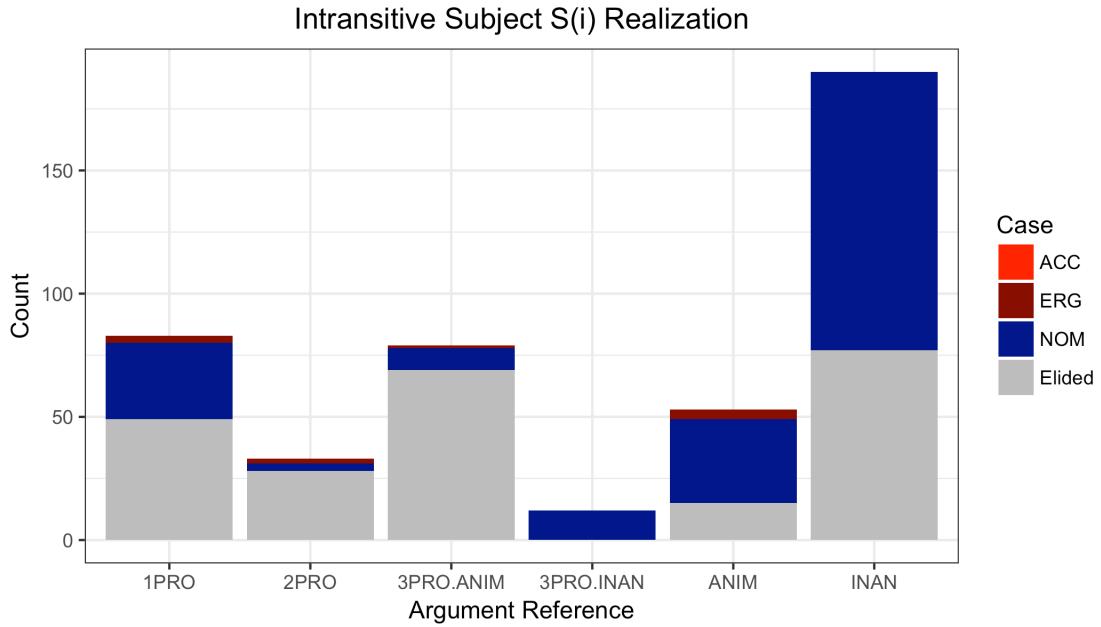


Figure 4.28: Intransitive Subject Types in the NNSP Sample

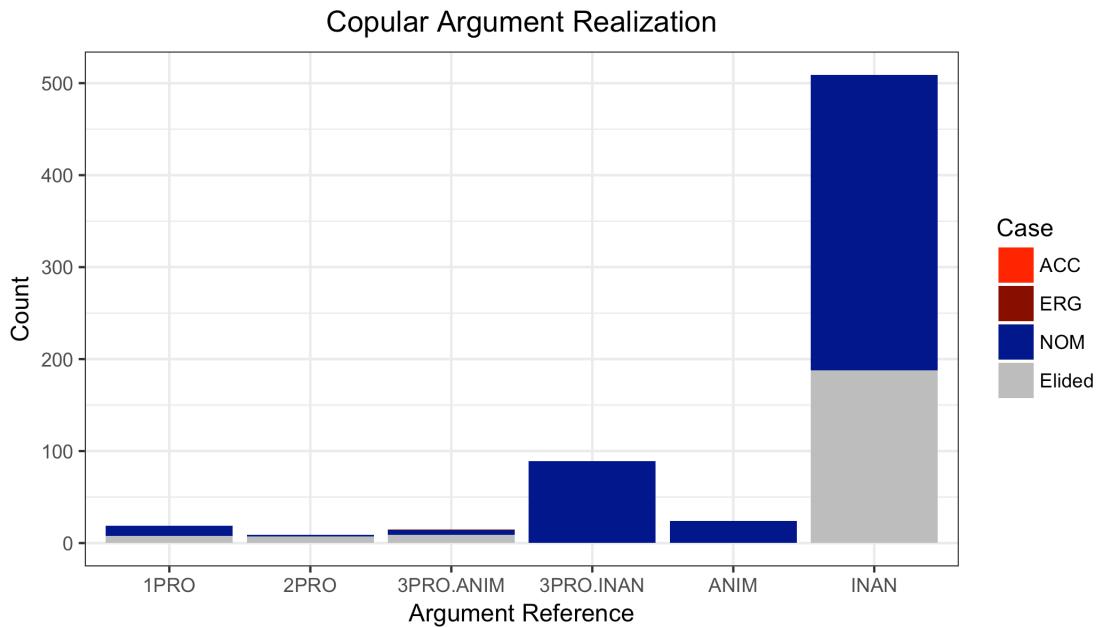


Figure 4.29: Copular Subject Types in the NNSP Sample

results are given in Figure (4.30). Of the 64 examples of dative marking (*-lāi* in a ditransitive or as an experiencer with copular or intransitive clauses), a surprisingly high 14.1% of the arguments have inanimate reference. Another interesting finding

is infrequent nominative unmarked forms on 4.7% of the indirect arguments. Instrumental arguments, marked with *-le*, are quite rare. Only 7 examples were found in the sample, and all are on arguments with inanimate reference.

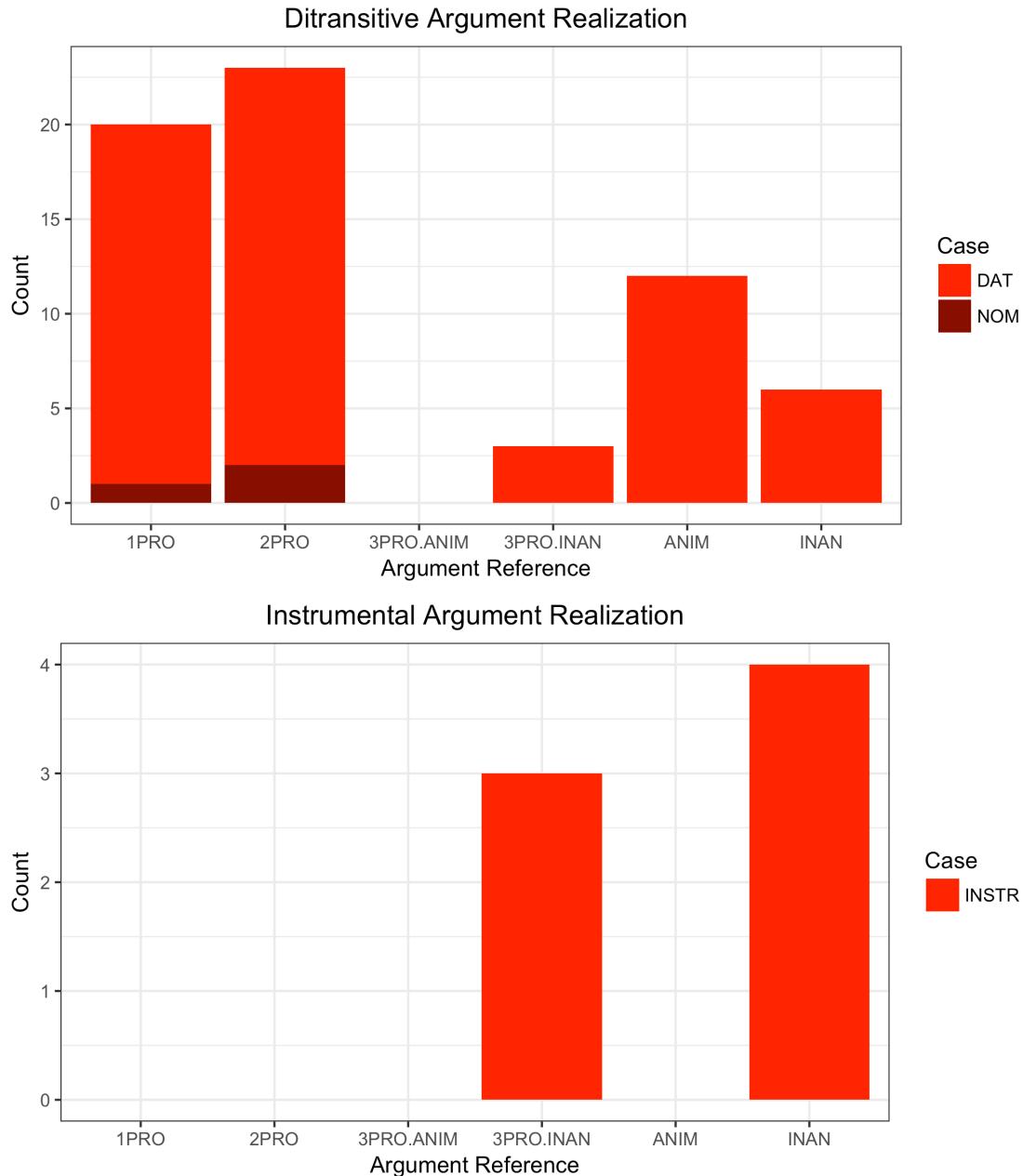


Figure 4.30: Oblique Argument Types in the NNSP Sample

Figure (4.31) tabulates the expressed arguments in the 809 transitive clauses which do not have indirect objects. Both arguments are overt 14.5% of the time, while both

are omitted about a third of the time.²⁹ This tendency toward the omission of one or both arguments is less strong among the 26 ditransitive constructions, in which all three arguments are overtly realized 42.3% of the time.

Both realized	O Omitted	S_t Omitted	Both Omitted
14.5% (n=117)	39.1% (n=316)	12.6% (n=102)	33.9% (n=274)

Figure 4.31: Number of Expressed Arguments in Transitive Clauses

To summarize:

- (1) The typical argument in the S_t position is a speech act participant. It is frequently omitted, which correlates with a tendency to be given and topical. Ergative marking correlates positively with the least frequent overt subject type (those with less animate reference) and negatively with the most frequent overt subject type (first person pronouns).
- (2) The typical argument in the O position is inanimate, and is less frequently omitted, thus more likely to be new information. The subject of copular clauses follows the same pattern.
- (3) The arguments in S_i may be animate or inanimate, and are less likely to be omitted.
- (4) In most transitive clauses either or both the S_t and O are omitted, leading to potential ambiguity, particularly if the S_t is unexpectedly inanimate or the O is unexpectedly animate.

The generalizations about ergative marking and the subject in section 4.2 can be partially explained by the frequency of different argument types in S_t position. In

29. If we consider only those which are lexically realized (excluding pronouns), then only 3.5% of the clauses have two overtly realized arguments. This is comparable to Du Bois' total of 1% in the Sacapultec sample, formulated as the **One Lexical Argument Constraint**, which Du Bois correlates with a pragmatic constraint on the number of new mentions typically allowed in an utterance (the **One New Mention Constraint**).

4.2 I noted that the ergative correlates positively with animacy and negatively with first person pronouns, and these observation correspond straightforwardly with the frequency of overt argument types. Ergative marking is more common on the least common types of subjects.

However, it is not simply the case that ergativity marks a deviation from the prototypical subject. Ergativity also correlates with definiteness (and topicality, as I show in 4.5), which is a property of a typical subject. In chapter 6, I argue that the correlations with animacy come from emphasizing the effector role, while the correlation with topicality and definiteness arises from general markedness principles.

4.4 Ergative marking and the Object

The disambiguation hypothesis suggests that ergative marking will be more likely either when the object is not overtly realized or when it is unmarked with the accusative case marker. Related to this point is the question of whether the systems of accusative marking and ergative marking are completely separate or are interrelated, as discussed by Dixon (1994). For example, ergative marking may be more likely when accusative objects are unmarked in the accusative case. Or ergative marking may be required when the object is more animate than the subject.

From the literature on object prototypes/proto-roles we may expect the ergative to correlate with a highly affected object. Those aspects of transitivity that have to do with the object include the following:

- (1) Affectedness: the object is highly affected by the event in question. It may come into existence as a result of the event, or it may be completely affected rather than partially affected.
- (2) Individuation: the object is human or animate, concrete, singular, and referential. Hopper and Thompson (1980)'s conception of the prototypical object is at

odds with that of Aissen (2003) for whom the prototypical object is indefinite and inanimate.

4.4.1 Ergative marking and the Realization of the Object

In a Nepali transitive clause, the direct object may be overt or elided. If it is overt, it may be a bare nominative form or marked with the accusative *-lāi* postposition. This can be conceptualized as three “settings” of argument realization: elided, unmarked nominative, and marked accusative. The same three settings are available to the transitive subject: as elided, unmarked nominative, or marked ergative.³⁰

In investigating a possible global interaction between subject and object realization, it is important to keep in mind that there will always be local motivations for a particular realization. The semantic properties of the argument referent, along with the exigencies of discourse structure (as discussed in the next subsection), have an effect on whether the argument is overt and/or case-marked. A weak correlation between subject and object realization is potentially an epiphenomenon of these local concerns, along with the overall tendency for arguments of certain types to be placed in certain positions (as discussed in the previous subsection).

On the other hand, Dixon (1994) describes ergative split systems in which a global interaction is grammatically operationalized. An example of this would be a system in which ergative marking is possible *only* when the object is elided. This is also the case for inverse case-marking languages, which can be thought of as languages with ergative splits conditioned by a global interaction of the semantic properties of S_t and O: inverse marking occurs when the object is higher along the Nominal hierarchy than the subject.

It is clear that Nepali is not such a language, despite some evidence of the type

30. Realization of common nouns with a pronominal form is another possibility which I'm leaving aside for now.

that Abadie presents that suggests that ergative marking is required when the object is omitted and it would otherwise lead to ambiguity. In these cases, the respondents found the nominative form to correspond more closely to an interpretation in which the overt argument is the object. This is true for (161c) even though that reading does not make sense in the given context:

- (161) a. Context: I just dropped a bit of samosa on the road. What will happen to it?
- b. *carā-le khān-cha*
bird-ERG eat-PRES.3.SG
'A bird will eat (it).' [TD]
- c. *carā khān-cha*
bird eat-PRES.3.SG
?'(Something) eats bird.' [TD]
- (162) a. Context: What will happen if a leopard decides to attack a tiger?
- b. *bāgh-le mār-idin-cha*
tiger-ERG kill-BEN-PRES.3.SG
'The tiger will kill (it).' [BA]
- c. *bāgh mār-idin-cha*
tiger kill-BEN-PRES.3.SG
'(It) will kill the tiger.' [BA]

In the Kathmandu Survey I included five questions of this type with omitted objects. Figure (4.32) compares (transitive imperfective) clauses with omitted objects to those with overt objects. It is not so much that there is a preference for the ergative when the object is elided, but rather that the nominative form is somewhat dispreferred for a significant proportion of speakers. Roughly half of the respondents (a conspicuous 47.0%) dislike the nominative form when the object is elided.

In the NNSP I examined all transitive imperfective clauses (excluding conditional, passive, and modal constructions). I examined whether or not the object was elided

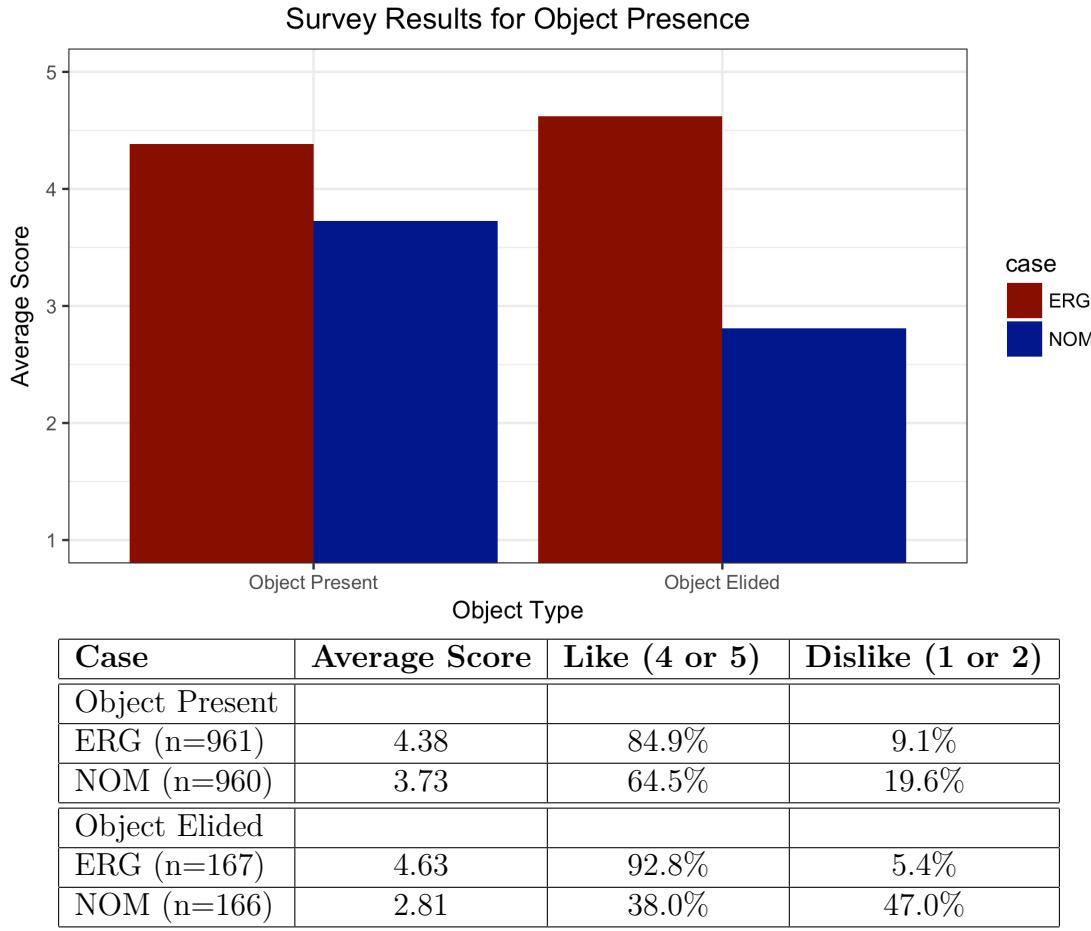


Figure 4.32: Survey Judgements for Transitive Clauses with Overt/Elided Objects

and whether or not the object was case-marked by *-lāi*. The results, presented in Figure (4.33) are comparable to the survey judgments. We find a (slight) increase in the percentage of ergative marking when the object is elided. However, we also find an increase in ergative marking when the object is case-marked accusative rather than bare nominative. These differences may or may not be significant, but even if they are, it is clear that we are not dealing with a global interaction in which, for example, ergative marking is completely required if the object is elided. The percentage of ergative marking ranged from 46-64%. Accusative marking is heavily associated with pronouns and definite animates, whereas elision is associated with information structural concerns like givenness. It is likely that any patterning we see here is an epiphenomenon of these local factors.

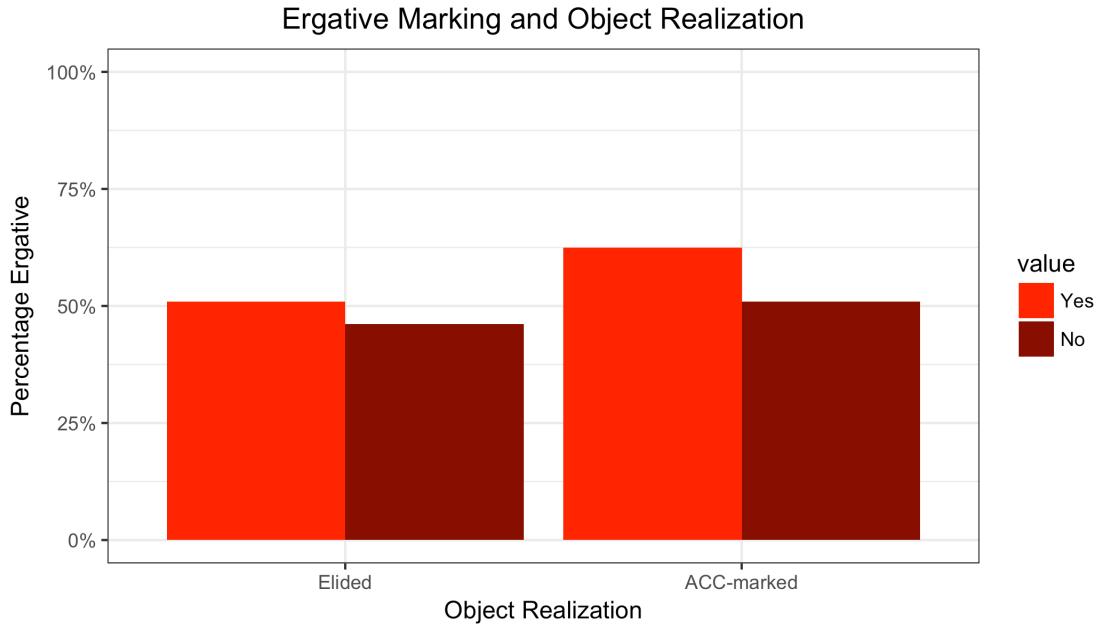


Figure 4.33: Corpus Results for Object Realization

Finally, we might consider ergative marking to be required particularly when the O is higher along the Nominal Hierarchy than the S_t, such as when the subject is an animate/inanimate common noun and the object is a local pronoun. These constructions are quite rare (I found six examples among the transitive imperfectives), but it appears that even this is more of a tendency than a requirement. The first two are examples we have already seen, while the third is a psychological predicate in which the subject, a feeling, is never case-marked ergative.

- (163) a. *garmi-mā katraījko-le pol-can ta ma-lāi ke*
 summer-LOC corduroy-ERG burn-PRES.3.PL FOC I-ACC what
 ‘In the summer, corduroy (fabric) burns me.’ [V001001004; F4]
- b. *tapāi-lāi yo kalar dherai myāc gar-cha yo bhandā*
 you.HON-ACC this color much match do-PRES.3.SG this COMP
 ‘This color matches you a lot better than this one.’ [V001001004; M7]
- c. *ani janawār-ko dar lāg-dainathyo ?*
 and animal-GEN fear-PST.HAB.3.SG.NEG ?
 ‘And weren’t you afraid of animals?’ [V001002005; M7]

4.4.2 Affectedness

Another property of high transitivity that we might expect to correlate with the usage of the ergative marker is the extent to which the object is affected by the action. Fauconnier (2011) considers the two primary cluster properties of the transitive subject to be the ability to affect the object and instigate the action. We might expect, then, that the ergative would be associated with a construal in which the object is completely affected. However, this does not appear to be the case. In the following example, the chair might be tipping slightly or tipping over completely, and the nominative/ergative alternation exists either way:

- (164) *mai-(le) kursi-lāi lāta hān-chu*
I.OBL-(ERG) chair-ACC tip hit-PRES.1.SG
'I tipped the chair.' [BB]

The difficulty here is in entangling properties of the event, for which a perfective aspect would imply that the object was completely affected and an imperfective aspect often implies that it has not been completely affected. We might expect an alternation in something like "She was kicking the dog" and "She was kicking at the dog," in which the second sentence in English implies that she did not actually make contact with her foot. However, this is not expressed by an ergative/nominative alternation in Nepali.

- (165) a. ***us-le*** *kukur-lāi lāthe hān-dai-thyo*
PRO.OBL-ERG dog-ACC kick strike-CONT-PST.HAB.3.SG
'(S)he was kicking the dog.' [BB]
- b. ***us-le*** *kukur-lāi lāthe hān-dai-thyo* *tara kukur-lāi*
PRO.OBL-ERG dog-ACC kick strike-CONT-PST.HAB.3.SG but dog-ACC
lāg-ena
contact-PRES.3.SG.NEG
'(S)he was kicking the dog but missed.' [BB]

Or consider the action of plucking out a splinter, and an alternation between 'plucking

out a splinter' and 'plucking at a splinter.' This too is not expressible by a simple ergative alternation, but rather requires a larger clause to get a similar meaning across:

- (166) a. ***us-le*** *hāt-bāṭa kāḍā jhik-dai-chā*
 3.PRO.OBL-ERG hand-ABL splinter pluck-CONT-PRES.3.SG
 '(S)he is plucking out the splinter in (her) hand.' [BB]³¹
- b. ***us-le*** *hāt-bāṭa kāḍā jhik-ne khoj-dāi-chā*
 3.PRO.OBL-ERG splinter pluck-INF try-CONT-PRES.3.SG
 '(S)he is trying to pluck out the splinter in (her) hand.' [BB]

4.4.3 Connectedness to the Object

I have observed one effect of nominative/alternation as it relates to the interpretation of the object, which I loosely refer to as "object connectedness." Multiple elicitation respondents (UK, BB, and SB) expressed a particular intuition about the minimal pair below:

- (167) a. ***ma ghar*** *ban-āu-dai-chu*
 I house build-CAUS-PROG-PRES.1.SG
 'I am building a house.' [UK]
- b. ***maile ghar*** *ban-āu-dai-chu*
 I.ERG house build-CAUS-PROG-PRES.1.SG
 'I am building a house (for myself).' [UK]

The ergative form emphasizes that the object has a closer relation to the subject. In this case, respondents expressed the intuition I am building the house for myself. I might or might not be a construction worker by trade, and in both cases the event is ongoing, but with (167b) the implication is that I am building a house for myself.

A similar example came up in discussing the NNSP interviews with BB. In this example, a customer is looking at t-shirts in a shop and tells the shopkeeper which

31. Double check this.

shirt he wants to buy.

- (168) a. *ma yo shirt lān-chu*
I this shirt take-PRES.1.SG
'I will take this shirt.' [BB]
- b. *yo shirt lān-chu mai-le*
this shirt take-PRES.1.SG I.OBL-(ERG)
'I will take this shirt.' [BB]

In the first context, he may be buying the shirt for a friend of his, but in the second it is clear that the shirt is for him. The subject is postposed, which is somewhat more common for ergative-marked clauses, as I discuss in section 4.5.1.

4.4.4 Summary

Associations between properties of the object and ergative marking on the subject are less clearly evident. This should be expected, because the relationship between the S_t is indirect and mediated by the event. There are no categorical observations, but there is some evidence that ergative marking is more common with elided objects, thus providing support for the disambiguation hypothesis. It may also be more common on accusative-marked objects, which are more likely to be animate or human.

Feature	Predicted Correlation
Elision	(Possibly)
Affectedness	No
Individuation	Yes

Figure 4.34: Ergative Correlations and the Object

I do not find the ergative participating in alternations in which the object is partially or completely affected by the event. Thus, I describe *-le* as an Effector rather than an Affectee: it must be involved in enacting the event, but it is not necessarily implicated in its effect on the object. From section 4.1.5 on individual-level predication we saw that one interpretation of the usage of the ergative was that

the object was more individuated (hence *ma kām garchu* “I’m doing work” versus *maile kām garchu* “I’m doing a job”).

4.5 Ergativity and the Discourse

Ergative-marking is correlated with topicality and topic positions as well as categorical propositions, which are ultimately expressions of discourse prominence.

4.5.1 Topic, Focus, and Word Order

Abadie (1974) and Verbeke (2011) both reject the notion that the *le*-marked element will necessarily be in a focused position, but it may represent a tendency. The subject of a transitive clause will typically be a topic, and topics are often elided because they represent given information rather than new information. This is why the *S_t* position has the greatest percentage of elided arguments (72%). In other words, topics are not typically prominent. Marking a subject imparts some level of prominence (leading to the claim of Grierson (1904a) and others that the ergative imparts emphasis), and thus we might expect it to correlate with a contrastive topic.

There are not obvious prosodic contours associated with focus or contrastive topic in Nepali, but rather discourse particles that impart focus (*ta*) and topicalization (*cāhī*). Furthermore, different positions in the clause are associated with different information structural categories:

- (169) [Topic] [Completive] [Focus] V [Background]

Elements in the **Topic** position and **Background** position contain old information. The difference is that the topic element is prominent and background element is not. **Focus** and **Completive** both provide new information, but focus is prominent and background is not.

The elicitation respondents generally allowed both the nominative and ergative whether the subject is focused or not, which suggests that there is no inherent and categorical connection between discourse structure and the expression of the ergative. There is perhaps a general tendency for the ergative marked element to contain given information, which means that they tend to be avoided in strongly focal positions.

In the Kathmandu Survey I compared some questions in which the focus is on the subject versus on the object. The results should be considered with caution because the absence of discourse particles and the fact that every argument is overt leads to somewhat awkward phrasing.³² Definitively testing a correlation between various kinds of focus and the expression of ergativity will require a survey specifically dedicated to examining these factors. From the results in Figure (4.35) results we can conjecture that nominative and ergative is available for most speakers, although the nominative is dispreferred by a majority of speakers when the focus is on the object rather than the subject. Ergative marking does not appear to be particularly correlated with focus.

(B1)

Q: *Ke ketā-haru-le nayā cal.citra herlān?* Will the boys watch the new film?

A: *Keti-haru-(le) herdaichan tara ketā-haru-le herdainan.* The GIRLS are watching it but the boys aren't.

(B2)

Q: *Ke usle sabai masalā ta hāmilāi cāhine?* Will s(h)e buy all the spices we need?³³

A: *U(Usle) sukumel kincha tara lwān kindaina.* (S)he will buy CARDAMOM but not cloves.

32. As pointed out to me by SB in particular.

33. I believe a more natural phrasing would be *Ke usle sabai hāmilāi cāhine masalā kincha.*

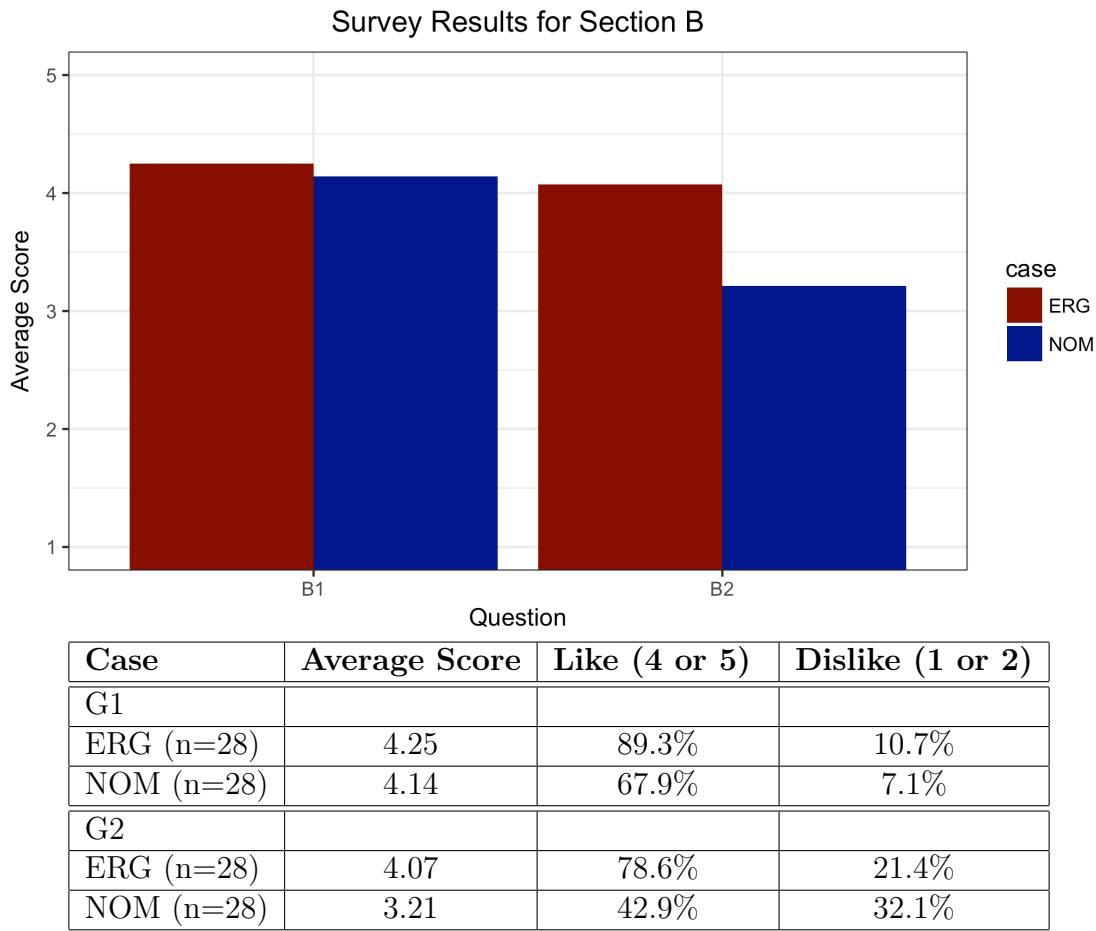


Figure 4.35: Survey Results for Section B

It is fairly rare for an ergative-marked element to be followed by the focus marker *ta* or the topic marker *cāhī*, although there are scattered examples in the corpus. I found three examples with *ta* and three with *cāhī*.

- (170) a. *sāno bha-epani lāun-dina maile ta yo ta*
 small COP.PERF-EVEN wear-PRES.3.SG.NEG I.ERG FOC this FOC
 ‘Even if it is smaller, I will not wear this.’ [V001001004; M9]
- b. *uniharu-le cāhī tyo mansthiti nai tyo ban-ā-era*
 PRO.PL-ERG TOP that mentality EMPH that build-CAUS-CONJ
 \bar{a} -ekā hūdā rah-echan
 come-PERF.PL COP.WHILE stay-MIR.PERF.3.PL
 ‘(As for the them), it is apparent that they arrive having built up that

sort of mentality.'

The focus and topic markers are not rare in the corpus; I counted 276 utterances which contained a *ta* marker and 109 with *cāhi*.³⁴ However, they are somewhat rare on pronouns, and they are even rarer with ergative-marked pronouns. To take the example of the first person singular, of the 64 overt instantiations of the nominative form *ma*, it was followed by *ta* 9.4% of the time (n=6) and by *cāhi* 6.3% of the time (n=4). In the ergative form, out of the 67 examples, the *ta* marker appeared 1.5% of the time (n=1) and the *cāhi* marker 3.0% of the time (n=2).

	Topic	Completive	Focus	Background
Nominative (n=22)	68.2% (n=15)	9.0% (n=2)	4.5% (n=1)	18.2% (n=4)
Ergative (n=32)	50.0% (n=16)	9.4% (n=3)	6.3% (n=2)	31.3% (n=11)

Figure 4.36: Argument Positions in Transitive Present Clauses

I examined the S_t argument placement in the 54 transitive simple present non-modal main clauses in the NNSP sample. I coded the placement as "Topic" if the S_t is the first element, "Completive" if it is the second (for example, if it is preceded by a core argument or any type of adverb or phrase), "Focus" if it is immediately preverbal and also not the first element, and "Background" if the subject follows the verb.³⁵ Figure (4.36) tabulates the results for nominative and ergative forms of the subject. For S_t in general, the most common position is in Topic position (the default position in an SOV default word order language). In both cases, it is least common in the Focus position, and somewhat rare in the Completive position, but more common in the Background position. The main difference between ergative and nominative forms of the verb is that the ergative is somewhat more common in the Background position.

34. I do not include in this total 21 examples of *cāini*, although they might be related and both likely stem from the verb *cāhinu*.

35. There were three examples which I call "VP-fronted" in which the verb itself is placed in the first topic position and all the remaining material comes afterward. There were three examples, all with ergative-marked subjects, and in the figure I consider the subject to be in Backgrounded position in this case.

These results accord with the intuitions of elicitation respondents. In particular BB noted that the ergative form sounds much more natural than the ergative form as it is postposed in the following examples from the NNSP sample:

- (171) a. *ani pheri 'lagāu' bhan-chau timi-le*
 and again 'wear.IMP' say-PRES.2.SG you-ERG
 'And then once again 'wear it!' you'll say.' [V001001004; M7]
- b. *yo lā-nu par-yo bhane, tei lān-chu*
 this take-NON.FIN need-PERF.3.SG COND, that.EMPH take-PRES.1.SG
 hai maile
 PRT I.ERG
 'If I have to take this one, I'll that one as well.' [V001001004; M7]
- c. *mil-ā-era pani artha aa-ena ke, kasto*
 arrange-CAUSE-CONJ also right come-PERF.3.SG.NEG what, how
 kurā gar-cháu timi-le ?
 thing do-PRES.2.SG you-ERG ?
 'Even after discounting it is not enough, what are you talking about?'
 [V001001004; M7]

So S_t arguments in general tend to be in the Topic or Background position, which both consist of old information but differ in terms of prominence. Ergative marked objects, interestingly, are somewhat more prevalent in the Background position, which goes against the general idea that the ergative is associated with greater prominence.

4.5.2 Categorical Propositions

The notion that ergative marking relates to a thetic/categorical distinction is based upon observations that elicitation respondents have made about sentences like the following:

- (172) a. *prakash/prakash-le curot khān-cha*
 Prakash/Prakash-ERG cigarette eat-CONT-PRES.3.SG
 'Prakash is smoking/smokes a cigarette.' [PK]

- b. *shikāri/shikāri-le* *mrigā samāt-dai-cha*
 hunter/hunter-le deer catch-CONT-PRES.3.SG
 ‘A/the hunter is catching deer.’ [TD]
- c. *jogi/jogi-le* *siddha māg-dai-cha*
 holy.man/holy.man-ERG alms beg-CONT-PRES.3.SG
 ‘A/the holy man is begging for alms.’ [BB]
- d. *tyo khaire/khare-le* *foto*
 that white.person/white.person-ERG photo
kic-dai-ho
 photograph-CONT-PRES.3.SG
 ‘That tourist is taking a picture.’ [RM]

(172a) is ambiguous between a habitual, ongoing, and future-oriented reading, while the rest are marked as continuous. The subject in (172a) is also necessarily definite, being a proper name, and the subject in (172d) is definite because it is a determiner, but the other two are ambiguous between definite or indefinite readings because Nepali does not have obligatory determiners that carry presuppositions of definiteness as English does. In each sentence both the nominative and ergative form of the subject is grammatical.

A general intuition of elicitation respondents is that the clause with the ergative form is more “about” the subject. The ergative marking provides an extra statement: it makes clear that the action belongs particularly to the subject, that this entity in particular is being picked out as the subject of the clause. The action might describe a particular habit or addiction in (172a), or be an action that particularly characterizes the subject, as in (172b).

These intuitions are the motivation behind considering Kuroda (1972)’s thetic/categorical propositions as a distinction relevant to the expression of ergativity in Nepali. The categorical proposition consists of the apprehension of the logical subject (the topic of the sentence, i.e., what it is about) and then the predication of that subject. This is contrasted with a thetic proposition, which is a statement about a state of affairs

with no particular entity of the clause being singled out as what the clause is about.

In the simple present tense as in (172a), it will tend to be associated with individual-level predication, but does not require the predication to be individual-level. Habitual and generic readings must be categorical propositions, but categorical propositions can also be stage-level (Ladusaw 2000). For example, in response to a particular question about the activities of a particular hunter, “The hunter is hunting deer right now” is a categorical proposition. This means that Butt and Poudel (2007)’s observations about the ergative being associated with habitual readings can be extended to other imperfective tenses by making it a general property of the proposition rather than simply a property of the predicate.

In the terminology of topicality, we can consider a categorical proposition to contain a particular entity that is the topic, whereas in a thetic proposition the entire clause is the topic. Note that the topic (logical subject in Kuroda’s terms) can be any element of the clause (noun phrases, prepositional phrases, etc.), although it is typically the S_t of transitive clauses. The contrastive topic marker in Nepali, *cāhi*, is a very clear example of a grammatical device that apprehends a particular element of the clause as the topic. It also emphasizes a contrast between other possible topics which can be a signal that the speaker is changing the subject by introducing a new topic, similar to the “As for” construction in English.

The ergative marker, *-le*, on the other hand, may only attach to the S_t . It does not, as we have seen, only mark contrastive topics. So the argument is not that the ergative marker is a topic marker in Nepali, but rather that it correlates specifically with categorical propositions in which the S_t is the topic.

4.5.3 QUD, Definiteness, and Aboutness

A categorical proposition is about a particular entity and a thetic proposition is about a state of affairs. In both Nepali and English, a statement like “Prakash is smoking a

cigarette” could be either thetic or categorical depending upon whether it is construed as a general statement about the world or a statement particularly about Prakash. Only the categorical reading is possible with explicit topicalization: “As for Prakash, he’s smoking a cigarette.” Under a question-based model of discourse structure as in Roberts (1996), every felicitous statement provides an answer to an implicitly or explicitly-stated question. The topic of a sentence in this sense is the Question Under Discussion that is addressed by the statement. The QUD determines whether the response is thetic or categorical:

- (173) a. Thetic-Supporting Question: Why is it smoky in here?

Answer: Prakash is smoking a cigarette.

- b. Categorical-Supporting Question: What is Prakash doing right now?

Answer: Prakash is smoking a cigarette.

Ladusaw (2000) argues that the subject of a categorical proposition must be strongly-construed, and Kuroda (1972) notes that categorical propositions are associated with definite determiners in English. Similarly, topics typically consist of old information that is known in the context of the discourse.³⁶ This means that while the S_t in the clause of a thetic proposition may be definite or indefinite, the topic of a categorical proposition generally must be definite. In the following sentences, imagine that there is a particular hunter known to both speakers to have been walking around the woods nearby.

- (174) a. Thetic-Supporting Question: Why was there just a loud bang?

Answer: A hunter is outside hunting deer. / The hunter is outside hunting deer.

- b. Categorical-Supporting Question: What is the hunter doing on our property?

36. While some languages contain topic markers that specifically introduce new discourse referents, *-le* is more often associated with old information.

Answer: #A hunter is outside hunting deer. / The hunter is outside hunting deer.

More generally, a question about a general state of affairs can be answered with a categorical proposition, but a question about a particular entity cannot be answered with a thetic statement:

- (175) a. Thetic-Supporting Question: What is happening in this photograph?

Answer: A holy man is begging for alms. A tourist is taking a picture of the market. / As to the holy man here, he's begging for alms. That tourist is taking a picture of the market.

- b. Categorical-Supporting Question: In this photograph, what is the holy man doing and what is the tourist doing?

Answer: #A holy man is begging for alms. A tourist is taking a picture of the market. / As to the holy man here, he's begging for alms. That tourist is taking a picture of the market.

Therefore, if ergative marking is particularly associated with categorical propositions, then we expect (1) that the ergative S_t must be definite, and (2) that the nominative S_t cannot be the answer to a question about a specific entity (categorical-supporting).

Furthermore, the QUD is related to a given topic X in the following way: it is a subquestion of the more general question “What about X?” So if the statement “The hunter is outside hunting deer” is categorical, then it provides a partial answer to the question “What about the hunter?” In general, we should find the ergative to be preferred when the question is explicitly “What about X?”³⁷

- (176) a. *amritā ciniyā bhāshā sik-che ani sonām ni ?*
Amrita Chinese language learn-PRES.3.SG.F and Sonam AND ?

37. A fairly literal equivalent expression in Nepali would be *ke ko bārema*. In the Kathmandu Survey I alternated this with a topic marker, but I understand this to be a mistake in the survey design because both are fairly unnatural ways of asking the question as opposed to the *ni* discourse particle.

‘Amrita studies Chinese. And what about Sonam?’ [BA]

- b. *sonam-(le) koriyan bhāshā sik-cha*
Sonam-(ERG) korean language learn-PRES.3.SG
‘Sonam studies Korean.’ [BA]

Judgments from elicitation respondents suggest that there is a general tendency for the ergative to be associated with categorical propositions, but it is not a one-to-one correspondence. As with the individual-level predication theory, there is a reasonable amount of variation in responses. The strongest inclination is for the ergative-marked element to be presupposed in the discourse. Even if the question is thetic, the ergative is only used when the subject is known to the speakers:

- (177) a. *bahira ke hun-dai-cha* ?
outside what COP-CONT-PRES.3.SG ?
‘What is happening outside?’ [TD]
- b. *shikāri-le mrigā samāt-dai-cha*
hunter-ERG deer catch-
‘The hunter is hunting a deer.’ [TD]

Note however that we have already seen some examples in which the subject is indefinite, so while this may be a tendency it is not a categorical one. In any case, indefinite subjects can be topics in some circumstances.

- (178) *carā-le khān-cha*
bird-ERG eat-PRES.3.SG
‘A bird will eat (it).’ [TD]

Respondents varied in whether or not they accepted the nominative form with an explicitly categorical-supporting question. Some disliked it (PK, TD, RM), but others found it acceptable. Judgments sometimes changed when I asked the same question at a later date, which is a strong indication that the QUD is not the only factor determining the expression of the ergative.

- (179) a. Context: What is Prakash doing right now?
- b. *prakāsh/prakash-le curot piun-dai-cha*
 prakash/prakash-ERG cigarette drink-CONT-PRES.3.SG
 ‘Prakash is smoking a cigarette.’ [PK]
- c. Context: Why does it stink in this room?
- d. *prakāsh curot piun-dai-cha*
 prakash-ERG cigarette drink-CONT-PRES.3.SG
 ‘Prakash is smoking a cigarette.’ [PK]

Section (I) of the Kathmandu Survey specifically contrasted a Thetic/Categorical QUD with the same response in a four-way contrast. In section (E), each trial specifically contained a “What about X?” question. In both cases, there survey results did not show any particular pattern beyond the general trend for the ergative to be preferred on all sentences. This may simply indicate that whether or not there is a tendency for categorical propositions to be associated with the ergative form, both forms are generally acceptable. The results are in Figure (4.37).

4.5.4 Ergative Marking and the (Implied) Question

A comment which frequently arose in discussions about the nature of the QUD was that the ergative form in a response follows an ergative form in the question:

- (180) *ko/kas-le khānā pak-āun-dai-cha ?*
 who/who-ERG food cook-CAUS-CONT-PRES.3.SG ?
 ‘Who is cooking food?’ [RM]

- (181) *priyā/priyā-le khānā pak-āun-dai-chu*
 Priya/.Priya-ERG food cook-CAUS-CONT-PRES.1.SG
 ‘Priya is cooking food.’ [RM]

If the question is *Ko?* then the response is nominative, and if the question is *Kasle?* then the response is ergative. I attempted to circumvent this issue in some cases by having the question be about a particular entity but expressed in a copular

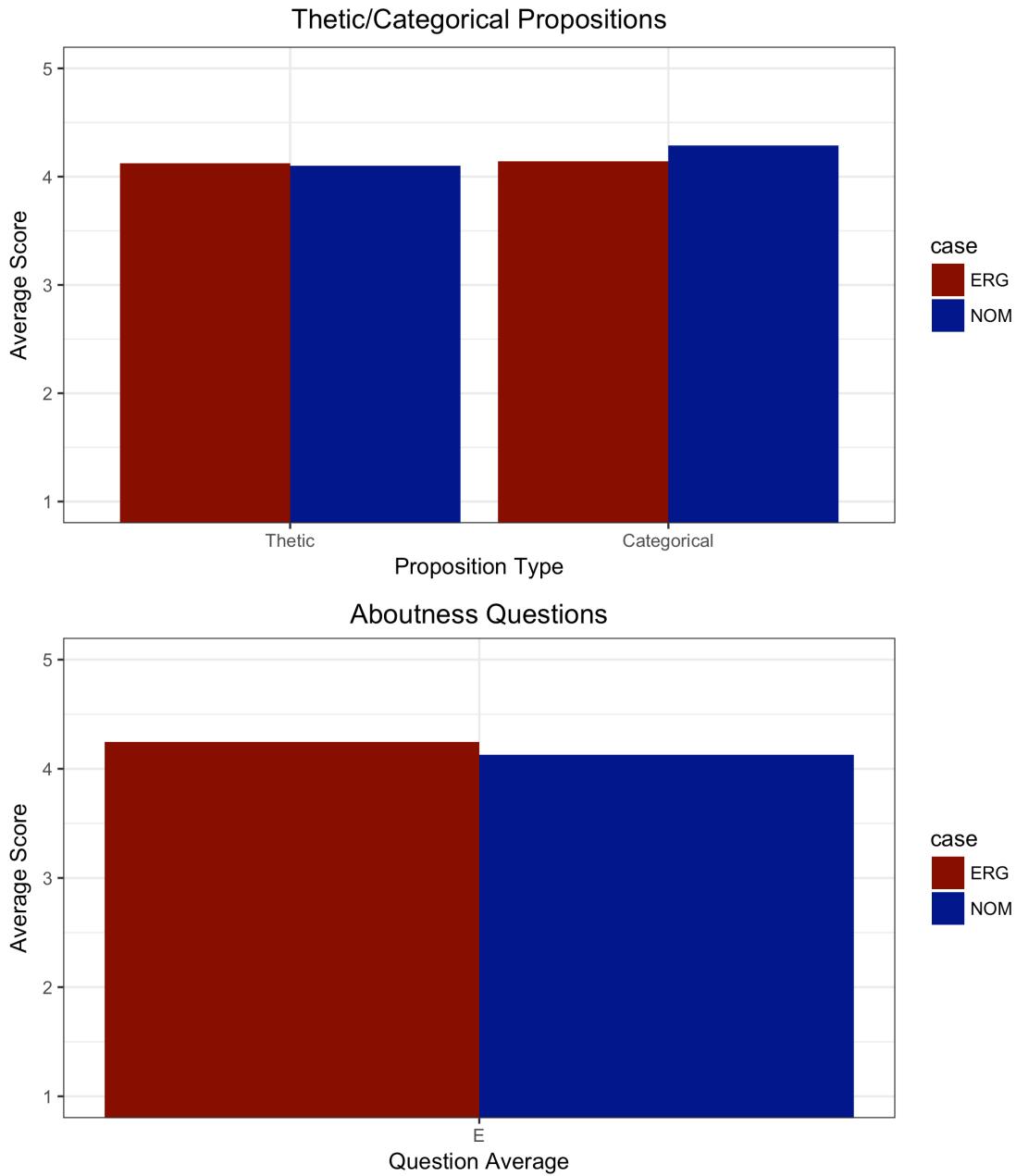


Figure 4.37: Ergative Marking Results for Categorical Propositions

clause. For example, for the categorical statement “The hunter is hunting deer?” the question “Where is the hunter?” could be considered to set up a categorical response (although TD noted that the sentence here is about a location rather than a hunter). This effect may have biased the results of the survey. I tabulated, for each transitive imperfective trial question, whether the S_t of the question was in

the nominative case, ergative case, or whether the question was a copular clause or some other circumlocution. The results, given in Figure (4.38), indicate that the ergative is still preferred in each case, but if the question uses the nominative case, the nominative case is somewhat more acceptable.

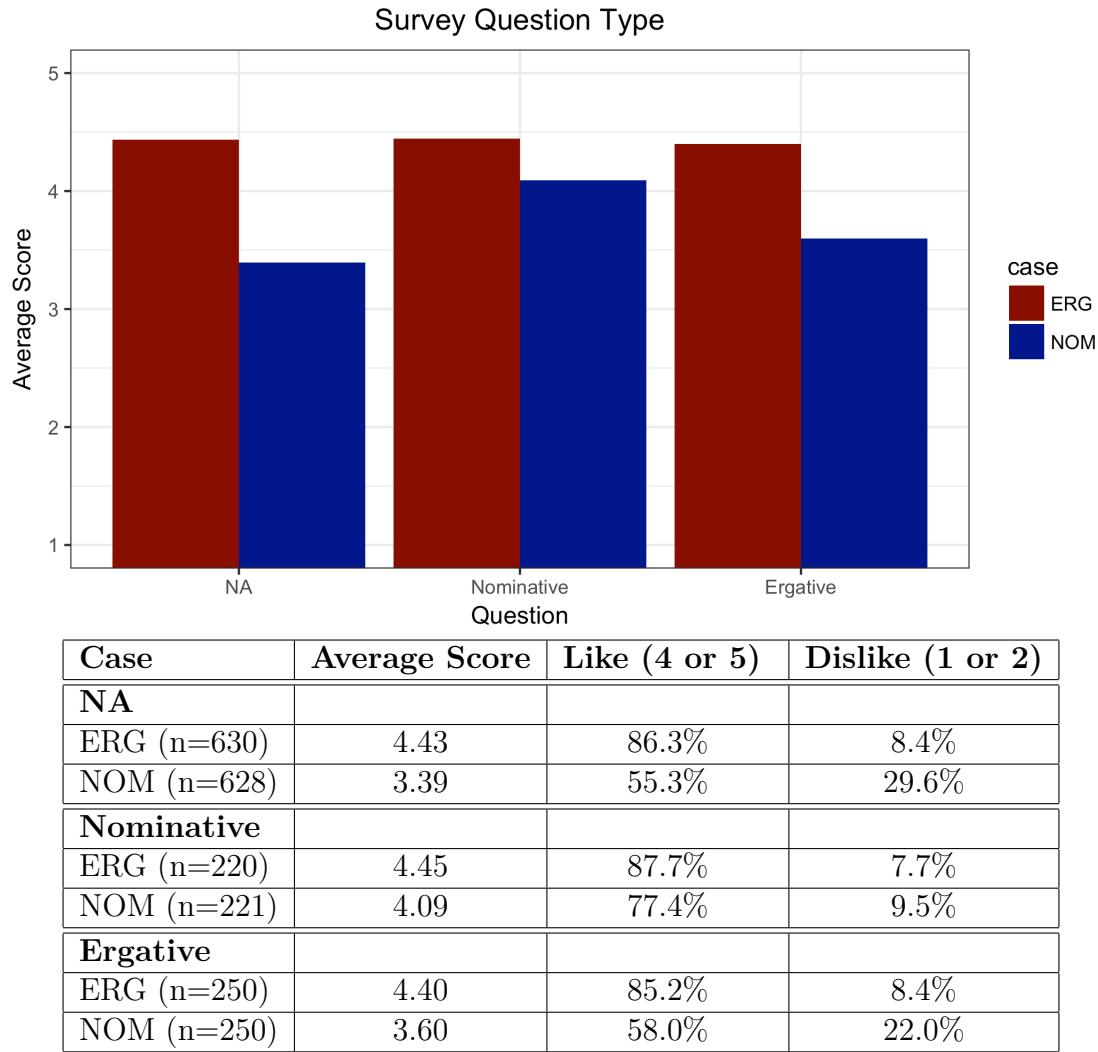


Figure 4.38: Ergative Marking Results and Case on the Survey Question

However, the mere availability of a nominative question word *Ko* in a transitive question is problematic for the theory that a categorical proposition must be marked with the ergative. The implication is that there can be a question underlying the statement *Priya khānā pakāundaicha* ‘Priya.NOM is cooking food,’ which is *Ko khānā pakāundaicha?* ‘Who.NOM is cooking food?’ But this question is requesting infor-

mation about a particular entity and not a state of affairs. So the response must be categorical. Therefore there cannot be a one-to-one correspondence between ergative marking and categorical propositions.

4.5.5 Summary

The associations in this section all derive from the ergative having a greater prominence in the discourse. The S_t position is typically the topic, and we find a correlation between the ergative and prominent topics (we also find it associated with backgrounded non-topical given information, which is a puzzle). Related to the notion of characterizing predicates and predicate prominence, we find that ergative-marked subjects are associated with categorical propositions. These associations do not relate to the subject as an effector of the event, but rather result from the direction of attention to the subject itself. In chapter 6 I argue that this is inherent to an optional case marking system, in which the speaker's choice to use the case marker will increase its discourse prominence.

4.6 Overall Summary of Feature Correlations

Figure (4.39) is a summary of the feature correlations discussed in this chapter. This analysis broadly supports the major observations that have been made in the literature about Nepali, as well as most of the correlations we expect from literature on OEM and transitivity. However, the only associations I take to be truly categorical are (1) the restriction of ergative marking to transitive clauses; and (2) the association between ergative and perfective verb forms. Every other association represents a tendency of greater or lesser strength. These represent pragmatic inferences that a listener might make about the usage of the ergative in a domain where it is not categorically determined. Thus native speakers will differ in their judgments.

All of these associations ultimately derive from two sources. The first involves the meaning of the ergative marker on the S_t of a transitive event. A prototypical transitive event is one in which the subject instigates, enacts, and completes an event which has an effect on an object. This is the source of feature associations related to a transitivity prototype and an S_t prototype. In general, the usage of the ergative marker is associated with highly transitive events. However, unlike with Hindi, ergative marking has no association with Agency or Volitionality. Furthermore, ergative marking is associated with subjects that are unexpected or infrequently found in S_t position. I will argue in chapter 6 this arises from general markedness principles.

Feature	Correlation	Categorical	Source
Interpretation of the Event			
Perfective verb forms	Positive	Yes	Effector
Intransitive clauses	Negative	Yes	Effector
Copular clauses	Negative	Yes	Effector
Individual-level predicates	Positive	No	Prominence
Characterizing predicates	Positive	No	Prominence
Realis Mode	Positive	No	Effector
Interpretation of the Subject			
Animacy in Common Nouns	Negative	No	Prominence/Effector
First Person Pronouns	Negative	No	Prominence/Effector
Kind Readings	Positive	No	Prominence
Strong Construal of Quantifiers	Positive	No	Prominence
Agency	None	-	-
Volitionality	None	-	-
Interpretation of the Object			
Elided Objects	Positive	No	Prominence
Affectedness	None	-	-
Individuation	Positive	No	Effector
Discourse Associations			
Focus	None	-	-
Topic	Positive	No	Prominence
Definiteness	Positive	No	Prominence
Categorical Propositions	Positive	No	Prominence

Figure 4.39: Overall Summary of Ergative Feature Correlations

Secondly, there are feature associations that do not seem to involve the effectiveness of the event per se, but rather correlate with discourse prominence. The marked S_t is associated with categorical propositions, and therefore definiteness, topicality, kind readings of the subject, strong construals, and characterizing predicates. Discourse prominence is an inherent feature of any optional case marking system, and

we expect most if not all of these features to also be found in other OCM systems like differential object marking.

In chapter 5 I discuss ergativity from a syntactic point of view before returning to this markedness-based account in chapter 6.

Chapter 5

The Syntactic Profile of Nepali Ergativity

Modern linguists first began to study ergativity as a unified phenomenon following Dixon’s 1972 monograph on Dyirbal, and during most of that decade ergativity was primarily studied through a functional-typological lens. During the following two decades syntacticians of all theoretical persuasions began to take interest in ergative languages (DeLancey 2004). For formal syntacticians in the generativist tradition, ergative patterning provided a challenge to the traditional understanding of case theory derived from Government and Binding Theory (Baker and Bobaljik 2017). This has led to a substantial theoretical literature on the nature of ergativity as a syntactic phenomenon.

In the Discussions chapter to follow this one, I analyze the ergative case marker in terms of its semantic and pragmatic functions. I will argue that the *-le* marker contributes the same semantic meaning as an ergative and as an instrumental case marker. It marks the given argument as the effector of the event represented by the clause.

This approach says little about the role that syntax plays in ergative patterning.

The purpose of this chapter is to profile the Nepali system among an overall typology of ergative languages in order to demonstrate that Nepali ergativity is a morphological phenomenon which has a minimal effect on structure.

In the first part of this chapter, I demonstrate that Nepali lacks ergative alignment in those grammatical domains where it would be expected in a syntactically ergative language: as a syntactic pivot in subordinate clauses, as well as in adjectival and verbal cross-referencing. Ergative patterning is thus restricted to the domain of nominal case morphology.

In the second part of this chapter I discuss three theories of ergative case that make differing predictions about ergative alignment: structural case, inherent case, and dependent case. Ergativity is assigned by the syntax for structural case and inherent case, and it is assigned by the morphology in dependent case. The ability of ergative marking to appear in non-finite clauses and adherence to the Marantz case generalization against derived ergatives indicates that Nepali ergativity is likely a dependent case. Furthermore, Nepali does not fit neatly into the typology of ergative patterning given by Legate (2008)'s inherent case analysis because ergative case is found in relative clauses.

5.1 Morphological and Syntactic Ergativity

Ergative patterning in a language may be present at one level of structure and not another. We have been considering ergative patterning at the morphological level, particularly as manifested in the case morphology. Ergative morphology may or may not correspond to ergative patterning at the syntactic level. The question, as articulated by Anderson (1977), is whether ergativity in a given language is syntactically “deep” or “shallow.” Anderson argues that ergativity is relatively superficial in most languages. Anderson’s perspective is a historical one, in which ergative morphology

can be thought of as a fossilized relic of an earlier form of the language; a syntactic construction (such as the passive) is reanalyzed and morphologically reappropriated without any effect on the underlying syntax.

5.1.1 Subjecthood Diagnostics

In support of this notion is the observation that for many morphologically ergative languages, syntactic diagnostics for subjecthood pattern along nominative-accusative lines. In other words, a particular diagnostic will group together S_i/S_t against O (nominative-accusative) even though case-marking groups S_i/O against S_t (ergative-absolutive). The common subjecthood diagnostics are reflexivization, coordination, and control (Keenan 1985). For all three of these diagnostics Nepali exhibits a nominative-accusative syntactic pattern.

Reflexivization

Reflexive pronouns in Nepali (as in other Indo-Aryan languages) are coreferential with the subject of the clause. The general form of the reflexive pronoun in Nepali is *āphu/āphai*.

- (182) a. *ma_i āphu_i-lāi dekh-chu*
I self-ACC see-PRES.1.SG
'I see myself.' [BB]
- b. *mai-le āphu_i-lāi dekh-ē*
I.OBL-ERG self-ACC see-PERF.1.SG
'I saw myself.' [BB]
- c. *ma_i-lāi *āphu_i-le dekh-ē*
I-ACC *self-ERG see-PERF.1.SG
*'Myself saw me.' [BB]

For both examples (182a) and (182b), the reflexive pronoun must be coreferential with the subject of the sentence. This is true whether the subject is unmarked, as

in the imperfective clause in (182a), or marked ergative, as in the perfective clause in (182b). This represents a nominative-accusative alignment pattern in the syntax which differs from the morphological expression of ergative marking in (182b).

With an ergative-absolutive syntactic alignment we would expect the pattern in (182c) to be grammatical instead: a reflexive pronoun in a transitive perfective clause would be coreferential with O.¹

The genitive form of the reflexive pronoun is *aphno*, and it too must be coreferential with the subject of the clause. This contrasts with the genitive form of regular pronoun *us-ko* ‘his’, which may refer to another contextually relevant entity. In this context, KS is presented with a situation in which Ram and Vijay are at a dinner party.

- (183) a. *rām_i aphno_i mec-mā bas-yo*
 ram self.GEN chair-LOC sit-PERF.3.SG
 ‘Ram sat in his (Ram’s) chair.’ [KS]
- b. *rām_i us_{i/j}-ko mec-mā bas-yo*
 ram PRO.OBL-GEN chair-LOC sit-PERF.3.SG

1. As a reflexive pronoun, then, *āphu* should never be able to take an ergative case. Because the referent of *aphu* is human, it cannot be an instrumental either. Yet the corpus contains many instances of *āphu-le* and *āphai-le*. This is due to an extension of the reflexive pronoun in which it is referential with a contextually relevant discourse entity:

- (1) a. *ainā-mā kas-le lugā dekh-cha ?*
 mirror-LOC who.OBL-ERG clothes see-PRES.3.SG ?
 ‘Who sees clothes in the mirror?’ [BB]
- b. *aphāi-le lugā dekh-chu*
 REFL.EMPH-ERG clothes see-PRES.1.SG
 ‘I (myself) see clothes (in the mirror).’ [BB]

This raises the question of whether some version of (182c) could be grammatical under this extended interpretation of the reflexive. BB speaker finds this sentence to be potentially grammatical but somewhat stilted:

- (2) *aphāi-le ma-lāi ainā-mā dekh-chu*
 REFL.EMPH-ERG I-ACC mirror-LOC see-PRES.1.SG
 ‘I (myself) see me in the mirror.’ [BB]

In any case, this extended usage of the reflexive pronoun is not germane to the point at hand, because it does not have to be coreferential with the subject of the clause.

‘Ram sat in his (Ram’s/Vijay’s) chair.’ [KS]

In a transitive clause, *aphno* must be coreferential with the S_t , indicative of an accusative pattern, rather than the O, which would be indicative of an ergative syntactic alignment.

- (184) a. *vijāy_i-le rām_j-lāi aphno_i mec-mā bas-ā-yo*
vijay-ERG ram-ACC self.GEN chair-LOC sit-CAUS-PERF.3.SG
‘Vijay seated Ram in his (Vijay’s) chair.’ [KS]

Coordination

As noted by Wallace (1982), Nepali allows a subject coordination in multi-clausal sentences even when there is an ergative-nominative clash:

- (185) [*āja ekdam kabādi pu lagā-era*] \oslash *ā-eko chu ke*
[today very trashy PRT wear-CONJ] \oslash come-PRES.PERF.1.SG what
maile
I.ERG
‘Today I came here wearing trashy (clothes).’ [V001001001; M3]
- (186) *tapāñ-le* [\oslash *tyo jasto “jangal wāk” jā-daakheri*] *bāgh*
you.HON-ERG [\oslash that how “Jungle Walk” go-WHILE] tiger
dekh-nu bhaeko cha ki chaina ?
see-PRES.PERF.HON.3.SG or COP.PRES.3.SG.NEG ?
‘Going on a jungle walk, have you ever seen a tiger?’ [V001002005; M7]

In the first example, the subject of both clauses is the speaker. The outer clause is perfective but unaccusative intransitive, requiring a nominative, and the inner clause is perfective and transitive, allowing an ergative. The (postposed) overt subject is ergative, and it controls both clauses. In the second example, the outer clause is perfective and transitive, requiring an ergative subject, while the inner clause is unaccusative, requiring a nominative subject. The ergative subject controls both clauses.

Genetti (1988) is a study of the similar phenomenon of multiple clause chains in the Tibeto-Burman language Newari. She concludes that topicality plays the decisive role in determining the case of the subject in multiple clause chains. This may be the true for Nepali as well. Regardless, the referentiality between S_i and S_t in multiple clauses, whether they carry absolutive or ergative morphological case, is indicative of a nominative-accusative syntactic pattern.

Control

Another subject diagnostic is that of subject coreference with control. In a matrix clause that takes a subordinate clause as a complement of the VP, the subject is coreferential with the syntactic subject of the subordinate clause.

- (187) $[ketā-le]_i \{ \emptyset_i \text{ kukur-lāi hirk-āu-na } \} \text{ khoj-yo}$
 boy_i-ERG {∅_i dog-OBJ hit-CAUS-INF } seek-PERF.3.SG
 ‘The boy tried to beat the dog.’

The NP *ketā* “boy” is coreferential with the deleted S_t of the subordinate clause rather than the O of the subordinate clause.² It is ungrammatical for the syntactic subject to be coreferential with the O of the subordinate clause:

- (188) * $[kukur-le]_i \{ \emptyset_i \text{ ketā-le hirk-āu-na } \} \text{ khoj-yo}$
 *dog_i-ERG {∅_i boy-ERG hit-CAUS-INF } seek-PERF.3.SG
 *‘The dog tried to get the boy to beat him.’

For all three of these subject diagnostics Nepali picks out S_t and S_i as the subject regardless of morphological case marking, suggesting that ergativity in Nepali is relatively superficial. Anderson (1977) [and possibly also Butt 2017] notes that this is true for Hindi, and it may be considered a general feature of Indo-Aryan.

2. This particular example is based off of Li (2007) (footnote 5). I changed the matrix verb from Li’s “The boy wanted to beat the dog” because the Nepali verb *cahānu* is often used in an impersonal form *cahinu* which takes a dative subject, and the given example was easier to elicit.

Syntactic ergativity in Dyirbal

In contrast to Nepali, there are some languages that do show a deeper ergative patterning. The most famous of these is the Pama-Nyungan language Dyirbal. The following example demonstrates syntactic ergativity in subject coordination for this language:

- (189) a. *ηuma banaga-n^yu*
father.ABS return-NONFUT
'Father returned.'(Dixon 1994: 10)
- b. *ηuma yabu-ηgu bura-n*
father.ABS mother-ERG see-NONFUT
'Mother saw Father.'(Dixon 1994: 10)
- c. *ηuma banaga-n^yu yabu-ηgu bura-n*
father.ABS return-NONFUT mother-ERG see-NONFUT
'Father returned and Mother saw (him).'(Dixon 1994: 12)

From (189a) and (189b) we can see that Dyirbal has an ergative morphological structure: S_i in (189a) and O in (189b) are absolutive, in opposition to the ergative-marked S_t of (189b). In (189c) there is coordination between the two clauses "Father returned" and "Mother saw him", but here it is the O of the transitive second clause that is deleted because it is coreferential with the S_i of the first clause. The syntactic organization of the clause is ergative-absolutive, providing a contrast with the Nepali example of coordination.³

In sum, we have found one point of variation in languages with ergative morphology. Languages like Nepali have a straightforward nominative-accusative syntactic organization, while languages like Dyirbal show ergative patterning of subjects in at least some syntactic constructions.

3. Van de Visser (2003: 179) notes, however, that Dyirbal does not have reflexive pronouns that pattern ergatively, and that there do not appear to be any languages which do.

5.1.2 Agreement

Recall that Nepali has a completely nominative-accusative verbal agreement pattern. The Nepali verb agrees uniformly with S_t and S_i regardless of whether S_t is case-marked ergative or nominative. In this respect Nepali differs from nearly every other Indo-Aryan language with ergative case-marking morphology, which exhibit some form of O agreement (Deo and Sharma 2006).

Regardless of whether verbal agreement represents an underlying syntactic organization or it is considered to be an entirely morphological phenomenon, the absence of ergative-absolutive verbal agreement in Nepali is evidence that ergativity is constrained to a relatively small part of the grammar of the language.

5.1.3 Lack of Covert Ergative Case in Adjectives

Deo and Sharma (2006) present evidence for ergative patterning in adjectives even when it is not overt on nouns, which can be considered evidence that ergativity has a deeper structural component. However, this is not the case for Nepali. In 3.2.5 we discussed the lack of ergative morphology in Marathi local pronouns. The first person singular $m\bar{i}$ has only the nominative form, but in the perfective agreement is with O. This indicates that the subject is underlyingly ergative:

- (190) $m\bar{i} \quad ek \quad \bar{a}mb\bar{a} \quad kh\bar{a}-ll\bar{a}$
I.ERG one mango.NOM eat-PERF.3.SG
'I ate a mango.' (Deo and Sharma 2006: 14)

Further evidence that the subject is covertly ergative comes from adjectival agreement. Adjectives show oblique inflection when they modify an ergative subject and nominative inflection when they modify a nominative subject:

- (191) a. $vedy\bar{a} \quad ash\bar{a} \quad m\bar{i} \quad ek \quad \bar{a}mb\bar{a} \quad kh\bar{a}-l\bar{a}$
foolish.OBL like.OBL I.ERG one mango.NOM eat-PERF.3.SG
'Foolish me ate a mango.' (Deo and Sharma 2006: 14)

- b. *vedī ashī mī ek āmbā khā-te*
 foolish.NOM like.LIKE I.NOM one mango eat-PRES.1.SG
 ‘Foolish me eats a mango.’ (Deo and Sharma 2006: 14)

(191a) is a perfective clause so the subject is ergative. The adjectives modifying *mī* are in the oblique case. (191b) is an imperfective clause so the expected pattern is nominative. The adjectives modifying *mī* are in the nominative case. Thus for Marathi there is evidence of ergative alignment even when it is not visible in the morphology.

Nepali does not have O agreement like Marathi, but there is adjectival agreement and a similar oblique inflection. Historically, nouns like *keto* ‘boy’ would inflect with an oblique case in the plural or when case-marked by ergative (*ketā-le*), accusative (*ketā-lāi*), or genitive (*ketā-ko*). Today this distinction is lost in almost all cases. Even at the turn of the 20th century Grierson observed that “the oblique and direct forms are used interchangeably”(Grierson 1904a: 23). BB used *ketā* generally but noted that *keto* specifies a young boy.

- (192) a. *sāno ketā bahirā cha*
 small boy-PL outside COP.3.SG
 ‘The very small boy is outside.’ [BB]
- b. *sānā ketā-haru bahirā chan*
 small.OBL boy-PL outside COP.3.PL
 ‘The small boys are outside.’ [BB]

The adjective *sāno* modifies the subject, and when the subject is plural, as in the second example, it takes an oblique inflection. Accusative case-marked objects also trigger oblique inflection:

- (193) *mai-le sānā khetā-lāi dekh-ē*
 I.OBL-ERG small.OBL boy-ACC see-PERF.1.SG
 ‘I saw a small boy.’ [BB]

However, ergative-marked subjects do not trigger oblique inflection:

- (194) *euta sāno kethā-le* *āap* *khāyo*
one small boy.OBL-ERG mango eat-PERF.3.SG
‘The small boy ate a mango.’ [BB]

So ergative case is treated identically to nominative case for the sake of adjectival agreement, and there does not appear to be evidence for covert ergative case.

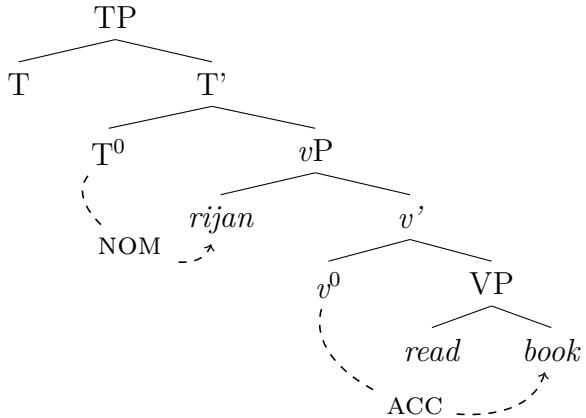
5.2 Theories of Syntactic Ergativity

Within the tradition of generative syntax, there are three basic ideas about how ergative case is assigned. Ergative case marking may be considered a **structural case**, an **inherent case**, or a **dependent case**. The first two theories presume the existence of a separate abstract case which is assigned in the syntax and is realized (perhaps imperfectly) by the available morphological structure. For dependent case theories, ergativity is entirely morphological in nature.

5.2.1 Structural Case, Inherent Case, and Dependent Case

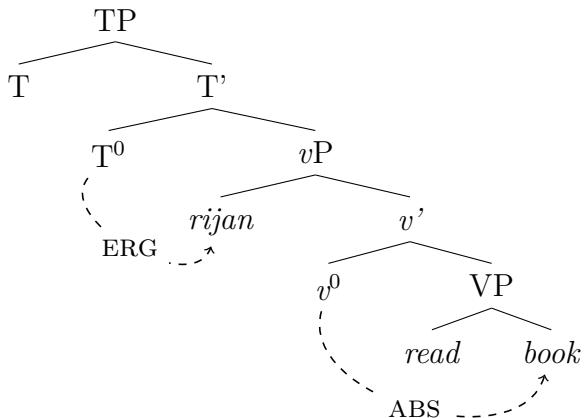
Bobaljik (1993) and Laka (1993) are structural analyses of ergative case. **Structural case** is assigned based on the position of the argument within the syntactic structure of a clause. Most structural case theories of ergative-marking argue that ergative case is assigned by the Tense head in the same way that nominative case is assigned in a nominative-accusative language. This is an illustration of “Rijan reads a book” in a hypothetical nominative-accusative language.

- (195) Structural Case Assignment in a nominative-accusative



The external argument “Rijan” receives nominative case from T^0 (perhaps after it moves to subject position at T), and the internal argument “book” receives structural accusative case from v^0 . The assignment of nominative case is structural because it is assigned by a clause head rather than a head in the verb phrase. Similarly, a structural theory of ergative case may posit that ergative case is assigned by T^0 . The internal argument is assigned absolute case by v^0 .⁴

(196) Structural Case Assignment in an ergative-absolutive language



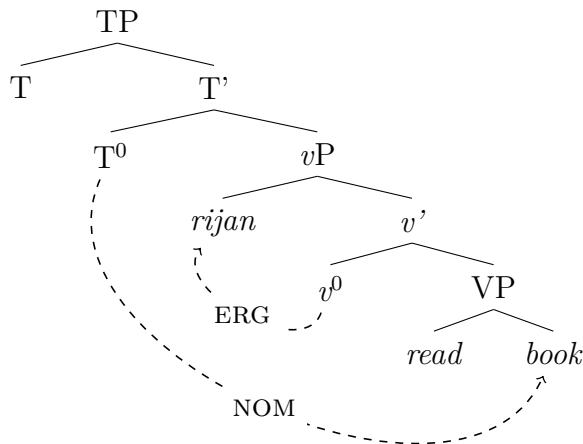
Inherent case analyses of ergativity include Woolford (1997), Laka (2006), Legate (2008), and Legate (2012). Inherent case is assigned locally to an argument in its base generated position (an argument cannot move up to a position to get inherent case).

4. Variations on this viewpoint include accusative case assignment by v^0 in languages with ergative-absolutive-accusative, with absolute unmarked, or ergative case assignment by v^0 .

This position may be associated with a particular thematic role, such as an agent thematic role being generated in the specifier of vP .

In different theories and manifestations, the internal argument may be assigned nominative/absolutive case by T^0 or v^0 , or else it may be a default unmarked case. Crucially, ergative case is assigned by v^0 to its specifier, where the external argument of a transitive clause originates.

(197) Inherent Case Assignment in an ergative language



Finally, Marantz (1991) argues that accusative and ergative case both constitute **Dependent cases**. This theory is further developed in Coon (2013), Baker (2015), and Baker and Bobaljik (2017). Dependent case is assigned to one of the arguments in a VP on the condition that another argument is present in the same clause. Therefore, ergative and accusative case may only be assigned in transitive clauses. If case is assigned to the lower argument of the VP then it is accusative case; if case is assigned to the higher argument then it is ergative case.

Dependent case is entirely morphological, so there is no abstract syntactic case assignment. Dependent case will be assigned separately from overall syntactic structure after the assignment of lexical cases (cases that are assigned idiosyncratically by particular verbs).

5.2.2 Structural versus Inherent Case

The distinction between structural and inherent case is a feature of the Principle-and-Parameters syntax model, and the distinction has carried over into subsequent models including the Minimalist Program (Laka 2006). It arises from the observation that case and semantic function (semantic role) are not equivalent, and in many languages it is necessary to distinguish between (a) structural case that is assigned to an argument by virtue of it being in a particular syntactic position, (b) case that is always associated with a particular thematic role, and (c) case that is assigned idiosyncratically by the verb. These latter two are examples of inherent case. We can see examples of (b) and (c) in Nepali: the postposition *-mā* always marks an argument with a locative theta role regardless of its position in the clause, and in experiencer constructions with the verb *lagnu* the subject must be marked with the accusative *-lāi*:

- (198) *sikka bui-mā khas-yo*
coin floor-LOC fall-PERF.3.SG
'The coin fell on the ground.' [TD]

- (199) *bhok-le us-lāi rīŋatā lāg-yo*
hunger-INSTR PRO.OBL-ACC dizzy feel-PERF.3.SG
'He feels dizzy from hunger.' (Schmidt 1993: 680)

If ergative case is structural, then it should be possible to disassociate ergative case from its semantic role. Laka (2006) compares ergativity in Burushaski with ergativity in Basque, and argues that ergative case is structural in Burushaski but inherent in Basque. For Burushaski, an agent is marked in ergative case in a transitive clause but absolute in an intransitive (unergative) clause:

- (200) a. *ne hír-e phaló bót-i*
the.M man-ERG seed.PL.ABS sow-3.SG.M
'The man planted the seeds.' (Laka 2006: 375)

- b. *ne hir yált-i*
the.M man.ABS yawn-PRET.3.SG.M
‘The man yawned.’ (Laka 2006: 375)

In Basque, by contrast, an agent will always be marked ergative, and the equivalent sentence requires a light verb to make the sentence transitive. This is evidence for Laka that Basque ergativity is inherent rather than structural.

- (201) *gizon-a-k aharrausi egi-n du*
man-DET-ERG yawn do-PERF has
‘The man yawned.’ (Laka 2006: 377)

For Nepali intransitive clauses, Li (2007)’s argument is that for unergative intransitive clauses there is a distinction based on the lexical semantics of verb, in which ergative marking is disallowed with telic agentive verbs and optional with atelic agentive verbs:

- (202) a. *rajnitigya lumbini jā-nu bhayo*
politician.ABS lumbini go-PERF.HON.3.SG
‘The politician went to Lumbini.’ [ST]
- b. *sahuji-(le) jahile.pani khok-nu huncha*
shopkeeper-(ERG) always cough-INF HON.PRES.3.SG
‘The shopkeeper is always coughing.’ [ST]

This would suggest that ergative case is structural. However, in 4.1.2 I argued that clauses like (202a) are unaccusative in Nepali. If I am correct, then this does not constitute evidence either way.

In any case, Legate (2012: 182) criticizes the general argument that ergativity must be structural simply because transitive and intransitives behave differently. For example, a transitivity restriction may also be found on datives for some languages in which dative case is straightforwardly inherent.⁵

Another feature of structural case is that an argument may be assigned different cases depending upon the larger syntactic structure of the clause (Baker and Bobaljik

5. However, imposing a transitivity restriction on Nepali would be somewhat more complicated because marking is optional rather than obligatory.

2017: 3). A particular argument may alternate case within the larger syntactic structure. This contrasts with inherent case, for which case is assigned only to an argument in its original position, and so there will not be case alternations in different syntactic environments.

As a general rule, Nepali ergative case is unaffected by the syntactic environment. This suggests that ergative case in Nepali is not structural.

Nonfinite Clauses

Accusative, nominative and ergative case marking is available to arguments in nonfinite clauses. The fact that ergative marking is possible in nonfinite clauses suggests that ergativity cannot be structural. Structural ergativity is assigned by the T head which would be absent in nonfinite clauses.

- (203) *[anu-le aushadi na-kinn-unjel] sut-dina*
[anu-ERG medicine NEG-buy-until] sleep-PRES.3.SG.NEG
'Until Anu buys medicine she will not sleep.'

In general, the ergative marker is not affected by the nature of the clause. In purposive clauses like the following, there is obligatory subject control, and there can be an ergative-nominative mismatch:

- (204) a. *[anu sutna]-lāi ghar gā-yi*
[anu sleep]-ACC house go-PERF.3.SG.F
'Anu went home in order to sleep.'
- b. *anu_i-le [ø_i sutna]-lāi aushadi kinn-in*
anu_i-ERG [ø_i sleep]-ACC medicine buy-PERF.3.SG.F
'Anu bought medicine in order to sleep.'

The same is true for nominalized clauses:

- (205) a. *hijo ā-eki keti-le kitāb lekh-eki*
yesterday come-PERF.3.SG.F girl-ERG book write-PERF.3.SG.F
che
COP.3.SG.F

‘The girl who arrived yesterday wrote a book.

- b. *hijo nāc-eki keti-le kitāb lekh-eki*
 yesterday dance-PERF.3.SG.F girl-ERG book write-PERF.3.SG.F
che
 COP.3.SG.F
- ‘The girl who danced yesterday wrote a book.

5.2.3 Marantz’ Ergative Case Generalization

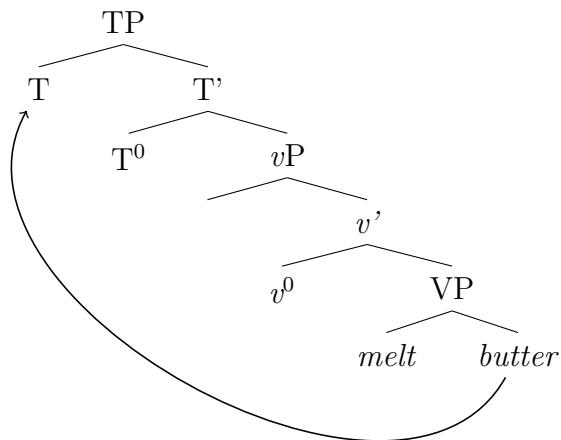
In proposing a dependent case analysis for the ergative, Marantz (1991) makes a strong prediction about the inability of internal arguments to obtain ergative case:

Ergative case generalization: Even when ergative case may go on the subject of an intransitive clause, ergative case will not appear on a derived subject. (Marantz 1991: 13)

An example of a derived subject is the argument of an unaccusative intransitive. The argument of an unaccusative, which is typically a theme, is presumed to originate as the internal argument of a VP and move up to subject position (Perlmutter 1978).

The following example depicts the derivation of *makkan pagl-iyo* “The butter melted”:

- (206) Possible derivation of an unaccusative intransitive



Marantz' generalization predicts that even languages which allow ergative marking on unergatives will disallow it on unaccusatives and any other derived subject. This makes sense in a dependent case analysis, because ergative case is assigned to the higher of two arguments in base position, and the theme argument is internal. For languages which allow ergative marking on unergative accusatives, a dependent case analysis may argue that there is in fact a covert internal argument. But an unaccusative should not be able to get ergative case.

The Marantz ergative case generalization is a natural consequence of an inherent case analysis as well. Ergative case should be unavailable to derived subjects because they do not originate in the external argument position. The ergative case generalization is not a natural consequence of a structural case analysis of ergative marking.

As discussed in 4.1.2, ergative case is never possible on unaccusative intransitive verbs. This provides a nice substantiation of the ergative case generalization for Nepali:

- (207) *das minet pacchi ciyā pāk-cha*
ten minute after tea.ABS cook-PRES.3.SG
'The tea will cook in ten minutes.' [BA]
- (208) *jahāj ādi-mā dub-iyo*
ship.ABS storm-LOC sink-PERF.3.SG
'The ship sank in the storm.' [BA]
- (209) *ghām-mā makkan pagl-iyo*
sun-LOC butter.ABS melt-PERF.3.SG
'The butter melted in the sun.' [ST]

Another example of a derived subject is a passive construction, in which the object moves to subject position. In Nepali, the demoted subject may be marked by the postposition *-dvāra*, (or, according to Verma (1976) the ablative *-bāta*) although for most of my respondents this usage sounded artificial and overly academic, and there

were no corpus examples of case-marked demoted subjects. For the subject of a passive construction, ergative case is not possible, as predicted by the ergative case generalization:

- (210) *khānā pāk-ā-i-dai-cha*
food cook-CAUS-PASS-CONT-PRES.3.SG
'The food is being cooked.' [BB]
- (211) *mrigā-(lāi) mār-i-yo*
deer-(ACC) kill-PASS-PERF.3.SG
'The deer was killed.' [BB]
- (212) *sitā-lāi kitāb di-i-yo*
sita-DAT book give-PASS-PERF.3.SG
'A book was given to Sita.' [BB]

For an inherent case analysis, ergative case is assigned by the head where the NP gets its thematic role. For a dependent case analysis, the relevant factor is whether there are multiple NPs in the same domain (Baker and Bobaljik 2017: 5). In both theories, a critical test of the ergative case generalization comes from constructions with multiple internal arguments, such as the passive of a double object construction or the applicative of an unaccusative verb (Legate 2012: 182).

The common difficulty with passive double object constructions, as noted by Legate, is that for many languages, including Nepali, the indirect object is obligatorily marked dative. This may or may not satisfy a transitivity restriction. In the examples below, *-lāi* is optional on the object of a bivalent verb but obligatory on the indirect object of a double object verb:

- (213) *rām-le kitāb-(lāi) padh-yo*
ram-ERG book-(ACC) read-PERF.3.SG
'Ram read a book.' [BB]
- (214) *rām-le sitā-lāi kitāb di-yo*
ram-ERG sita-DATIVE book give-PERF.3.SG
'Ram gave Sita a book.' [BB]

So (212) does not allow an ergative on “book” either because of the ergative case generalization or because it does not satisfy a transitivity restriction. Furthermore, Nepali does not appear to have the sort of applicatives that would be useful for testing the hypothesis.

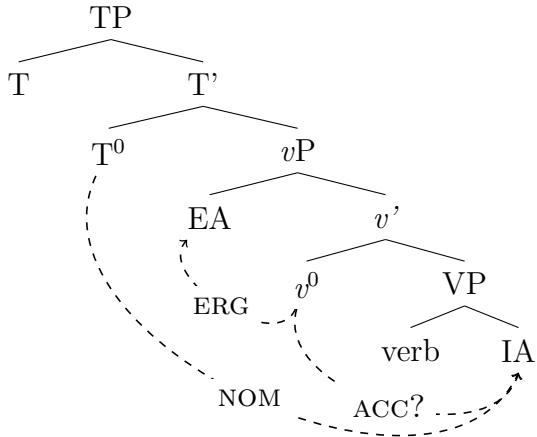
As discussed in 4.1.2, there are some verbs in Nepali that could arguably be considered unaccusative transitives, i.e., verbs with two internal arguments. For these verbs, we do find ergative marking on the subject:

- (215) a. *gaidā-haru-le hāti-lāi gher-e*
 rhino-PL-ERG elephant-ACC surround-PERF.3.PL
 ‘The rhinos surrounded the elephant.’
- b. *jangal-le upatyakā-lāi gher-yo*
 forest-ERG valley-ACC surround-PERF.3.SG
 ‘The forest covered the whole valley.’
- c. *kamilā-haru-le khānā-lāi dhāk-e*
 ant-PL-ERG food-ACC cover-PERF.3.PL
 ‘The ants covered the food.’

It may be the case that the subjects of (215a) and (215c) are agents rather than themes, but this is less plausible for (215b). If true, this would potentially be evidence to support a dependent case analysis, because ergative marking is indeed possible in the presence of another argument, regardless of semantic role.

5.2.4 Legate’s Typology of Ergative Case

Legate (2008)’s inherent case analysis of ergativity accounts for some of the differences between ergative languages by appealing to the morphological interpretation of inherent case.



Nominative case is assigned structurally by T to the highest available argument. Ergative case is assigned by the *v* head to the external argument (EA) in base position. So the external argument of an unergative intransitive will get ergative case if there is no transitivity restriction, otherwise they will get nominative case. The internal argument of an unaccusative will get nominative case.

In Legate's typology there are two kinds of ergative languages, ABS = NOM languages and ABS = DEF (also called ABS = NOM & ACC) languages. Absolutive is not a separate case. Rather, the morphological manifestation of inherent case leads to the appearance of an ergative-absolutive pattern. The only difference between the two types of languages is that for ABS = NOM languages the *v* head assigns structural accusative case to the internal argument (IA), and for ABS = DEF languages it does not (hence the question mark in the diagram).

For ABS = NOM languages, absolute case is just nominative case assigned by the T head, which searches down and assigns nominative case to the internal argument (IA). Because the *v* head does not assign accusative case, the internal argument is available to get nominative case.

For ABS = DEF languages, the *v* head always assigns accusative case to the internal argument. However, there is no morphological expression of accusative case,

and the IA is left unmarked. Thus the nominative S_i and the (accusative) O are both unmarked.

Legate compares the ergative patterning of Warlpiri, Nieuean, Enga, and Hindi with Georgian, categorizing the first four as ABS = DEF and the last as ABS = NOM. She also categorizes languages with ergative splits according to nominal hierarchy as ABS = DEF, and insofar as Nepali counts as a language of that type we can a priori consider Nepali to be of that type. Legate develops five diagnostics for distinguishing ABS = DEF and ABS = NOM, which are summarized by Legate (Legate 2012: 181):

1. ABS = DEF: lacks NOM and ACC morphology
ABS = NOM: may have independent ACC morphology
2. ABS = DEF: caseless DPs bear ABS
ABS = NOM: (no prediction)
3. ABS = DEF: if nonfinite allow for only a subset of cases, ABS is unavailable for S but available for O
ABS = NOM: if nonfinite allow for only a subset of cases, ABS is unavailable for both S and O
4. ABS = DEF: multiple absolutives are possible (e.g. indirect/applicative/adpositional objects)
ABS = NOM: absolute is unique
5. ABS = DEF: A/S agreement, or S agreement
ABS = NOM: S/O agreement

I will examine the behavior of Nepali for each of these diagnostics.

Lack of Accusative Morphology

Legate's ABS = DEF requires that nominative and accusative case marking have the morphological (unmarked) form. Therefore, such a language will not have overt

accusative case marking. This appears to be the case for Nieuean, Enga, and Warlpiri. Hindi has a system of marking direct and indirect objects that is similar to Nepali, which Legate considers to be dative marking. In fact, both Hindi and Nepali require this case marker on indirect objects, but it is variable on direct objects.

Assume the Nepali case marker *-lāi* is not a marker of accusative structural case, but rather an inherent dative case marker. Accusative case is then morphologically unmarked in Nepali. So this diagnostic indicates that Nepali, like Hindi, could be ABS = DEF.

Caseless DPs

If absolutive is the morphological default, then it should be found in constructions in which no abstract case features are assigned to a DP. Legate gives the example of hanging-topic left-dislocation, which is found in Hindi and Nepali. The hanging topic is not given any abstract case marking:

- (216) *tyo keti, tes-le timi-lāi dekh-i ?*
 that girl.ABS, 3.OBL.LOW-ERG you-ACC see-PERF.3.F.SG ?
 ‘That girl, did she see you?’ [BB]

It is not possible for there to be ergative case marking on “that girl,” which suggests that the absolute form is the default. The absolute form is also the only form available when the hanging topic is coreferential with an accusative-marked O, as in (217a), or an *S_i*, as in (217b).

- (217) a. *tyo keti, tai-le dekh-is us-lāi?*
 that girl.ABS, 2.PRO.OBL.LOW-ERG see-perf.2.sg.low PRO.OBL-ACC?
 ‘That girl, did you see her?’ [BB]
- b. *tyo keti, u āja ākh-i ho ?*
 that girl.ABS, 3.PRO today come-PERF.3.F.SG COP ?
 ‘That girl, did she arrive today?’ [BB]

This again follows the prediction for ABS = DEF languages, although ABS = NOM languages make no prediction on this either way.

Cases in Nonfinite Contexts

For ABS = DEF languages, S_i and O have a different Abstract Case, even though the morphological expression of that case is the same. S_i is nominative, and O is accusative. This makes a prediction for nonfinite contexts: absolutive on S should be unavailable because there is no finite T head, but absolutive on O should remain unavailable. Legate demonstrates that this is indeed the case for Hindi nominalized clauses, in which absolutive on S becomes unavailable:

- (218) a. *[rām-ke baithne]-par māā-ne us-ko khānā di-yaa*
 [Ram-GEN sit.nonfin]-LOC mother-ERG him-DAT food give-PERF
 ‘When Ram sat down, Mother gave him food.’ (Legate 2008: 65)
- b. *ilā-ne [rām-ke darvāzā kholne]-par anu-ko dāātĀ*
 ila-ERG [Ram-GEN door.ABS open.NONFIN]-LOC anu-DAT scold.PERF
 ‘Ila scolded Anu on Ram’s opening the door.’ (Legate 2008: 65)

In (218a) the subject of the nonfinite nominalized clause is “Ram”, but it cannot get nominative case because there is no T head, so it is marked genitive. In (218b) the object “door” can get accusative case, so it is marked absolutive.

For ABS = NOM languages, Legate predicts that neither O nor S should be available in nonfinite contexts, and this is true in Georgian, for which the genitive is required in both contexts.

Nepali, unlike either of these patterns, shows no restrictions for S_i or O in nominalized nonfinite clauses:

- (219) a. *[timi-lāi dekh-eko kethi] bahirā che*
 [you-ACC see-PERF girl] outside COP.3.SG.F
 ‘The girl who saw you is outside.’ [BB]

- b. [*timi-le dekh-eko kethi*] *bahirā che*
 [you-ERG see-PERF girl] outside COP.3.SG.F
 ‘The girl who you saw is outside.’ [BB]
- c. [*timi-lāi dekh-eko kethi*] *kitāb lekh-i*
 [you-ACC see-PERF girl] book write-PERF.3.SG.LOW.F
 ‘The girl who saw you wrote a book.’ [BB]
- d. [*timi-le dekh-eko kethi*] *kitāb lekh-i*
 [you-ERG see-PERF girl] book write-PERF.3.SG.LOW.F
 ‘The girl who you saw wrote a book.’ [BB]

So this diagnostic does not appear to be relevant for Nepali. I was unable to find any other type of nonfinite clause in which there are restrictions on case.

Multiple Absolutives

Because the realization of S_i and O as absolutive is the same, for ABS = DEF languages it is possible to have multiple absolutive arguments in the same clause. The facts for Hindi and Nepali are quite similar here. In a verb with imperfective aspect, it is quite possible for a clause to have multiple absolutives. In other words, there is differential ergative marking on the subject and differential object marking on the object, and these are somewhat independent of each other:

- (220) *rām harek din eutā āap khaan-cha*
ram every day single mango eat-PRES.3.SG
 ‘Ram eats a mango every day.’ [BB]

This contrasts with Georgian, in which accusative marking and ergative marking are in complementary distribution such that clauses with two absolutives are not possible:

- (221) a. *glex-i tesavs simind-s*
 peasant-NOM he.sows.it corn-DAT
 ‘The peasant is sowing the corn.’(Legate 2008: 65)

- b. *glex-ma datesa simind-i*
 peasant-ERG he.sowed.it corn-NOM
 ‘The peasant sowed corn.’(Legate 2008: 65)

Agreement

Finally, Legate distinguishes between agreement patterns for ABS = DEF languages and ABS = NOM languages. Some ABS = DEF languages allow agreement with the (ergative-marked) S_t , and some do not. But they should not normally allow agreement with O. Thus the prediction is that for ABS = DEF languages agreement should be with S_i and O or just S_i . For ABS = NOM languages the O is assigned nominative case so there should be agreement with S_i and O.

Hindi does in fact have O agreement when S_t is ergative and O is unmarked. Legate explains this as “aggressive agreement,” such that T looks for something to agree with when nothing else is available. However, this argument is less tenable in other Indo-Aryan languages which presumably show ABS = DEF characteristics. In Gujarati, for example, there is O agreement even when O is case-marked (Deo and Sharma 2006: 73). In any case Nepali, with its straightforward S_t / S_i agreement, falls into the expected ABS = DEF category.

5.2.5 Conclusions

Ergative patterning in Indo-Aryan languages does not as a rule exhibit the deep syntactic ergativity found in Dyirbal. However, even compared to related languages ergative case-marking in Nepali has a minimal impact on the syntax. Hindi and Marathi exhibit ergative patterning in verbal agreement, adjectival cross-reference, and a sensitivity to case marking in subordinate and relative clauses. None of this is found in Nepali.

A complete syntactic analysis of Nepali would likely need to consider ergativity as a dependent case, such that ergativity is assigned in the presence of an object.

Adherence to the Marantz Case Generalization indicates that it is not structural, nor does it neatly fit into Legate's inherent case analysis. The dependent case analysis is fairly straightforward if we follow the conclusion from 4.1.2 that ergative case is restricted to transitive clauses. The analysis is a little trickier if we accept that ergative marking is possible with unergative intransitives, because it would have to explain why ergativity is variable in both perfective and imperfective clauses with unergative intransitives.

The main visible trace of ergativity in the syntax is a strong adherence to the Marantz Ergative Case Generalization, which could be framed in terms of the semantic roles of the case-marked subject. Ergativity in Nepali is largely morphological and has relatively little impact on the underlying syntactic organization.

Chapter 6

Discussions

三十輻共一轂，當其無，有車之用。

Thirty spokes converge on a hub
but it's the emptiness
that makes a wheel work

Dàodé Jīng Chapter 11

(translation by Pine 2009)

In this chapter, I discuss a prototype analysis of the *-le* marker as the Effector of an event, and argue that all of the feature associations discussed in chapter 4 arise either from this meaning or from general markedness principles that increase the discourse salience of the marked form in opposition to its absence.

In section 6.1 I present the Effector analysis and illustrate how it unifies many of the feature correlations we've seen. In 6.2, I discuss the relationship between semantic markedness and variability in optional marking systems. I argue that markedness represents a cline of opposition from inclusive asymmetry to polar opposition, and this cline is associated with the strength of pragmatic implicatures. At the far end of

this cline an opposition is grammaticalized as a semantic entailment, which I demonstrate with optional gender marking in English. I then apply this notion to the Nepali ergative marker to derive the associations we find between ergativity and semantic properties of the subject. In 6.3 I discuss discourse prominence and its relation to categorical propositions and characterizing predicates. In 6.4 I discuss the grammaticalization of obligatory associations related to ergative marking and event structure.

6.1 A Prototype Analysis of the Nepali Ergative

All of the feature correlations we have seen arise from the pragmatic usage of the ergative marker in grammatical domains where it varies with the unmarked nominative form. Most of the feature correlations are predicted by the Transitivity Hypothesis, for which ergativity is associated with higher transitivity. The prediction is that some of these features of transitivity may be present whenever a language has an ergative alignment pattern in case marking, although the hypothesis does not predict which particular correlations will be present in any given language.

Crucially, the particular feature correlations come about through the meaning of the morphological form itself. The ergative postposition *-le* is homophonous with the instrumental postposition. If we consider the semantic contribution of *-le* to be the same whether it marks a core argument (as an ergative) or an oblique argument (as an instrumental), then we can explain many of these puzzling feature correlations. This will be the focus of the current chapter, in which I will argue that the best characterization of this meaning is that *-le* marks an **Effector** of the event described by the clause.

6.1.1 Peripheral Usages of *-le*

First, let us review the various usages of *-le* outside of the main subject marking paradigm. As an **instrumental marker**, *-le* marks oblique arguments.

- (222) a. *maile camcā-le bhāt khā-ē*
I.ERG spoon-INSTR rice eat-PERF.1.SG
'I ate rice with a spoon.' [TD]
- b. *dudh-le ketā-haru-lāi pos-cha*
milk-INSTR child.OBL-PL-ACC nourish-PRES.3.SG
'Through milk (one) nourishes children.' [SB]
- c. *aba kāt-ne hisāb-le tyo gar-era diskāunt*
now cut-NONFIN account-INSTR that do-CONJ discount
gar-idi-nus
do-BEN-IMPER.HON
'Now, doing it with a "cutting calculation," give me a discount.' [V001001004;
F4]
- d. *bishnu siment-le/#kāmdā-haru-le/#rām-le ghar*
Bishnu cement-INSTR/#worker-PL-INSTR/#Ram-INSTR house
ban-āu-dai-cha
build-CAUS-CONT-PRES.3.SG
'Bishnu is building a house using cement/#workers/#Ram.' [BB]

In (222a) the spoon is a tool by which I, the agent, enact the event of eating rice. In (222b) it is a little more difficult to tell whether 'milk' is an inanimate ergative-marked subject ("Milk nourishes children") or an instrumental ("Through milk (one) nourishes children"). This is because in either case 'milk' is in the same place along a causal chain by which milk nourishes children. The question of whether or not 'milk' is the subject, which is not at all obvious to native speakers, is the extent to which 'milk' is construable as the initiator of the event. It is answered by determining whether or not the verb can agree with the argument, and from SB's judgments I believe that it is in fact an instrumental. The third example requires some context: the customer has asked the shopkeeper whether the shopkeeper tailors the clothes

after they are purchased. The shopkeeper replies that this service is not included and the customer will have to tailor it herself. She responds with (222c), that he should determine the price using a “cutting calculation” (i.e., taking the additional price of tailoring into account). Here it is clear that the cutting calculation has not instigated the event of determining the price, but it is rather an important component in the process of making that determination. (222d) demonstrates that this instrumental usage of *-le* appears to be restricted to non-volitional participants (BB allowed *rām-le* in this sentence only under the assumption that this was the name of a particular type of tool).

Secondly, we find *-le* marking verb forms which are either perfective (*-eko*) or non-finite (*-na*). These **reason clauses**, following Butt and Poudel (2007), are a form of instrumental in which the entire subordinate clause is implicated in enacting the event in the main clause:

- (223) a. *pāunā ḣaun-na-le* *ma timro bihā-mā jā-na*
 guest come-NON.FIN-INSTR I your wedding-LOC come-NON.FIN
 pā-inā
 get-PERF.1.SG.NEG
- ‘Because of guests’ coming, I could not go to your wedding.’ (Butt and Poudel 2007: 10)
- b. “*aphai-le*” *bhan-na-le* *āphno paisā tir-era ...*
 SELF-ERG say-NON.FIN-INSTR SELF.GEN money pay-CONJ ...
- ‘By saying “myself”, (I mean) paying my own money.’ [V001001004; M7]

Thirdly, there are ERG/ACC alternations with modal constructions of obligation:

- (224) *rām-le/rām-lāi* *āp khā-nu par-cha*
 Ram-erg/Ram-ACC mango eat-NON.FIN need-PRES.3.SG
- ‘Ram must eat mangoes.’ [AG]

Here the element marked by *-le* or *-lāi* is volitional, and the alternation correlates with an internalized obligation versus an external force. One way of translating the

sentence with *-le* would be “Ram is *obliged* to eat mangoes,” whereas with *-lāi* it would be “Ram is *forced* to eat mangoes.” The form with the ergative emphasizes that *Ram* is the specific entity who has been designated to enact the event, whereas with the accusative the emphasis is on the effect that the event has on Ram.

For all three of these usages of *-le*, as well as on subjects in main clauses, the form marks an **Effector** of the event. This is not something which is the instigator or main cause of the event, but it is a participant which is crucial in enacting the event and effecting its completion.

6.1.2 Agents and Effectors

As discussed in section 3.2.4, Dowty (1991)’s Agent Protorole includes the following properties: Volitionality, Sentience/Perception, Causation, and Movement. As discussed in section 3.2.5, Næss (2004) takes the typical S_t to be Controlling and Unaffected, while Fauconnier (2011) takes it to be an Instigator and an Affector. In each case, we can distinguish between Agentive properties (Volitionality, Sentience/Perception, Controlling, Instigator) which are properties of instigating an event, and Effector properties (Causation, Movement, Unaffected, Affector) which are properties of enacting/completing an event. Both Agentive and Effector properties are found in prototypical Agents, but Effector properties are also shared by instrumentals. Dowty notes that the entailments for an instrumental argument are Causation and Movement, but not Volitionality or Sentience/Perception (Dowty 1991: 577).¹ These properties are used in different ways in each theory, but in each we can draw a distinction between those properties which relate to the initiation/instigation/ultimate cause/control of the event, and those that relate to its being enacted and completed.

1. Dowty’s entailments pertain to particular predicates, but they constitute a prototype because different predicates will have different entailments. This is distinct from my usage of the term entailment below, in which a case marker entails certain properties and pragmatically implicates others.

A prototypical transitive subject has all of these properties, while instruments (in a wide sense of the word) are any arguments which are *not* agents but which participate in enacting and completing the event. This is schematized in Figure (6.1).

Properties of S_t		
Properties of Instruments		
Dowty (1991)	Sentience/Perception	Causation
Næss (2004)	Controlling	Unaffected
Fauconnier (2011)	Instigator	Affector

Figure 6.1: Properties of Transitive Subjects and Instruments

From Croft (2012: 282) I borrow the following definition of a transitive event prototype: “The most prototypical transmission of force relationship that is profiled by a simple verb in a canonical Transitive argument structure construction is by an initiator with mental capacities exercising her/his control acting on a physical endpoint.” In such an event, the transitive subject will tend to be a volitional, controlling agent who initiates the event. Other participants in the event who are on the left side of the causal chain (that is, those that are involved with effecting the event rather than being affected by it) will be marked as antecedent obliques. Thus the transmission of force will naturally go from Subject to Instrument to Object. While the subject will typically be active throughout the duration of the event, the instrument is implicated in the enacting of the event but not its initiation. From these observations I give the following definition of an Effector:

- (1) **Effector:** a participant implicated in effecting an event.

I prefer the term Effector rather than Causator to emphasize that this participant plays a significant role in effecting or enacting the event but it is not necessarily the

primary instigator of the event or its controller (although it can be). I also distinguish it from Affecter, which implies that this participant has an effect on the O participant, whereas I want to emphasize the role that the *-le* has in effecting the event but remain neutral about its role in having a particular impact on the object. For those other features which are involved in instigating and controlling the event, I reserve the term Agentive.

The prototypical S_t is both an Effector and an Agentive. An instrument is just an Effector. The semantic meaning of the Nepali morphological form *-le* is that the participant to which it attaches is an Effector. The Nepali ergative form can therefore be schematized as an Effector on a transitive subject:

$$(2) \text{ ERG} = S_t + \text{INSTR}$$

The S_t will typically have both Agentive and Effector properties, and the addition of the *-le* marker simply emphasizes the Effector properties. In a perfective clause it is part of an obligatory paradigm in which the S_t must be marked. In imperfective clauses *-le* varies with the nominative. It does not change the truth conditions of the clause because it does not add any new information, but rather the use of the marker makes prominent the subject as an effector of the event.

This is similar to Holisky (1987)'s analysis of NOM/ERG alternations in Tsova-Tush. Holisky also distinguishes between Agentive and Effector roles, and notes that the ergative in that language optionally picks out an Agentive as distinct from an Effector. This leads to pragmatic implications of control and volitionality. It is compatible with my analysis, because different morphological forms in different languages will target slightly different properties. The Tsova-Tush ergative includes agentive properties, but the Nepali ergative does not.

Many of the feature associations we have seen are attributable to the Effector meaning of the ergative. Specifically, all the properties associated with Hopper and Thompson (1980)'s Transitivity Hypothesis derive from *-le* emphasizing the effector

role of the S_t in a transitive clause. Transitivity and Perfectivity are grammaticalized associations which I discuss in section 6.4, but the other associations are implicatures of greater or lesser strength. For example, there is the association with individuated objects from section 4.1.5:

- (225) *u/usle māsu khān-daina*
 PRO/PRO.ERG meat eat-PRES.3.SG.NEG
 '(S)he doesn't eat meat' / '(S)he won't eat meat.'

Speakers may interpret the difference between these two sentences based on whether *māsu* refers to meat in general or meat in a specific instance. The ergative correlates with a usage in which the subject has a greater effect on the object and the event is higher along a cline of transitivity. Similarly, the usage of the ergative is correlated with other features of high transitive like realis mode (hence its preference in definite future verb forms over hypothetical or simple present tense verb forms) and kinesis (considering its possible preference in action-oriented verbs over non-action oriented verbs). I discuss the relationship between Ergativity and Event Structure in section 6.4.

This is a pragmatic implicature rather than a semantic entailment, because the S_t is an effector of the event whether it is marked with *-le* or not. Marking serves to emphasize its role as an effector. This does not necessarily mean that marking it as an effector implies that it is not agentive, because a typical S_t is both. Therefore, we would not necessarily expect to find *-le* to be more common with indirect causation. This association could potentially develop over time as an opposition between nominative and ergative forms, but there is not very much evidence that this has happened in Nepali.²

2. One possible instance of this is TD's intuition that the ergative form in *Prakāsh-le curot khāncha* "Prakash is smoking a cigarette" implies that he was forced to do so by his friends. Also, the usage of *-le* in obligational modals is associated with internalized necessity but it may also emphasize that the subject has less agency in the matter. But these are marginal cases.

The Nepali pattern differs from that of many other OEM languages in that ergative alternations do not correlate with volitionality or agency. This is a consequence of the meaning of *-le*, which does not have Agentive properties. In Hindi, for which there exist alternations that correlate with volitionality, the ergative and instrumental are not isomorphic. In a language in which the ergative and instrumental have the same form and can be argued to contribute the same meaning, ergativity will not correlate with volitionality but will have many of the other features of transitivity discussed here.

This analysis implies that morphological forms with different nuances of meaning may take part in ergative and accusative case marking systems. “Ergative” and “instrumental” markers will have a different range of meanings in different languages and these meanings may shift over time, take on new properties, and become grammaticalized into paradigms (a topic of the next section). For example, Figure (6.2) overlays the *-le* and *-lāi* Nepali marker on Croft’s conceptual space for participant roles, compared with the *-se*, *-ne* and *-ko* markers of Hindi (Croft 2012: 280).

This explanation covers many of the features we have seen, but it does not cover those that are more problematic from the perspective of the Transitivity Hypothesis: ergative marking is more common on unexpected (inanimate, non-first person) subjects, and it is correlated with discourse prominence, characterizing predicates, and categorical propositions.

Discourse prominence is inherent to markedness in an optional case marking system. In the next section, I will develop a markedness-based account of variable and obligatory case marking systems, taking as an example the use of gendered morphology in English. I apply this to ergative marking to derive the discourse prominence-based feature associations we find with Nepali, and conclude with a discussion of grammaticalization and event structure.

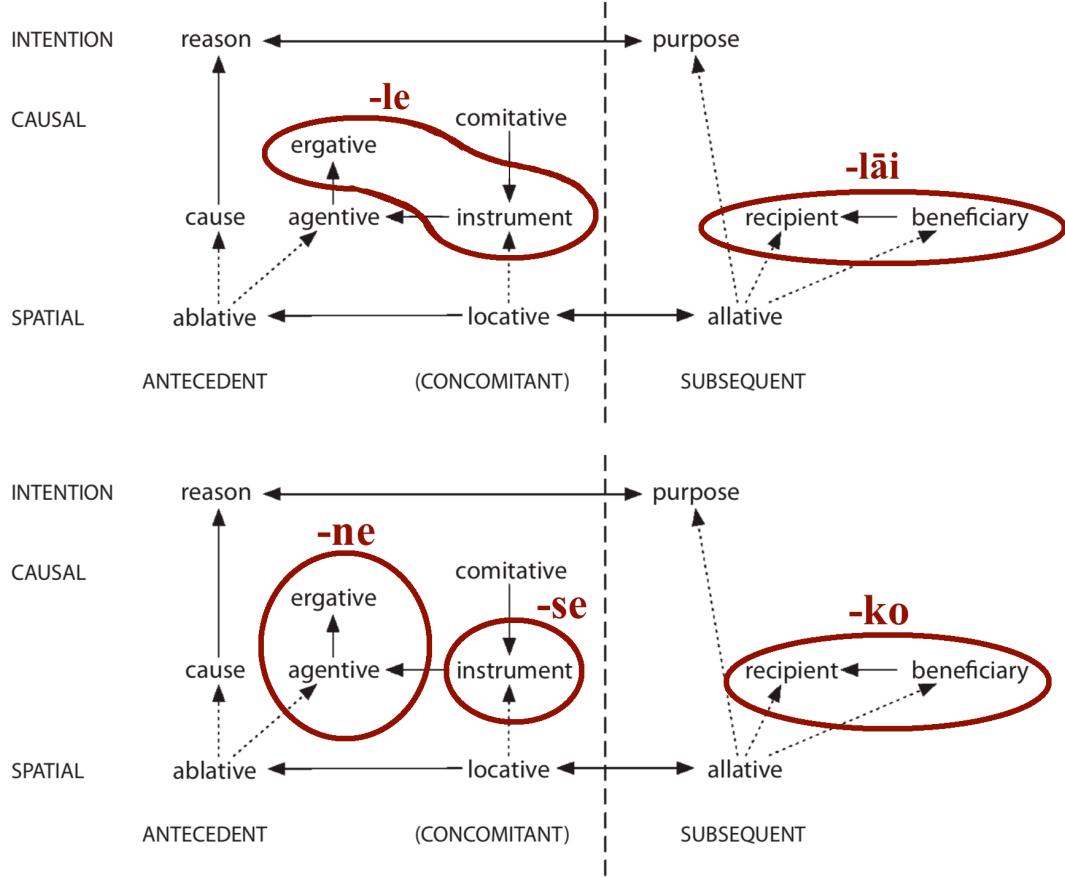


Figure 6.2: Nepali and Hindi Case Markers in Croft’s Conceptual Space for Participant Roles

6.2 On Markedness and Variability

The fundamental argument presented here is that all of the feature correlations relating to the case expression of S_t are attributable to an oppositional asymmetry between the marked ergative form (*-le*) and the unmarked nominative form. In this section I argue that there is a scale of markedness that ranges from inclusive asymmetry to polar opposition. Outside of perfective transitive clauses, the usage of the Nepali ergative is pragmatically conditioned and is associated with gradient tendencies rather than categorical divisions, indicating that the ergative is implicated in an opposition of inclusive asymmetry rather than polar opposition. The distinction between obligatory and optional case marking is due to a grammaticalization of this

opposition between the presence of a marker and its absence.

6.2.1 Opposition with a Zero Form

The clearest cases of semantic markedness are those in which there is an opposition between the presence and absence of a form. In Nepali, this is generally the case for ergative marking of the S_t (-*le*/-∅) and for accusative marking on the O (-*lāi*/-∅), as well as marking of number on noun phrases (-*haru*/-∅). In the simplest case there is an iconic relationship between a marked form and a marked meaning, such that the additional form corresponds with a more restricted meaning. This is **formal markedness**.

Markedness asymmetries that exist between two overt forms will tend to be more complex, because it is rarely the case that two overt forms are truly oppositional. For example, various lexical pairs represent asymmetric oppositions in English (*tall/short*, *big/small*). The first of the pair is considered to be the unmarked because of examples like the following:

- (226) She is five feet seven inches *tall/#short*.

However, “tall” and “short” are not simply opposed, either symmetrically or asymmetrically. The word “short” can also be a description of length in the separate opposition *short/long*. Only “tall” can describe the size of a latte, and only “short” can describe an insufficient amount of cash. While they may represent oppositions in many contexts, a full lexical entry for each of these terms would necessarily include properties which distinguish them beyond whether or not they refer to a height as [+/-DIMINITIVE].

Similarly, Jakobson (1936) takes the instrumental and dative cases in Russian to represent a markedness asymmetry, in which the instrumental is [+PERIPHERAL] and the dative is [-PERIPHERAL], for which “a peripheral case presupposes the presence

of a central point in the context of the utterance, which the peripheral case helps determine” (Jakobson 2011a: 79). These cases presumably have distinct regular and idiosyncratic meanings beyond whether or not the case-marked argument is peripheral.

Oppositions between zero forms can also shift in meaning over time, and when they take on additional meanings they become less symmetrical. This is true for many gendered lexical items English, in which the historical opposition may simply have been between a zero form and a feminine marker, but over time the full feminine lexical item takes on divergent and often negative connotations: hence *governor/governess*, *master/mistress*, *fox/vixen*.

The ergative/nominative case alternation in Nepali is largely an alternation between a case-marker and a zero-form (excepting a small number of oblique forms). This makes it particularly amenable to an analysis based on semantic markedness. We can make the assumption that fundamentally the only meanings at issue are those that are contributed by the one morphological form, the postposition (*-le*).

More broadly, this analysis will be applicable to any language with optional ergative marking, as defined by McGregor (2010) as a system in which the ergative marker alternates with its absence. The issues brought up here will have relevance for any optional case marking system, and to ergative-marking patterning in general.

6.2.2 Obligatory Marking and Grammatical Context

In applying markedness theory to the semantics and pragmatics of argument realization, it is important to distinguish between obligatory, grammaticalized oppositions on the one hand and variable, pragmatically-conditioned oppositions on the other. This will be represented as a gradient scale of markedness that corresponds to the extent to which the two forms are grammaticalized as a symmetric opposition. An example of obligatory opposition is the grammatical category of number in English

noun phrases.

- (227) a. I baked the **biscuit**.
b. I baked the **biscuit-s**.

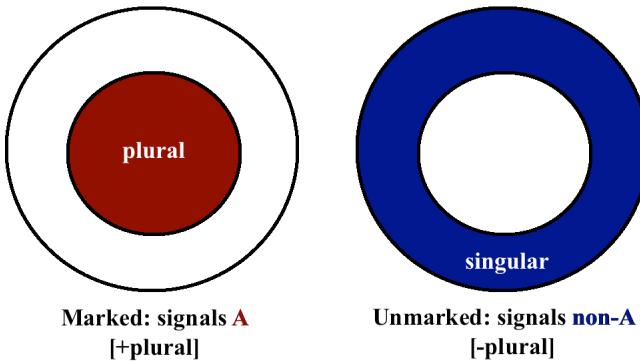
In (227a), the singular form of ‘biscuit,’ a zero-form with no overt number marking, entails that the denoted entity is singular. In (227b), the plural marker (-s) has the effect of entailing that the denoted entity is a plurality. I cannot say “I baked the biscuit” to refer specifically to the act of baking several biscuits, and I cannot say “I baked the biscuits” to refer specifically to the act of baking a single biscuit. In other words, plural marking is obligatory in English. Encoded in the grammar is a requirement that a (count) noun phrase be marked to specify whether it is singular or plural.³

Thus in this grammatical context the only available interpretation is the minus interpretation. The presence of (-s) on an NP signals the property [+PLURAL], and the absence of (-s) signals [-PLURAL]. The first half of (6.3) is a visual representation of this pattern.

This contrasts with the Nepali plural marker *-haru*. The Nepali plural marker also signals the property [+PLURAL], but generally its absence says nothing about whether the noun phrase referent is singular or plural. The only interpretation available is the zero-interpretation (depicted in the second part of Figure 6.3). This, ultimately, is what is meant when a grammatical category is described as “optional”: it signals a particular meaning when it is present, but it does not participate in a grammaticalized (polar) opposition with a zero form (Andrews 1990). Its usage is pragmatically conditioned in the sense that whether or not a form is present in a particular clause depends upon the context of the discourse, the conventions of conversational strategy (e.g., Gricean constraints on quantity and relevance), and the whims of the speaker

3. Excepting forms in which the singular and plural are identical, as in ‘sheep’, ‘fish’, etc.

English (-s): Minus Interpretation polar opposition



Nepali (-haru): Zero Interpretation inclusive asymmetry

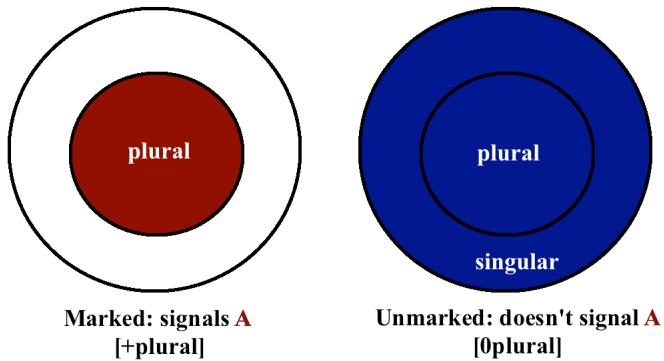


Figure 6.3: Noun Phrase Number Marking in English and Nepali

in choosing which information to present. It is like an adjective in that it simply adds extra information when it is included. However, with an adjective there is even less of a sense in which the form is in an oppositional relationship with its absence. The apprehension of two forms as being in an oppositional relationship, as I discuss below, is a matter of degrees.

- (228) a. *maile yo biskut ban-ā-ē*
I.ERG this biscuit build-CAUS-PERF.1.SG
'I made this biscuit/these biscuits.'
- b. *maile yo biskut-haru ban-ā-ē*
I.ERG this biscuit-PL build-CAUS-PERF.1.SG

‘I made these biscuits.’⁴

- c. *maile yo mit̥o biskut-haru ban-ā-̑*
I.ERG this delicious biscuit-PL build-CAUS-PERF.1.SG
‘I made these delicious biscuits.’

If I were to hand someone a platter of biscuits and say, in English, “I baked this biscuit,” the natural interpretation of number would be a minus interpretation. The recipient of the biscuits would try to make sense of this statement by assuming I am referring to a single one in the group. In Nepali, the statement *maile yo biskut banā̑* could naturally refer to the entire batch.

The primary interpretation for Nepali number-marking on argument noun phrases is a zero-interpretation, therefore number marking is optional. This is also the case for gender-marking in Nepali. And in the imperfective grammatical domain, this is the case for ergative marking on the S_t . For each of these systems, the absence of marking signals nothing.

In English, the minus-interpretation prevails with number-marking on argument noun phrases, gender marking in English pronouns, and in accusative marking on pronominal forms. To the extent that the minus interpretation is required in a particular grammatical context, marking is obligatory. A good example of optional marking in English is derivational morphology that encodes gender (in particular, the ‘-ess’ feminine marker).⁵

The zero-interpretation is apparent from examples like (229a). Both ‘lion’ and ‘lioness’ are possible in a context for which the most likely interpretation is a female referent. This suggests that the word ‘lion’ is unspecified for gender. Even when agreement in gender is required by the grammar, as in (229b), most English speakers

4. There is also variation in plural agreement with the determiner *yo/yi* ‘this’/‘these’.

5. The ‘who’/‘whom’ distinction, for those English speakers who allow both in O position, is another example of an optional opposition in English (suggested to me by Larry Horn). Here the distinction has come to be associated with stylistic register.

will still accept ‘lion.’

- (229) a. Lions/Lionesses give birth to litters of between one and four cubs.
b. The lion/lioness licked her cubs.

However, in other cases with ‘-ess’ the minus-interpretation is more common. This can be problematic for oppositions like ‘actor’ and ‘actress.’ On the one hand, ‘actor’ can be unspecified for gender as in (230a). On the other hand, many speakers of English today would not accept (230b), although this usage is shifting. And there is a ‘slipperiness’ to (230c) (to borrow the term from Waugh 1982), between an interpretation which includes female performers (zero-interpretation) and one which excludes them (minus-interpretation).

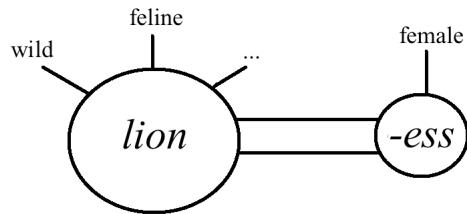
- (230) a. The Screen Actors Guild and its affiliated unions represent nearly 200,000 film actors and other media professionals.
b. ?The actor received universal acclaim for her nuanced portrayal of a computer hacker.
c. Who is the greatest actor of the 1990s?

This is significant, because it illustrates that markedness opposition is only a relevant concept to a particular context. In this case, a significant part of that context is the lexical item to which ‘-ess’ attaches. Acquisition of the lexical items ‘lion’ and ‘actor’ will be slightly different for every speaker of English. They will include a slightly different cluster of properties because they will be based on different experiences, but there will almost always be enough overlap to allow for communication. Another part of the context comes from the discourse, in which some of these properties will be emphasized over others. So gender might not be a particularly salient property for the term ‘lion’, but it might be made more salient in the context of a particular description of lion behavior in a nature documentary.

6.2.3 Gradient Markedness in English Gender Marking

The issue with the ‘actor’/‘actress’ opposition is the extent to which masculinity is an inherently salient property of the lexical item ‘actor’. Or, equivalently, we can describe it as the extent to which *-ess*/∅ are in polar opposition in the context of the word ‘actor’.

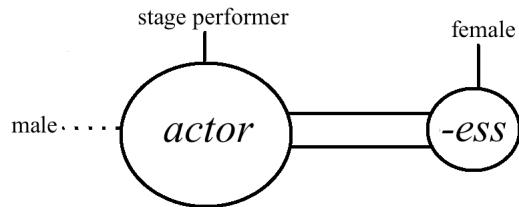
When the zero-interpretation is the only one available, the opposition is one of complete inclusive asymmetry. This is the case when the given context does not carry strong connotations of gender, as with the lexical item ‘lion.’ The diagram below depicts the semantic entailments (the properties connected by solid lines) of the lexical item ‘lion’ and the ‘-ess’ feminine marker (with the double line representing affixation).



The sense in which ‘lion’/‘lioness’ are in opposition is fairly minimal, because the property of masculinity is not a salient property of ‘lion’ (outside of a particular discourse context in which gender is relevant). So we have a situation of inclusive asymmetry, in which the marked form ‘lion’ is unspecified for gender and ‘lioness’ has the exact same meaning with an additional specification of gender. The *-ess* marker in such a situation is highly variable and subject to speaker choice: a speaker might use either ‘lion’ or ‘lioness’ in any particular context. It is very nearly equivalent to a noun phrase with a descriptive adjective, as in ‘female lion.’

The ‘actor’/‘actress’ opposition is more ambiguous between a zero-interpretation and a minus interpretation in which ‘actor’ is considered the masculine form of a pair.

It is also ambiguous because different speakers will differ in the extent to which they consider the terms to be in opposition. This is not because masculinity is an entailed part of the meaning of ‘actor.’ Rather, for cultural and social reasons masculinity is somewhat more salient as a property (represented below with dotted lines indicating an implicature). It is not an entailed part of the meaning to the extent that a sentence like “My favorite actor is female” is possible.

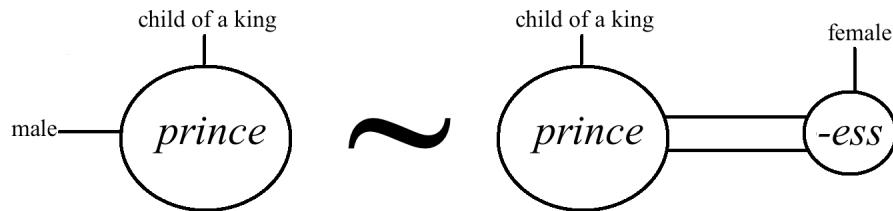


This is a marked context for ‘-ess’ because of the clash between the masculine property of the root and the feminine property of the marker. If the root form provides a marked context, then the marker will be more common, and the balance of the scales will tip from inclusive asymmetry (zero-interpretation) towards polar opposition (minus-interpretation). To the extent that a zero-interpretation is still possible, the marker will still be optional in the sense that both its presence and absence are grammatical in a given clause. As the scale tips towards polar opposition, both the usage and the non-usage of a particular morpheme will be considered meaningful. In this situation, there is a greater likelihood that deviating from a perceived norm will have pragmatic implications. This is why the ‘actor’/‘actress’ opposition is slippery in a way that ‘lion’/‘lioness’ is not.

Thus the term ‘actress’ is marked because of the salience of masculinity as a property in the root form. The masculine property is not entailed in the meaning of ‘actor,’ but for some speakers it is a strong implicature that is closer to a semantic entailment. For these speakers ‘actor’/‘actress’ nearly represents a polar opposition.

This is why many people reject the word ‘actress’ altogether: they wish to avoid the implication of a minus-interpretation (that the prototypical, default actor is male).⁶

Finally, there are some oppositions for which the minus-interpretation is the only one available. This means that there is a simple polar opposition between a masculine/feminine pair. This is the point at which marking becomes obligatory. A good example of this type of opposition is the pair ‘prince’/‘princess.’ For most speakers, the word ‘prince’ must refer to the male form and ‘princess’ to the female form, and the cover term ‘prince’ cannot refer to both.⁷



English lexical items which take the *-ess* marker represent a scale of markedness contexts based upon the extent to which masculinity is a salient property. At the point where the property of masculinity becomes an entailment for the root lexical item, there is a full apprehension of *prince/princess* as an opposition. The pair become part of an obligatory paradigm. This scale of markedness is schematized in Figure (6.4).

This scale represents a synchronic representation of markedness contexts, but it can also describe the process by which an optional opposition shifts and becomes more marked over time until the opposition becomes part of a paradigm. Lehmann (1989) describes such a connection between markedness and grammaticalization in

6. See, for example, Sims (2017) on the push for gender-neutral award categories.

7. There is *some* historical evidence for the usage of ‘prince’ with a zero interpretation. In response to her councillor Robert Cecil telling her that she must go to bed, Queen Elizabeth I is said to have exclaimed, “Must! Is *must* a word to be addressed to princes? Little man, little man! Thy father, if he had been alive, durst not have used that word” (from Knowles 2014, suggested by Larry Horn). But I believe many modern English speakers would reject this gender-neutral usage of ‘prince.’

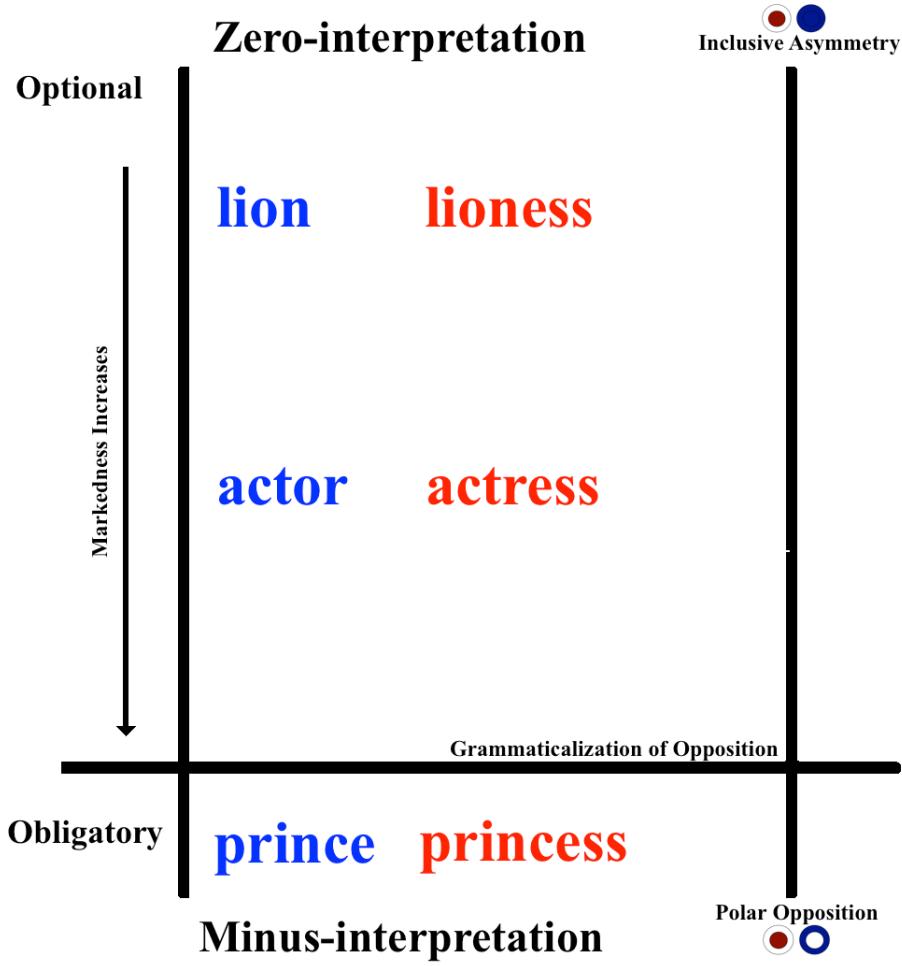


Figure 6.4: Scale of Markedness

the creation of the paradigm of definite/indefinite articles in English, which originally came from non-obligatory markers (the numeral ‘one’ and the demonstrative ‘that’).

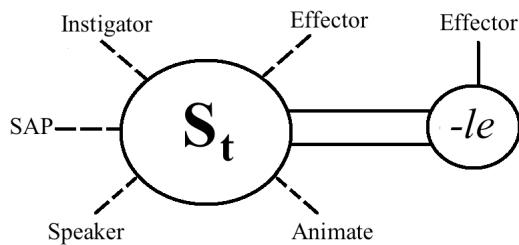
Case marking of core arguments is a different domain of grammar than that of gendered derivational morphology, but the same interplay between markedness and grammaticalization leads to the patterns we find in Nepali case marking.

6.2.4 Gradient Markedness on the Transitive Subject

In the Observation section I concluded that Nepali does not have a categorical split based upon the semantics of the S_t referent. Rather, ergative marking is probabilistically less common on first person pronouns, more common with (non-human) animate

common nouns, and even more common with inanimate common nouns.

Overt arguments in the S_t position are typically local (SAP), particularly first person pronouns. They are overwhelmingly animate. However, none of these generalizations are categorical. Inanimate subjects are rare but possible in Nepali, and the Instigator and Effector properties are part of a markedness prototype. Therefore, they represent implicatures of varying strengths. We can represent the role of *-le* in the grammatical context of an S_t below. In the diagram below, implicatures are represented by dashed lines, and semantic entailments are solid lines:



The speaker has a choice of marking a particular overt transitive subject with *-le* or leaving it unmarked. There are different possible motivations for usage.

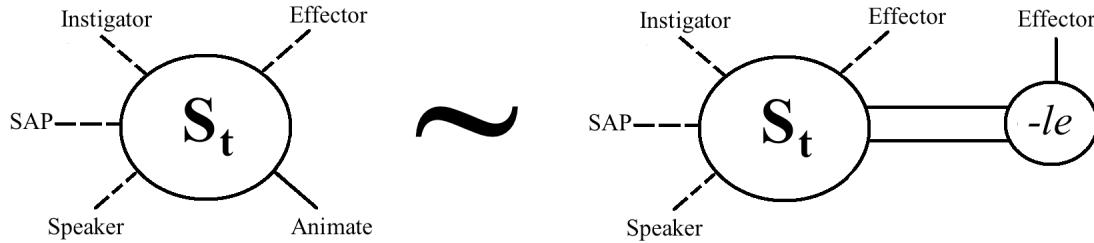
- (1) If the referent of the argument is inanimate, it is in a (semantically) marked position as an S_t , and the speaker will be very likely to use *-le* to distinguish the argument as the effector of the event rather than an affected object. It is not crucial that the argument is disambiguated as the S_t in particular, but merely that it is on the effecting side of a force-dynamic event rather than on the affected side. This is why some inanimate arguments are seemingly ambiguous between subjects and instrumentals.⁸
- (2) Similarly, if the subject referent has an unexpected role in the sentence (as in “The mouse ate the cat”), the usage of *-le* is more likely. It is one strategy a

8. “This explains the extremely common syncretism of ergative or passive agent case with instrumental; the case-form does not refer to agentivity (on which supposition its use for non-agentive instruments is anomalous), but rather to activity in the initial phase of the event, which notion is equally applicable to agents and instruments” (DeLancey 1981: 634)

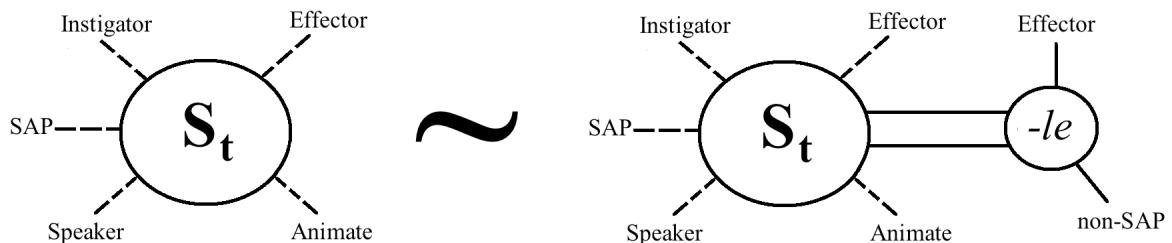
speaker will use to disambiguate the roles of each participant.

- (3) If the referent of the argument is a speech act participant, and particularly if it is the speaker, then it is in a (semantically) unmarked position as an S_t . The speaker will have no need to distinguish it from an affected object, which would typically be case-marked if it is an SAP. In most cases the speaker will not use *-le*. As with *lion/lioness*, the usage of the marked element simply provides extra information by entailing that the entity is an effector of the event. Hence the intuition that “I am building a house” with an ergative-marked subject implies that the house is being built by and for the speaker.
- (4) The usage of *-le* creates a marked form which is inherently more prominent in discourse. It may also be used to imply a particular interpretation of the event. I discuss this in the next sections.

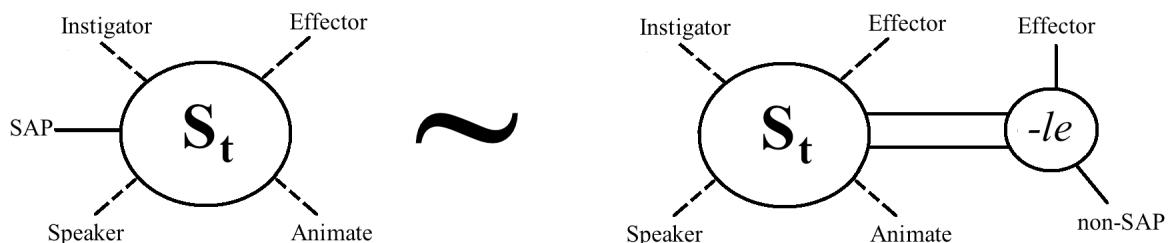
Let us assume a slightly different system in which inanimate subjects are obligatorily ergative-marked in Nepali. This may in fact be true for some speakers of Nepali today, or it may become true in the future. This would imply the grammaticalization of an opposition between animate nouns (for which the ergative is variable with the nominative) and inanimate nouns (for which the ergative is obligatory). The mechanism involved would be the implicature that an S_t is animate, which would become an entailment in the nominative form. Thus, a paradigm would be created in which the nominative form is in opposition with the ergative form. The variability still exists as before, but only for subjects with animate reference. The only form available to subjects with inanimate reference is the ergative.



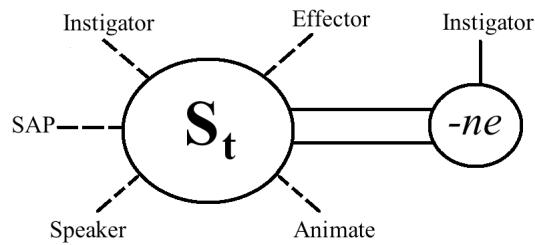
Alternatively, we could imagine a system in which local pronouns cannot be ergative-marked in Nepali. There would be a categorical split between local pronouns and other types of nouns. This would be somewhat like the Marathi system, except there would be ergative/nominative alternations outside of the local domain. This requires an alteration to the core semantic meaning of the ergative marker. Specifically, there would be an extra entailment that would exclude *-le* on local pronouns:



If a system emerges such that ergative marking is obligatory in one domain (non-local pronouns) and nominative marking is obligatory in elsewhere, then there is an entailment on both the nominative form (that it must be a local pronoun) and on the ergative form (that it cannot), which separates their usage into different categorical domains:



Finally, let us return to a system in which there are no categorical splits based on the semantics of the noun phrase. In Hindi and Tsova-Tush we find pragmatic alternations associated with Agentive properties (Volitionality and Control, but following Fauconnier's terms I tentatively subsume these under the Instigator proto-property). This indicates that the ergative marker in these languages has an Instigator entailment. The properties emphasized by choosing to mark a subject with the ergative are those that pertain to an Agent instigating an event.



Volitional alternations that are found in one language but not another are due to the precise semantic contribution of the subject marker. The Hindi instrumental is not homophonous with the ergative, and so there is no synchronic evidence to the learner that there should be an affinity in meaning. It would be a useful typological study to compare languages with ergative/instrumental syncretisms to those without and see whether they are less likely to participate in alternations associated with volitionality.

Differential object marking in Nepali and other Indo-Aryan languages may be worked out in a similar way. Here I will briefly sketch the basics of such an analysis. The relevant fact about the *-lāi* accusative marker (as well as Hindi *-ko*) is that it is homophonous with the dative marker, which obligatorily marks the goal in a ditransitive construction. The core meaning of *-lāi* is that the marked element is a Recipient of the action, and there is a strong implicature in Nepali that the marked element is animate. The typical object is a common noun (animate or inanimate), and it has the general property of being Affected by the event. The marker will be

very common on Proper Nouns and Pronouns and uncommon on inanimate nouns. In Hindi the connection is an entailment: the nominative object must be a common noun.

These analyses does not make reference to either Neutralization or Markedness Reversal, because markedness oppositions are defined solely within a given context. Markedness arises from the interaction of prototype properties in a local environment. There is no inherent reason for marking on the S_t and O to behave in opposite ways, except that some (but not all) relevant properties of the S_t prototype have opposing values to those of the O prototype.

6.3 Variable Ergativity and Prominence

An optional case marking system is one in which a case marker alternates with its absence. In such a system, it will always be the case that the marked form has an increased salience in the discourse. The speaker chooses to draw attention to this element or some aspect of its properties with the addition of a morphological form. This choice may have different pragmatic motivations, but no matter the intention and interpretation, one of the effects will be that the given element is more prominent. This is ultimately a feature of markedness: a morphologically marked form is inherently prominent in a variable system.

Grierson (1904a), Clark (1963) and others have noted that the Nepali ergative form is more “emphatic.” McGregor (2010)’s crosslinguistic analysis of optional ergative marking associates usage of the ergative form with prominent subjects and usage of the unmarked nominative (in his terms, non-usage of the ergative form) with backgrounded subjects. Prominence is associated with different properties in different languages, but these may include contrastive focus, unexpected subjects, and subjects that are high in agency or potency.

I believe that agency/potency is a separate issue that depends upon the semantics of the given marker (as discussed in the previous section). The other two features are inherent to any optional case marking system: the case-marked form of an argument will be more likely in a marked context, i.e., when the argument is unexpected. Contrastive focus (or, rather, contrastive topic) is a consequence of marking the subject as a particular entity in the discourse. This is Kuroda's categorical proposition, in which a particular entity is apprehended, and then something is predicated of that entity. Ultimately it is a consequence of morphological markedness.

This does not mean that an ergative-marked element must be the subject of a categorical proposition, or be the topic, or have contrastive focus, or be unexpected. Being pragmatically-conditioned, the usage of the ergative marker will have multiple motivations, but all of them draw attention to the marked element. Prominence is a matter of degrees, and different elements may be made salient or backgrounded in the same clause. An overt argument is already somewhat prominent in a language for which both subjects and objects may be elided, and adding ergative marking increases its prominence. Usage of the contrastive topic marker *cāhī* increases it even more and is the preferred method for introducing a new topic.

Discourse prominence will be a relevant factor in any optional case marking system. In differential object marking systems as well, accusative case will tend to correlate with topicality of the object, as well as definiteness. Aissen (2003) predicts accusative marking to be associated with animacy and definiteness, and conversely for ergative marking to be associated with inanimacy and indefiniteness of the referent. In fact what we find is that both accusative marking and ergative marking correlate with definiteness. Definiteness is not an inherent property of accusative or ergative marking, but is rather a feature of markedness. Outside of the domain of S_t and O marking, plural marking will also be more common on definite referents.

In the realm of Nepali ergativity, we have seen several feature correlations that

have to do with definiteness, but none of them are categorical (with the possible exception of the quantifier *dheraijaso* requiring an ergative marker). There is the intuition that the ergative form is “picking out” a particular entity from a group, and the affinity for the ergative with strongly construed sets, particularly with ambiguous quantifiers. These are elements that tend to already be salient in the discourse. Furthermore, we do not generally find these correlations showing up in survey judgments. Perhaps this is because they are highly dependent upon the particular context and the extent to which the given arguments are already salient in the discourse.

One feature which requires further explanation is the correlation between subject and word order. Ergative-marked subjects are often found post-verbally in what Butt and King (1996) deem the position for backgrounded information. While arguments in these positions tend to contain old information rather than new information (i.e., are definite), they are also less prominent. This is unexpected coming from the position that ergative-marking is associated with prominence. More corpus work would be useful to determine exactly how common ergative-marked subjects are in each of these positions. Additionally, more research is needed to precisely determine the relation between word order and discourse structure in Nepali. Perhaps the subject isn’t backgrounded so much as the predicate is more salient in the discourse. This would be related to another curious fact about variable ergativity in Nepali, the correlation with individual-level predicates.

6.3.1 Predicate Prominence

Discourse prominence also affects the interpretation of the event. Butt and Poudel (2007) note a correlation between individual-level predicates and ergative marking in the simple present tense. I have argued that this is actually an effect of the tendency for ergative-marked subjects to be associated with categorical propositions. In Brentano’s double judgment, the first judgment consists of apprehending and affirm-

ing the existence of an entity, and the second judgment consists of predicating of it a particular property. A categorical judgment will be used on generic statements with subjects that are kinds, as in *kukur-le māsu khān-cha* “Dogs eat meat,” and we find that the ergative is more common on such statements. A categorical proposition is also generally the only possible response to a question about a particular entity rather than a state of affairs, which is why we find the tendency for a question with an ergative to have a response with an ergative.

I have also noted the tendency for the ergative to correlate with a characterizing reading of a predicate. Thus while *ma curot khān-chu* “I smoke” with a nominative subject may be a statement about an occasional practice, with an ergative subject the interpretation may be that smoking is a habit or an addiction. The predicate is construed as individual-level either way, but in the second case the predicate has more of an inherent connection to the subject. Similarly, the ergative may be associated with an occupation or some other quality that is inherent to the speaker.⁹

We find these interpretations particularly in the simple present verb form, for which the event is ambiguous between stage-level and individual-level, present-oriented and future-oriented. The interpretations are not categorical, and they can be cancelled by other considerations.

Predicate prominence appears to be another effect of discourse prominence on the subject. We expect categorical propositions to correlate with individual-level predicates and definiteness (Ladusaw 2000). In the simple present, this distinction may be moving in the direction of becoming operationalized as an aspectual distinction encoded by noun and verb morphology.¹⁰

9. For a different verb form, the continuous, we might also include here the intuition with *maile ghar ban-āu-dai-chu* ‘I am building a house’, where the ergative might be used to indicate that the speaker is building a house for her own usage.

10. This is already the case with the perfective verb forms requiring ergative marking on the subject, which I discuss in the next section.

In the simple present tense, one interpretation of the ergative is that the event expressed is individual-level, distinguishing a habitual from an ongoing reading of the event. This goes against a prediction of Hopper and Thompson (1980)'s Transitivity Hypothesis, because the prototypical transitive event is punctual rather than durative.¹¹ Indeed, this interpretation apparently conflicts with another possible interpretation that straightforwardly accords with the Transitivity Hypothesis:

- (231) Nominative correlates with a future/ongoing reading; ergative correlates with a habitual reading:

- a. *ma kām gar-chu*
I work do-PRES.1.SG
'I am doing work.' / 'I will do work.' [SB]
- b. *maile kām gar-chu*
I work do-PRES.1.SG
'I do work.' [SB]

- (232) Nominative correlates with a general activity; ergative correlates with a specific activity:

- a. *ma kām gar-chu*
I work do-PRES.1.SG
'I do work.' / 'I work.' [SB]
- b. *maile kām gar-chu*
I work do-PRES.1.SG
'I am doing the/a job.' [SB]

When asked about the difference between a nominative/ergative alternation with a simple present tense transitive sentence like those above, some consultants associated the habitual reading with the ergative, and others associated the habitual reading

11. Although Croft, following Hopper and Thompson, notes that the prototypical punctual event has a durative counterpart, which is "not as prototypical but is widely treated as another instance of the transitive event prototype" (Croft 2012: 355).

with the nominative. This is because of two competing possible interpretations of marking on the subject.

In (231b), the ergative-marked subject is more prominent. This is particularly the case with a first person pronoun, for which ergative marking is the least likely because it is the most frequent form in S_t position. The predicate is interpreted as more inherent to the subject, hence a habitual interpretation. This is Butt and Poudel (2007)'s observation.

Another possible interpretation does not relate to discourse prominence in itself but rather to the meaning of *-le* as an effector and as part of the ergative paradigm. In (232b) the event is construed as being more prototypically transitive, with a highly individuated object. This favors an ongoing interpretation rather than a habitual one.

The fact that these conflicting interpretations are both possible suggests that habituality is epiphenomenal, and it illustrates the pragmatic nature of variable ergative marking. The speaker may either be emphasizing the subject and its relation to the predicate, or the speaker may be emphasizing the transitivity of the event.

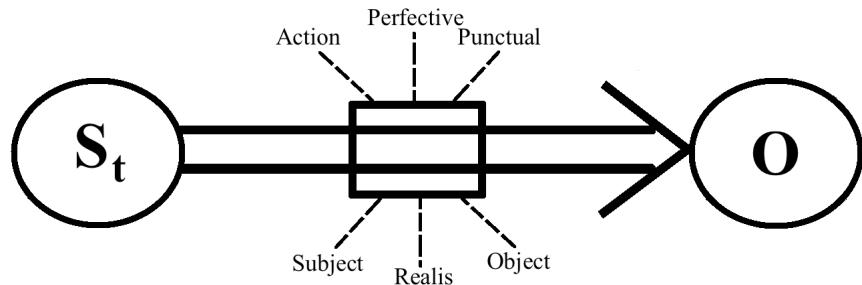
In the next subsection I discuss more generally the effect of ergative marking on event structure.

6.4 Ergativity and Event Structure

Most of the feature correlations that relate to event structure follow the predictions of Hopper and Thompson (1980)'s Transitivity Hypothesis. In general, these feature correlations are pragmatic rather than semantic; they do not represent categorical splits in the language. Rather, they are general tendencies. Hopper and Thompson caution that their hypothesis is only applicable to obligatory feature correlations, but with the exception of individual-level predication (discussed above), their predictions do in fact accord with these observations:

- (1) That ergative marking is restricted to transitives (following Hopper and Thompson's **Participant** property)
- (2) The preference for the ergative on the Definite Future over the Hypothetical Future (following Hopper and Thompson's **Mode**)
- (3) That ergative marking is obligatory with perfective transitives (following Hopper and Thompson's **Aspect**)
- (4) That ergativity is obligatory in the perfective main clause but optional in all subordinate clauses (also following **Mode**)

Below, the overall event structure of a transitive clause is represented with Hopper and Thompson's prototypical properties for transitive events. The double arrow represents the transmission of force from the subject to the object.¹² The square in the center is the verb and whatever morphological material is associated with the event described by the clause. The 'Subject' and 'Object' implicatures taken together represent requirements on the presence of S_t and O (the Participants property).



Ergative marking crosslinguistically correlates with these properties of a prototypical transitive event (particularly the correlation between ergative marking and perfectivity, which I discuss below). This is unexpected under the assumption that

12. This can be understood as a simplified version of Croft (2012)'s three-dimensional representation of causal and aspectual structure (Croft 2012: 212). The arguments are arranged from left to right representing the transmission of force. In Croft's formulation the overall event is structured into separate subevents for each participant, whereas I simplify the structure by associating the entire event with the morphological structure of the verb and state implicatures and entailments rather than representing them diagrammatically.

markedness is associated with deviations from a prototype. If the ergative marks *less* prototypical subjects (such as inanimates), then why should it mark *more* prototypical events?

Rather than seeking a deviation from the prototype, it is important to consider the semantic meaning contributed by the ergative and possible motivations for its usage. With an inanimate subject the preference for the ergative comes from the necessity to discriminate the argument as an effector of the event. The S_t position is semantically marked for inanimates. Compare this with a hypothetical correlation between ergative marking and kinesis: ergative marking is more common on action verbs like “eat” and less common on non-action verbs like “watch.”

- (233) a. The dog ate the rat.
b. The dog watched the rat.

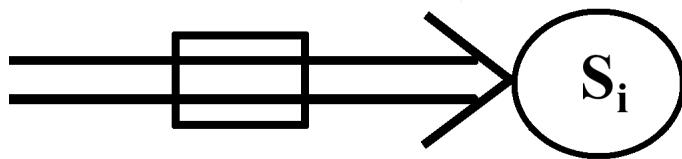
In both cases the ergative will be helpful to distinguish the S_t , and will be more common if the S_t referent is unexpected (as in “The rat ate the dog”). Another usage of the ergative may be to emphasize the dog’s role in effecting the event (or, in some languages, its volitionality/agency). The watching event is a deviation from the transitive prototype, while the eating event is more prototypical. But there is not a motivation to specify that the dog is an effector simply because the event is not a prototypical one. If the speaker wishes to emphasize the dog’s role in effecting the watching event, they can do so with the ergative, but this will be in opposition to other watching events, not to events in general. Furthermore, the motivation to specify the dog as an effector via ergative marking (or to imply its lack of effectiveness with the nominative) will be much more common with action verbs than with non-action verbs. The distinction will tend to be more relevant. If ergative marking comes to be used very commonly with action verbs, then there might be a grammaticalized entailment between ergative marking and the category of action verbs. In fact, entailments can occur between any of these related properties in the transitivity prototype, and this

forms the basis of the Transitivity Hypothesis.

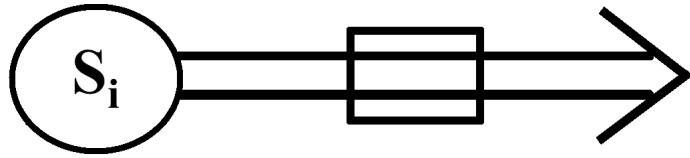
Different properties are grammaticalized as entailments in different languages. There are two domains which in Nepali represent categorical splits in ergative-nominative case marking. Ergative marking is disallowed in copular clauses and unaccusative intransitives. I have argued that most if not all reported instances of S_i marking with unergative intransitives should be analyzed as S_t marking on a transitive clause with an elided object. Thus, ergative marking is restricted to the S_t position in **Transitive clauses**. Secondly, ergative marking is obligatory in transitive clauses with **Perfective** verb forms, and variable elsewhere.

6.4.1 Ergative Marking in Transitive Clauses

If we define *-le* as an ergative marker, then its inability to appear in intransitives or copular clauses is straightforward. By its definition, it will be restricted to the S_t position of a transitive clause. If the contribution of this marker is, more broadly, as an Effector (or as an Instigator, or both), we might then expect an active-stative alignment to be possible. Its usage is expanded to include an S_i that effects the given event (i.e., the subject of unergative intransitives). It will be unavailable with unaccusative clauses like “The ghee melted” because the S_i here is being affected by the event rather than effecting it:



Whereas it may be available with unergative clauses like “I walked,” in which the lone participant does effect the event:

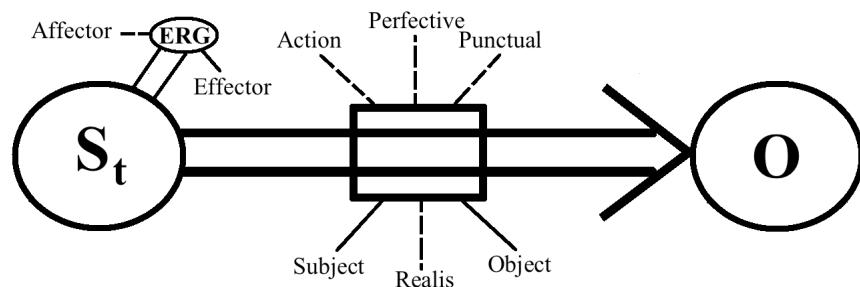


An instrumental argument is generally possible with intransitives (hence “I walked with a walking stick”). So if *-le* contributes the same meaning as it does in an instrumental, we might expect an active-stative alignment. This relates to the precise meaning of Effector, and my preference of this term over Fauconnier’s prototypical S_t property Affector. The question is whether it indicates an affect upon another participant, or whether it simply relates to the maintenance (and perhaps completion) of the event itself.

It is not entirely clear whether ergativity has an effect on the interpretation of the object in transitive clauses, as I discussed in the Observations section on ergative marking and the object. For example, it does not appear to correlate with an object being completely affected (e.g. in “I tipped over the chair”), nor does the nominative correlate with the object being partially affected. However, I have noted that the ergative may be correlated with a highly individuated object, as in *maile kām gar-chu* “I am doing a job.”

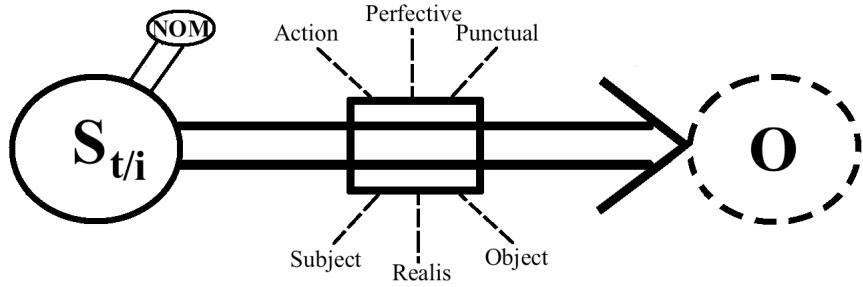
More significantly, in the Observations section on intransitives I made the case that *-le* is in fact restricted to S_t . Verbs like *khelnu* ‘to play’ and *gāunu* ‘to sing’ are typically transitive whether or not the object is overtly realized. For some speakers, it is possible to construe *khelnu* as intransitive, which explains the possibility of a nominative form in *ma khelē* “I played.” Unergative intransitives like *hidnu* ‘to walk’ or *vācnu* ‘to live’ may be construed as transitive if there is an implied object, which explains the possibility of an ergative form with *usle jiwān rāmrari vāc-yo* “He lived his life well.”

Thus *-le* takes on additional qualities when it is incorporated into the ergative pattern. These include both the effecting of an event and the affecting of an object, following the expected pattern of high transitivity. This apparently rises to the level of a semantic restriction against *-le* on S_i . If this is in fact a categorical distinction, then we must distinguish between (1) the meaning entailed by the postposition *-le* as an effector from (2) the meaning entailed by the ergative (S_t -*le*) in a transitive clause. The generalization given below relates specifically to ergative case:



If an S_t is case-marked ergative, there is an entailment (contributed by *-le*) that it is an effector of the event. There is also an implicature that the object is affected by the event.¹³ The verbal morphology entails the event properties relating to event structure. In particular, the lexical semantics of the verb root tell us whether there is an (overt or covert) subject and object. Ergative marking is only possible in a transitive clause, which is depicted diagrammatically above by the entailments that both a subject and an object must be present. This whole frame is in a paradigmatic opposition with the nominative form depicted below, for which there is no entailment that both a subject and object must be present:

13. This is not an entailment, because ergative marking is possible in transitive clauses in which the object is not affected by the event, e.g. *maile mitho awāj sunē* “I heard a beautiful sound.”



Thus nominative case may be found in both transitive and intransitive events, while ergative case is restricted to transitive events.

6.4.2 The Connection between Ergativity and Perfectivity

The other obligatory feature correlation is between perfectivity and ergativity. This is the most common source of an aspectual split in ergative/nominative patterning, both for Indo-Aryan and for ergative languages in general. Hopper and Thompson (1980) conclude that telicity (under which they include both lexically telic verbs and verbs with perfective morphology) is associated with high transitivity. However, the connection between ergativity and perfectivity is a particularly close association.

On the one hand, there is a very clear diachronic source for this connection. Ergativity arose in Indo-Aryan from the reanalysis of a deverbal construction into a general perfective form, and Nepali inherits ergative morphology in those modern verb forms which have perfective aspectual reference. The viewpoint of Anderson (1977) and Garrett (1990) is that the common association between ergativity and perfective aspect is entirely due to this grammaticalization pathway. It is an accident of history rather than evidence of an inherent semantic correlation.¹⁴

However, it is less clear that ergativity arose in Tibeto-Burman through this diachronic pathway, and yet in many of these languages ergative marking (or agentive

14. “Since ergative morphology is often triggered by perfective aspect in particular, it has been argued that the two have some intrinsic connection or share some inherent feature. However, as Anderson (1988: 340-49) argues, this type of theoretical claim results from a failure to appreciate the diachronic evidence” (Garrett 1990: 262).

marking) is sensitive to perfectivity, even if the correlation is pragmatic rather than semantic for some languages. Meanwhile, Nepali learners inherit a system in which ergative marking is variable in transitive clauses except in perfective clauses where it is obligatory. Where it is variable, ergativity is associated with many different factors relating to the transitivity of the clause.

Hopper and Thompson (1980)'s explanation for perfectivity as a marker of transitivity is based on their theory that transitivity is tied to prominence in discourse structure: perfective events tend to move the discourse forward, while imperfective events tend to be scene-setting. So here the connection between ergativity is theoretically motivated but indirect: ergativity is associated with transitivity, and perfectivity is also associated with transitivity.

A common intuition about the connection between ergativity and perfectivity relates to the observation that perfective clauses profile the result state rather than the initiation of the event. Thus Dixon (1994) notes that the ergative S_t/O grouping is more natural with completed events, because in non-realized events the focus is on the initiator of the action rather than the result, and initiators are S_i or S_t (Dixon 1994: 98-99). Similarly, DeLancey (1981) argues that nominative/accusative patterning profiles the agent and initiation of the event, while ergative/absolutive morphology profiles the patient and its result (the unmarked form is profiled in each case). The perfective profiles the result state (as do passive constructions), while the imperfective profiles the initiation of the event. Languages with perfective ergative-splits do not allow a clash between aspect and case-marking (DeLancey 1981: 646).

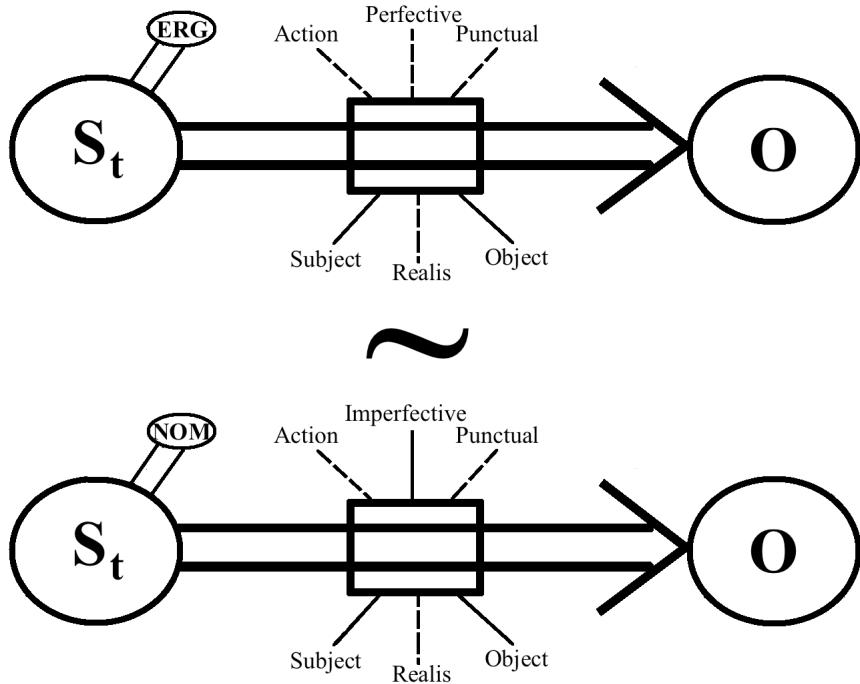
These arguments are intrinsically based on markedness. In the perfective, the subject is marked (both morphologically and semantically) because the result state and the effect on the object is being profiled. Næss (2004) explicitly invokes markedness principles in explaining this connection. She takes the main property of the S_t to be the Controller of the event, and the main property of the O to be Affectedness.

Within an accusative case marking system affectedness is the marked property, and within an ergative case marking system control is the marked property (Næss 2004: 1208). Since an imperfective event cannot be described in terms of its result state, it must be described in terms of its initiation by a controller, thus it takes an accusative alignment (Næss 2004: 1209).

Hopper and Thompson (1980)'s explanation fits the Nepali data better. The pattern is attributed to the inherent meaning of the ergative as a marker of transitivity (i.e. as an effector and/or agentive) rather than attributing it to markedness as deviation from a prototype. It strikes me as counter-intuitive that the ergative should be associated with control and volition and therefore be relegated to event structures in which the object and result state are profiled. Furthermore, the other explanations require that the ergative profile the initiation of the event, whereas the effector analysis specifically excludes the initiation of the event (and includes its completion).

The passive-to-ergative grammaticalization pathway described in Croft (2012: 255-57) is useful for this discussion. The passive voice is characterized by “the de-profiling of the causal segment from agent (initiator) to patient (endpoint)” (Croft 2012: 256). The grammaticalization pathway consists of the gradual restoration of that causal segment. Initially there is a passive form without an overt agent, and then this agent is reintroduced as a marked antecedent oblique (often an instrumental). At this point the causal chain is not profiled, but rather the result state. The process of reanalyzing the passive as an ergative consists of reinterpreting the oblique as an ergative agent and profiling the entire causal chain. This is accompanied by a shift in the usual highest-topicality discourse role: with passives it is the O, and with ergative it is the S_t. Croft notes that this “is also represented as the profiling of the agent subevent and thus its force-dynamic relationship to the patient subevent” (Croft 2012: 257). Thus the ergative profiles the entire causal chain rather than just the result state.

Regardless of the scenario we accept, in Nepali we find that ergative case is obligatory in perfective transitive clauses and variable elsewhere.



This is operationalized as an opposition between nominative and ergative case in transitive clauses (diagrammed above by the entailed presence of both a subject and object). Ergative marking may be found in both perfective and imperfective clauses, so there are no extra entailments. Nominative marking, however, entails the presence of imperfective aspect. This follows the Transitivity Hypothesis, because a marker of low transitivity (nominative case) is associated with another marker of low transitivity (verbal morphology with imperfective aspect).

Chapter 7

Conclusions

Nepali presents with a complex case marking pattern in which ergative case is obligatory in perfective transitive clauses, disallowed in intransitive clauses and copular clauses, and varies with the nominative elsewhere. Where ergative marking is variable, its usage correlates with properties of the Event, properties of the Subject, properties of the Object, and properties of the Discourse.

The study of pragmatic phenomena require the implementation of multiple strategies for collecting language data. The data for this investigation come from four converging lines of inquiry: descriptions of the Nepali pattern in the literature, targeted elicitations with native speakers, the implementation of a grammaticality judgment survey, and the analysis of a sample of the Nepali National Spoken Corpus.

I found ergative marking to be obligatory in perfective main clauses and variable in subordinate clauses. What appears to be active marking in intransitive clauses is analyzed as ergative marking in transitive clauses with covert objects. The only categorical split is the distinction between perfective and non-perfective verb forms. Every other association was found to be non-categorical.

These non-categorical associations include a positive correlation between subjects with inanimate reference and the expression of ergativity in common nouns, and a

negative correlation between first person pronouns and ergativity in the pronominal domain. This follows expected patterns of marking based on the types which are most frequent in discourse. Ergative marking is somewhat associated with highly individuated objects, but not with affected objects.

Ergative marking is positively associated with characterizing or individual-level predicates, kind readings, categorical propositions, and strong construals of quantifiers. There was no correlation found between ergative marking and agency or volitionality.

The main claim of this analysis is that the Nepali ergative marks an effector of the event described by the clause. This term refers to a participant which is implicated in enacting and effecting the event, but is not necessarily its main controller or instigator. As a component of the ergative case marking system, it has a pragmatic usage, implicating the subject as a participant in a prototypically transitive event. Aspects of this analysis contribute to the general theory of Optional Ergative Marking and its relation to argument proto-roles. Associations between the ergative and prototypical properties of a transitive event arise from the meaning of the ergative marker as an effector. This analysis also provides a straightforward explanation for the lack of volitional correlations in Nepali that we find in other languages with variable ergativity.

The other semantic and pragmatic features are associated with discourse prominence. These include the correlation with categorical propositions and characterizing predicates. Here the associations are attributable to general principles of semantic markedness. Variable ergativity represents the presence of pragmatic implicatures of various strengths. Gradient markedness oppositions can lead to the conventionalization of these associations into semantic entailments. This is demonstrated for English gender marking, the association between ergative marking and semantic properties of the transitive subject in Nepali, and the association between ergative marking and

Nepali perfective verb forms.

These conclusions present several possible avenues for future research. The first would simply involve the expansion of the data set by analyzing a larger portion of the NNSP and conducting more grammaticality judgment surveys on a larger pool speakers, which could be disseminated through an online marketplace. In this way, it would be possible to obtain enough data to make more confident statements about the comparative strength of various feature correlations. It would also illuminate the ways that this system varies alongside other kinds of dialectal variation. One particular point of interest would be to examine particular variants of this system and see if they correlate with properties of the speaker's L1 or parents' L1.

In that same vein, a major unanswered question on the problem is the extent to which Nepali variable ergativity arises from language contact with other languages of Nepal. Depending upon when variable ergativity first entered the system, we might expect Magar or Newari to have had a greater impact on the language. It would be helpful to use the same approach described here to analyze languages which have had a long history of contact with Nepali.

Finally, it would be useful to compare the Nepali system in depth to as many languages with variable ergativity as possible to uncover other patterns in the associations between argument realization and semantic/pragmatic features.

Bibliography

- Abadie, P. (1974). Nepali as an ergative language. *Linguistics of the Tibeto-Burman Area*, 1(1):156–177.
- Acharya, J. (1991). *A descriptive grammar of Nepali and an analyzed corpus*. Georgetown University Press.
- Ahearn, L. M. (2001a). *Invitations to love: Literacy, love letters, and social change in Nepal*. University of Michigan Press.
- Ahearn, L. M. (2001b). Language and agency. *Annual Review of Anthropology*, 30(1):109–137.
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21(3):435–483.
- Anderson, S. R. (1977). On mechanisms by which languages become ergative. *Mechanisms of syntactic change*, 2:317–364.
- Anderson, S. R. (1988). Morphological change. *Linguistics: the Cambridge survey*, 1:324–362.
- Andrews, E. (1990). *Markedness Theory: The Union of Asymmetry and Semiosis in Language*. Duke University Press.
- Baker, M. (2015). *Case*. Cambridge University Press.
- Baker, M. and Bobaljik, J. (2017). On inherent and dependent theories of ergative case. *The Oxford handbook of ergativity*, pages 111–134.
- Battistella, E. L. (1996). *The logic of markedness*. Oxford University Press.
- Begum, M. R. (2017). *Developing a Pilot Hindi Treebank Based on Computational Paninian Grammar*. PhD thesis, International Institute of Information Technology, Hyderabad.
- Bendix, E. H. (1974). Indo-Aryan and Tibeto-Burman contact as seen through Nepali and Newari verb tenses. *International journal of Dravidian linguistics*, 3(1):42–59.

- Bharati, A., Chaitanya, V., Sangal, R., and Ramakrishnamacharyulu, K. (1995). *Natural language processing: a Paninian perspective*. Prentice-Hall of India New Delhi.
- Bickel, B. (2011). Grammatical relations typology. In Song, J. J., editor, *The Oxford handbook of language typology*, pages 399—445. Oxford University Press.
- Bobaljik, J. (1993). On ergativity and ergative unergatives. *MIT working papers in linguistics*, 19(4588):334385.
- Brentano, F. C. (1874). *Psychologie vom empirischen Standpunkt*, volume 1. Duncker & Humblot.
- Bresnan, J., Cueni, A., Nikitina, T., and Baayen, R. H. (2007). Predicting the dative alternation. In *Cognitive foundations of interpretation*, pages 69–94. KNAW.
- Büring, D. (2003). On d-trees, beans, and b-accents. *Linguistics and philosophy*, 26(5):511–545.
- Burton, S. and Matthewson, L. (2015). Targeted construction storyboards in semantic fieldwork. *Methodologies in semantic fieldwork*, pages 135–156.
- Butt, M. (2001). A reexamination of the accusative to ergative shift in Indo-Aryan. *Time over matter: Diachronic perspectives on morphosyntax*, pages 105–141.
- Butt, M. and King, T. H. (1996). Structural topic and focus without movement. In *Proceedings of the First LFG Conference*. Stanford: CSLI Publications.
- Butt, M. and Poudel, T. (2007). Distribution of the ergative in nepali. *Manuscript, University of Konstanz*.
- Carlson, G. N. (1977a). A unified analysis of the English bare plural. *Linguistics and philosophy*, 1(3):413–457.
- Carlson, G. N. (1977b). *Reference to kinds in English*. PhD thesis, University of Massachusetts.
- Chafe, W. L. (1976). *Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Points of View*, pages 25–56. New York: Academic Press.
- Chelliah, S. L., Hyslop, G., et al. (2011). Introduction to special issue on optional case marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area*, 34(2):1.
- Chierchia, G. (1995). *Individual-Level Predicates as Inherent Generics*, chapter 3, pages 176–223. University of Chicago Press.
- Chvany, C. V. (1984). From Jakobson's cube as objet d'art to a new model of the grammatical sign. *International journal of Slavic linguistics and poetics*, 29:43–70.

- Clark, T. W. (1963). *Introduction to Nepali: a first-year language course*. School of Oriental and African studies, University of London.
- Comrie, B. (1978). Ergativity. In Lehman, W., editor, *Syntactic typology: studies in the phenomenology of language*, pages 329–394. University of Texas Press.
- Comrie, B. (2013a). Alignment of case marking of full noun phrases. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Comrie, B. (2013b). Alignment of case marking of pronouns. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Coon, J. (2013). *Aspects of split ergativity*. Oxford University Press on Demand.
- Croft, W. (1991). *Syntactic categories and grammatical relations: the cognitive organization of information*. Chicago: University of Chicago Press.
- Croft, W. (2003). Typology and universals. (2nd edition). Cambridge: Cambridge University Press.
- Croft, W. (2012). *Verbs: Aspect and causal structure*. OUP Oxford.
- De Beauvoir, S. and Parshley, H. M. (1953). *The second sex*. Vintage books New York.
- De Lacy, P. (2006). *Markedness: Reduction and preservation in phonology*, volume 112. Cambridge University Press.
- DeLancey, S. (1981). An interpretation of split ergativity and related patterns. *Language*, pages 626–657.
- DeLancey, S. (2004). The blue bird of ergativity. *Ergativity in Amazonia*, 3:1–37.
- DeLancey, S. (2011). ‘Optional’‘Ergativity’in Tibeto-Burman Languages. *Linguistics of the Tibeto-Burman Area*, 34(2):9–20.
- Deo, A. (2017). Marathi tense marking: a window into the lexical encoding of tense meanings. In *Penn Linguistics Conference 41*.
- Deo, A. and Sharma, D. (2006). Typological variation in the ergative morphology of Indo-Aryan languages.
- Deuba, T. (September 28, 2018). Prov 7 named Sudurpaschim amid objection from NC, RJP. *The Himalayan Times*.
- Diesing, M. (1992). Bare plural subjects and the derivation of logical representations. *Linguistic Inquiry*, 23(3):353–380.

- Dixon, R. M. (1972). *The Dyirbal language of north Queensland*, volume 9. CUP Archive.
- Dixon, R. M. (1979). Ergativity. *Language*, pages 59–138.
- Dixon, R. M. (1994). *Ergativity*. Cambridge University Press.
- Dowty, D. (1991). Thematic proto-roles and argument selection. *language*, pages 547–619.
- Du Bois, J. W. (1987). The discourse basis of ergativity. *Language*, pages 805–855.
- Eagle, S. (2008). The language situation in Nepal. *Language planning and policy*, 1:170–225.
- Einstein, A. and Infeld, L. (1961). *The Evolution of Physics: The Growths of Ideas from Early Concepts to Relativity and Quanta*. Cambridge University Press.
- Esri (2018). Topographic. *World Topographic Map*.
- Fauconnier, S. (2011). Differential agent marking and animacy. *Lingua*, 121(3):533–547.
- Féry, C. and Ishihara, S. (2016). Introduction. In Féry, C. and Ishihara, S., editors, *The Oxford Handbook of Information Structure*, pages 1–18. Oxford University Press.
- Féry, C., Pandey, P., and Kentner, G. (2016). The prosody of focus and givenness in Hindi and Indian English. *Studies in Language. International Journal sponsored by the Foundation “Foundations of Language”*, 40(2):302–339.
- Foley, W. A. (1986). *The Papuan languages of New Guinea*. Cambridge University Press.
- Garrett, A. (1990). The origin of NP split ergativity. *Language*, pages 261–296.
- Genetti, C. (1988). A syntactic correlate of topicality in Newari narrative. *Clause combining in grammar and discourse*, pages 29–48.
- Genetti, C. E. (1994). *A descriptive and historical account of the Dolakha Newari dialect*. PhD thesis, Inst. for the Study of Languages and Cultures of Asia and Africa, Tokyo Univ. of Foreign Studies Tokyo.
- Green, G. M. (1974). *Semantics and syntactic regularity*. Indiana Univ Pr.
- Grierson, G. A. (1904a). Specimens of Pahari languages and Gujuri. *Linguistic Survey of India*, 9(4).
- Grierson, G. A. (1904b). Specimens of the rajasthani and gujarati. *Linguistic Survey of India*, 9(2).

- Hammarström, H., Forkel, R., and Haspelmath, M. (2018). Glottolog 3.3. Max Planck Institute for the Science of Human History.
- Harris, A. (2019). *The Non-aspectual Meaning of African-American English ‘Aspect’ Markers*. PhD thesis, Yale University.
- Haspelmath, M. (2006). Markedness (And What to Replace It With). *Journal of Linguistics*, 42(1):25–70.
- Holisky, D. A. (1987). The case of the intransitive subject in Tsova-Tush (Batsbi). *Lingua*, 71(1-4):103–132.
- Hopper, P. J. and Thompson, S. A. (1980). Transitivity in grammar and discourse. *Language*, pages 251–299.
- Horn, L. R. (2001). *A Natural History of Negation*, chapter 3. Markedness and the Psychology of Negation, pages 154–192. CSLI Publications.
- Hutt, M. (1988). *Nepali: A national language and its literature*. Sterling Publishers and School of Oriental and African Studies.
- Hutt, M. (1996). Ethnic nationalism, refugees and Bhutan. *Journal of Refugee Studies*, 9(4):397–420.
- Hutt, M. and Subedi, A. (1999). *Teach Yourself Nepali. A complete course in understanding, speaking and writing Nepali*. Hodder Headline.
- Jackendoff, R. (1972). *Semantics in Generative Grammar*. MIT Press, Cambridge, MA.
- Jakobson, R. (1932). *Zur Struktur der Russischen Verbums*. Pražský Linguistický Kroužek.
- Jakobson, R. (1936). Beitrag zur allgemeinen kasuslehre: Gesamtbedeutungen der russischen kasus. *TCLP*, VI.
- Jakobson, R. (1939). Signe zéro. *Mélanges de linguistique offerts à Charles Bally*.
- Jakobson, R. (1957). *Shifters, verbal categories and the Russian verb*. Harvard University, Department of Slavic Languages and Literatures, Russian Language Project.
- Jakobson, R. (2011a). Contribution to the General Theory of Case: General Meanings of the Russian Cases. In Waugh, L. R. and Halle, M., editors, *Russian and Slavic grammar: Studies 1931-1981*, volume 106, chapter 6, pages 59–104. Walter de Gruyter.
- Jakobson, R. (2011b). Zero Sign. In Waugh, L. R. and Halle, M., editors, *Russian and Slavic grammar: Studies 1931-1981*, volume 106, chapter 11, pages 151–160. Walter de Gruyter.

- Jinitha, K. (2009). *Rajasuya, Niranunasika and Subhadraharana of Melpettur - A Grammatical Study*. PhD thesis, Sree Sankaracharya University of Sanskrit, Kalady.
- Keenan, E. L. (1984). Semantic correlates of the ergative/absolutive distinction.
- Keenan, E. L. (1985). Relative clauses. In Shopen, T., editor, *Language typology and syntactic description. Volume 2: Complex constructions*. Cambridge University Press.
- Keidan, A. (2011). The kāraka-vibhakti device as a heuristic tool for the compositional history of pāṇini's astādhyāyī. *Rivista degli studi orientali*, 84(1/4):273–288.
- Knowles, E. M., editor (2014). *Oxford Dictionary of Quotations*. Oxford University Press, eighth edition.
- Kratzer, A. (1995). *Individual-level predicates*, volume 125–175, chapter 2. University of Chicago Press.
- Krifka, M. (1999). Manner in dative alternation. In *West Coast Conference on Formal Linguistics*, volume 18, pages 260–271. Citeseer.
- Krifka, M. (2008). Basic Notions of Information Structure. *Acta Linguistica Hungarica*, 55:243–276.
- Kuroda, S.-Y. (1972). The categorical and the thetic judgment: Evidence from Japanese syntax. *Foundations of language*, pages 153–185.
- Kuroda, S.-Y. (1990). The categorical and the thetic judgment reconsidered. In *Mind, meaning and metaphysics*, pages 77–88. Springer.
- Ladd, R. (1996). *Intonational Phonology*. Bloomington: Indiana University Press.
- Ladusaw, W. (2000). Thetic and categorical, stage and individual, weak and strong. *Negation and Polarity. Syntactic and Semantic Perspectives*, Oxford University Press, Oxford, pages 232–242.
- Laka, I. (1993). Unergatives that Assign Ergative, Unaccusatives that Assign Accusative. *MIT Working Papers in Linguistics*, 18:149–172.
- Laka, I. (2006). On the nature of case in Basque: structural or inherent. *Organizing grammar*, pages 374–382.
- Legate, J. A. (2008). Morphological and abstract case. *Linguistic Inquiry*, 39(1):55–101.
- Legate, J. A. (2012). Types of ergativity. *Lingua*, 122(3):181–191.
- Lehmann, C. (1989). Markedness and grammaticalization. *Markedness in synchrony and diachrony*, pages 175–190.

- Leonetti, M. (1994). ‘Ser’ y ‘estar’: estado de la cuestión. *Barataria*, 1977:182–205.
- Levin, B., Hovav, M. R., and Keyser, S. J. (1995). *Unaccusativity: At the syntax-lexical semantics interface*, volume 26. MIT press.
- Lewis, D. (1975). Adverbs of quantification. *Formal semantics-the essential readings*, pages 178–188.
- Li, C. (2007). Split ergativity and split intransitivity in Nepali. *Lingua*, 117(8):1462–1482.
- Lidz, L. A. et al. (2011). Agentive marking in yongning na (mosuo). *Linguistics of the Tibeto-Burman Area*, 34(2):49.
- Lindemann, L. (2016). Nepali le as a marker of categorical subjecthood. *Formal Approaches to South Asian Languages*, 2.
- MacDonell, A. A. (1927). *A Sanskrit Grammar for students*. Oxford University Press.
- Malla, U. N. (2012). National Population and Housing Census 2011. Technical report, Government of Nepal Central Bureau of Statistics.
- Marantz, A. (1991). Case and licensing. Unpublished Manuscript.
- Marty, A. (1918). *Gesammelte Schriften*, volume II. Halle: Max Niemeyer Verlag.
- Masica, C. P. (1993). *The Indo-Aryan Languages*. Cambridge University Press.
- Mathesius, V. (1975). *A Functional Analysis of Present Day English on a General Linguistic Basis*. The Hague and Paris: Mouton.
- Matthews, D. (1998). *Course in Nepali*. Routledge.
- Mayer, M. (2003). *A boy, a dog, and a frog*. New York: Dial Books for Young Readers.
- McGregor, W. B. (2009). Typology of ergativity. *Language and Linguistics Compass*, 3(1):480–508.
- McGregor, W. B. (2010). Optional ergative case marking systems in a typological-semiotic perspective. *Lingua*, 120(7):1610–1636.
- McNally, L. (1998). On recent formal analyses of topic. In *The Tbilisi symposium on language, logic, and computation: Selected papers*, volume 14, pages 147–160. Stanford, CA: CSLI Publications.
- Meakins, F. (2009). The case of the shifty ergative marker a pragmatic shift in the ergative marker of one. *The role of semantic, pragmatic, and discourse factors in the development of case*, 108:59.

- Meyerhoff, M. (2008). Empirical problems with domain-based notions of "simple". In *Social Lives in Language Sociolinguistics and multilingual speech communities: Celebrating the work of Gillian Sankoff*, volume 24 of *Impact: Studies in Language and Society*, pages 327–357. John Benjamins Publishing.
- Michailovsky, B. (1996). L'inférentiel du népali. In Guenchéva, Z., editor, *L'Énonciation médiatisée. (Bibliothèque de l'Information Grammaticale.)*, pages 109–123. Louvain - Paris: Éditions Peeters.
- Milsark, G. L. (1974). *Existential sentences in English*. PhD thesis, Massachusetts Institute of Technology.
- Næss, Å. (2004). What markedness marks: the markedness problem with direct objects. *Lingua*, 114(9):1186–1212.
- Narasimhan, B., Di Tomaso, V., and Verspoor, C. M. (1996). Unaccusative or unergative? verbs of manner of motion. *Quaderni del Laboratorio di Llinguistica*, 10.
- Patil, U., Kentner, G., Gollrad, A., Kügler, F., Féry, C., and Vasishth, S. (2008). Focus, word order and intonation in hindi. *Journal of South Asian Linguistics*, 1(1).
- Patrika, N. (July 13, 2018). Naming new provinces. *Nepali Times*.
- Perlmutter, D. M. (1978). Impersonal passives and the unaccusative hypothesis. In *annual meeting of the Berkeley Linguistics Society*, volume 4, pages 157–190.
- Peterson, J. (2000). Evidentials, inferentials, and mirativity in Nepali. *Linguistics of the Tibeto-Burman Area*, 23(2):13–37.
- Pine, R. (2009). *Lao-tzu's Taoteching: With Selected Commentaries from the Past 2,000 Years*. Copper Canyon Press.
- Pokharel, M. P. (1998). Categorial splits in the use of -le in nepali. *Nepalese Linguistics: Journal of the Linguistics Society of Nepal*, 15:42–50.
- Portner, P. and Yabushita, K. (1998). The semantics and pragmatics of topic phrases. *Linguistics and Philosophy*, 21(2):117–157.
- Postal, P. M. (1970). On the surface verb 'remind'. *Linguistic inquiry*, 1(1):37–120.
- Poudel, T. (2008). Nepali ergativity: A historical perspective. *Presentation handout from*.
- Pradhān, P. (1944). *Nepālī Vyākaran*. MacMillan and Co., fourth edition.
- Prince, A. and Smolensky, P. (1993). Optimality Theory: Constraint Interaction in Generative Grammar. *Technical Report of the Rutgers Center for Cognitive Science*.

- Ramchand, G. et al. (1997). *Aspect and predication: The semantics of argument structure*. Oxford University Press.
- Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136.
- Rooth, M. (1985). *Association with Focus*. PhD thesis, University of Massachusetts at Amherst.
- Roy, I. A. (2013). *Nonverbal Predication: Copular Sentences at the Syntax-Semantics Interface*. Oxford University Press.
- Sánchez-Alonso, S. (2018). *The Cognitive Sources of Language Change and Variation: Connecting Synchronic Variation and Diachrony in Spanish Copula Use*. PhD thesis, Yale University.
- Schmidt, R. L. (1993). *A practical dictionary of Modern Nepali*. Schoenhofs Foreign Books.
- Shannon, T. (1986). Book notice of Michael Shairo's *The Sense of Grammar*. *Language*, 62.1:233–4.
- Shapiro, M. (1983). *The Sense of Grammar*. Bloomington: Indiana University Press.
- Sharma, S. and Gupta, V. (2012). Hybrid Approach for Punjabi Text Clustering. *International Journal of Computer Applications*, 52(1).
- Siewierska, A. (2013a). Alignment of verbal person marking. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Siewierska, A. (2013b). Verbal person marking. In Dryer, M. S. and Haspelmath, M., editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Silverstein, M. (1976). Hierarchy of features and ergativity. *Grammatical categories in Australian languages*.
- Simon, G. F. and Fennig, C. D., editors (2018). *Ethnologue: Languages of the World, Twenty-first Edition*. Dallas, Texas: SIL International.
- Sims, D. (April 12, 2017). Should Acting Prizes Be Gender-Neutral? *The Atlantic*.
- Stalnaker, R. (1974). Pragmatic Presuppositions. In Munitz, M. and Unger, P., editors, *Semantics and Philosophy*, pages 197–214. New York University Press.
- TFS Working Group (2011). The Woodchopper. Totem Field Storyboards. Retrieved from <http://www.totemfieldstoryboards.org> on Mar 11, 2019.

- Tournadre, N. (1991). The rhetorical use of the Tibetan ergative. *Linguistics of the Tibeto-Burman area*, 14(1):93–107.
- Trubetzkoy, N. (1931). Die phonologischen systeme. *Travaux du Cercle linguistique de Prague*, 4:96–116.
- Turin, M. (2007). *Linguistic diversity and the preservation of endangered languages: A case study from Nepal*. International Centre for Integrated Mountain Development.
- Turnbull, A. ([1923] 1982). *Nepali Grammar & Vocabulary*. Asian Educational Services.
- Turner, R. L. (1931). *A comparative and etymological dictionary of the Nepali language*. K. Paul, Trench, Trubner.
- Van de Visser, M. (2003). ‘Syntactic ergativity’ in Dyirbal and Balinese. *Linguistics in the Netherlands*, 20(1):177–188.
- Van Valin, R. D. and Wilkins, D. (1996). The case for ‘effector’: case roles, agents, and agency revisited. *Grammatical constructions: Their form and meaning*, pages 289–322.
- Velleman, L. and Beaver, D. (2016). Question-based models of information structure. In Féry, C. and Ishihara, S., editors, *The Oxford Handbook of Information Structure*, pages 86–107. Oxford University Press.
- Verbeke, S. (2011). *Ergativity and alignment in Indo-Aryan*. PhD thesis, Ghent University.
- Verbeke, S. (2013). Differential subject marking in nepali: The agent marker le in imperfective constructions.
- Verbeke, S. and De Cuypere, L. (2015). Differential subject marking in nepali imperfective constructions: A probabilistic grammar approach. *Studies in Language*, 39(1):1–23.
- Verkuyl, H. J. (1972). On the compositional nature of the aspects. *Foundations of Language Supplementary Series*, 15.
- Verma, M. K. (1976). The notion of subject and the data from Nepali. *The notion of subject in South Asian languages*, 2.
- Wallace, W. D. (1982). The Evolution of Ergative Syntax in Nepali. *Studies in the Linguistic Sciences*, 12(2):147–211.
- Wallace, W. D. (1985). *Subjects and subjecthood in Nepali: an analysis of Nepali clause structure and its challenges to relational grammar and government & binding*. PhD thesis, University of Illinois at Urbana-Champaign.

- Waugh, L. R. (1982). Marked and unmarked: A choice between unequals in semiotic structure. *Semiotica*, 38(3-4):299–318.
- Whitney, W. D. (1879). *A Sanskrit Grammar*. Breitkopf and Härtel.
- Williams, C. J. et al. (1980). *A grammar of Yuwaalaraay*. Department of Linguistics, Research School of Pacific Studies, The Australian National University.
- Willis, C. M. et al. (2011). Optional case marking in Darma (Tibeto-Burman). *Linguistics of the Tibeto-Burman Area*, 34(2):101.
- Woolford, E. (1997). Four-way case systems: Ergative, nominative, objective and accusative. *Natural Language & Linguistic Theory*, 15(1):181–227.
- Wright, D. (1877). *History of Nepal*. Asian Educational Services.
- Yadava, Y. P., Hardie, A., Lohani, R. R., Regmi, B. N., Gurung, S., Gurung, A., McEnery, T., Allwood, J., and Hall, P. (2008). Construction and annotation of a corpus of contemporary Nepali. *Corpora*, 3(2):213–225.
- Zeisler, B. (2012). Practical issues of pragmatic case marking issues in the Kenhat varieties of Ladakh. *Linguistics of the Tibeto-Burman Area*, 35(1).