

The
Munda
Languages

Edited by
Gregory D. S. Anderson

ROUTLEDGE LANGUAGE FAMILY SERIES

THE MUNDA LANGUAGES

The Munda group of languages of the Austroasiatic family are spoken within central and eastern India by almost ten million people. To date, they are among the least well-known and least documented languages of the Indian subcontinent.

This unprecedented and original work draws together a distinguished group of international experts in the field of Munda language research and presents current assessments of a wide range of typological and comparative–historical issues, providing agendas for future research.

Representing the current state of Munda linguistics, this volume provides detailed descriptions of almost all of the languages in the family.

Gregory D.S. Anderson is Director of the Living Tongues Institute for Endangered Languages. His key publications include: *The Munda Verb: Typological Perspectives* (2007), *Auxiliary Verb Constructions* (2006) and *Language Contact in South Central Siberia* (2005).

ROUTLEDGE LANGUAGE FAMILY SERIES

Each volume in this series contains an in-depth account of the members of some of the world's most important language families. Written by experts in each language, these accessible accounts provide detailed linguistic analysis and description. The contents are carefully structured to cover the natural system of classification: phonology, morphology, syntax, lexis, semantics, dialectology, and sociolinguistics.

Every volume contains extensive bibliographies for each language, a detailed index and tables, and maps and examples from the languages to demonstrate the linguistic features being described. The consistent format allows comparative study, not only between the languages in each volume, but also across all the volumes in the series.

The Austronesian Languages of Asia
and Madagascar

*Edited by Nikolaus Himmelmann and
Sander Adelaar*

The Bantu Languages

*Edited by Derek Nurse and
Gérard Philippson*

The Languages of the Caucasus

Edited by Alice Harris

The Celtic Languages

(new edition forthcoming)

Edited by Martin J. Ball and James Fife

The Dravidian Languages

Edited by Sanford B. Steever

The Germanic Languages

*Edited by Ekkehard Konig and
Johan van der Auwera*

The Indo-Aryan Languages

*Edited by George Cardona and
Dhanesh K. Jain*

The Indo-European Languages

*Edited by Paolo Ramat and
Anna Giacalone Ramat*

The Iranian Languages

Edited by Gernot Windfuhr

The Khoesan Languages

Edited by Rainer Vossen

The Mongolic Languages

Edited by Juha Janhunen

The Oceanic Languages

*Edited by John Lynch, Malcolm Ross
and Terry Crowley*

The Romance Languages

*Edited by Martin Harris and
Nigel Vincent*

The Semitic Languages

Edited by Robert Hetzron

The Sino-Tibetan Languages

*Edited by Graham Thurgood and
Randy J. Lapolla*

The Slavonic Languages

*Edited by Bernard Comrie and
Greville G. Corbett*

The Tai-Kadai Languages

(forthcoming)

Edited by Anthony Diller

The Turkic Languages

*Edited by Eva Csato and
Lars Johanson*

The Uralic Languages

Edited by Daniel Abondolo

The Munda Languages

Edited by Gregory D.S. Anderson

THE
MUNDA
LANGUAGES

Edited by
Gregory D.S. Anderson

First published 2008
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
711 Third Avenue, New York, NY 10017, USA

*Routledge is an imprint of the Taylor & Francis Group,
an informa business*

© 2008 Gregory D.S. Anderson for selection and editorial matter;
individual chapters, the respective contributors

Typeset in Times New Roman by
Newgen Imaging Systems (P) Ltd, Chennai, India

All rights reserved. No part of this book may be reprinted or reproduced
or utilized in any form or by any electronic, mechanical, or other means,
now known or hereafter invented, including photocopying and recording,
or in any information storage or retrieval system, without permission in
writing from the publishers.

British Library Cataloging in Publication Data

A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data

A catalog record for this book has been requested

ISBN10: 0-415-32890-X

ISBN13: 978-0-415-32890-6

To Norman

This page intentionally left blank

CONTENTS

<i>List of illustrations</i>	xi
<i>Notes on contributors</i>	xv
<i>Preface</i>	xvii
<i>List of abbreviations</i>	xix
1 Introduction to the <i>Munda Languages</i> <i>Gregory D.S. Anderson</i>	1
2 Santali <i>Arun Ghosh</i>	11
1 The language and its speakers	11
2 Phonology	20
3 Morphology	32
4 Syntax	74
5 Semantics/Discourse	85
6 Lexicon	87
7 Brief analysed text	88
3 Mundari <i>Toshiki Osada</i>	99
1 Introduction	99
2 Phonology	100
3 Morphology	106
4 Syntax	145
5 Semantics/Discourse	153
6 Lexicon	156
7 Brief analysed texts	156
4 Kera? Mundari <i>Masato Kobayashi and Ganesh Murmu</i>	165
1 Introduction	165
2 Phonology	168
3 Morphology	170
4 Syntax	184
5 Semantics/Discourse	188
6 Lexicon	188
7 Brief analyzed text	188
5 Ho and the other Kherwarian languages <i>Gregory D.S. Anderson, Toshiki Osada, and K. David Harrison</i>	195
1 Introduction	195
2 Phonology	198
3 Morphology	204

4	Syntax	231
5	Semantics/Discourse	236
6	Lexicon	237
7	Brief analyzed text	240
8	Ho orthography	250
6	Korku <i>Norman H. Zide</i>	256
1	Introduction	256
2	Phonology	257
3	Morphology	262
4	Syntax	288
5	Semantics/Discourse	293
6	Lexicon	293
7	Brief analyzed texts	293
7	Sora <i>Gregory D.S. Anderson and K. David Harrison</i>	299
1	Introduction	299
2	Phonology	301
3	Morphology	306
4	Syntax	362
5	Semantics/Discourse	370
6	Lexicon	371
7	Brief analysed texts	371
8	Gorum <i>Gregory D.S. Anderson and Felix Rau</i>	381
1	Introduction	381
2	Phonology	382
3	Morphology	387
4	Syntax	414
5	Semantics/Discourse	420
6	Lexicon	422
7	Brief analysed texts	423
9	Kharia <i>John Peterson</i>	434
1	Introduction	434
2	Phonology	435
3	Morphology	441
4	Syntax	483
5	Semantics	492
6	Lexicon	493
7	Brief analysed text	495
10	Juang <i>Manidepa Patnaik</i>	508
1	Introduction	508
2	Phonology	509
3	Morphology	513
4	Syntax	539
5	Semantics	548

6	Lexicon	548
7	Brief analyzed texts	549
11	Remo (Bonda) <i>Gregory D.S. Anderson and K. David Harrison</i>	557
1	Introduction	557
2	Phonology	558
3	Morphology	568
4	Syntax	608
5	Semantics/Discourse	617
6	Lexicon	619
7	Brief annotated sample text	623
12	Gutob <i>Arlo Griffiths</i>	633
1	Introduction	633
2	Phonology	637
3	Morphology	640
4	Syntax	665
5	Semantics/Discourse	669
6	Lexicon	670
7	Brief analyzed text: the Go?ter ritual	671
13	Gta? <i>Gregory D.S. Anderson</i>	682
1	Introduction	682
2	Phonology	682
3	Morphology	688
4	Syntax	743
5	Semantics/Discourse	754
6	Lexicon	756
7	Brief analyzed texts	756
14	On Nihali <i>Norman H. Zide</i>	764
1	Overview of Nihali	764
2	Proposed linguistic layers in Nihali	766
3	What kind of a (contact-shaped) language is Nihali?	772
	<i>Index</i>	777

This page intentionally left blank

ILLUSTRATIONS

MAPS

1.1	Approximate distribution of Munda languages	2
1.2	Approximate distribution of Munda languages of Orissa	3
4.1	Approximate area where Kera? Mundari is spoken	166

FIGURES

1.1	Tree diagram detailing 'traditional' classification of Munda languages	3
1.2	Tree diagram detailing revised traditional classification of Munda languages	4
1.3	Tree diagram of classification of Proto-South Munda languages as revised by Anderson	4
5.1	Mr K.C. Naik Biruli, Ho orator and scholar, 2005	250
5.2	The Ho orthography, a handwritten version from 2005, including number symbols (lower row), but excluding lower-case letters	251
5.3	A sample page from the Ho alphabet primer book showing the use of both upper- and lower-case forms	251
5.4	Handwritten sample of Ho orthography, with upper- and lower-case letters, 2005	252

TABLES

2.1	Distribution of Santali speakers in India	14
2.2	Santali vowels	20
2.3	Santali vowel combinations	24
2.4	Santali consonants	26
2.5	Native sub-system	27
2.6	Borrowed sub-system	27
2.7	Case markers	35
2.8	Personal pronouns	41
2.9	Interrogative pronouns	43
2.10	Indefinite pronouns	44
2.11	Other indefinites	44
2.12	Demonstratives	45
2.13	Demonstratives referring to sight and sound	45
2.14	Santali agreement markers	54
2.15	Direct object pronominal infixes	56
2.16	Aspect suffixes with active and middle markers	58

3.1	Mundari vowels	100
3.2	Mundari consonants	101
3.3	Mundari dialect comparison	101
3.4	Mundari word CV structure	103
3.5	Mundari harmony feature sets	105
3.6	Mundari genitive	109
3.7	Mundari pronouns	109
3.8	Mundari possessive pronouns – I	110
3.9	Mundari possessive pronouns – II	110
3.10	Mundari demonstratives – I	110
3.11	Mundari demonstratives – II	111
3.12	Mundari interrogatives	112
3.13	Mundari indefinites	112
3.14	Mundari numerals	114
3.15	Mundari verb template	119
3.16	Mundari agreement markers	120
4.1	Personal pronouns	172
4.2	Demonstratives	173
4.3	Object paradigm	178
5.1	A comparative vocabulary in the Kherwarian languages	238
9.1	The consonants	437
9.2	Inalienable possessive markers	444
9.3	The free-standing pronominals	446
9.4	Interrogatives	447
9.5	Kharia numerals 30–100	449
9.6	Infix <i>-NV-</i> with reduplication	453
9.7	The <i>-EL-AY</i> alternation	459
9.8	Person/Number/Honorific Marking on the Predicate	460
9.9	The basic TAM/voice categories	462
9.10	Predicating lexemes which appear in both the middle and active	466
9.11	‘Reflexivity’ and basic voice	467
9.12	‘Thoroughness’ and basic voice	467
9.13	Suddenness/(Non-)volitionality and basic voice	467
9.14	The present copulae, non-negative and negative	484

BOXES

5.1	Ho verb template	217
7.1	Sora finite verb template	331
7.2	Sora incorporation and complex verb template	360
8.1	Possible vowel combinations in Gorum	386
9.1	A schematic overview of the NP	441
9.2	Kharia numerals 1–10	448
9.3	Kharia numerals 11–19	449
9.4	Numerals borrowed from Sadani	450
9.5	Collectives	450
9.6	The ‘nominalizer’ <i>-NV-</i>	452

9.7	A schematic overview of the non-negated morphologically fully finite predicate in Kharia	456
9.8	A schematic overview of the negated morphologically fully finite predicate in Kharia	471
10.1	Reflexives in Juang I	522
10.2	Reflexives in Juang II	522
10.3	Reciprocals in Juang I	522
10.4	Reciprocals in Juang II	522
10.5	Juang noun structure	524
10.6	Juang basic verb structure	527
12.1	Positive finite verb forms	652
12.2	Converbs	653

This page intentionally left blank

CONTRIBUTORS

Gregory D.S. Anderson is Director of the Living Tongues Institute for Endangered Languages (Salem, Oregon, USA). He is an expert in and has published widely on the Munda languages, indigenous Siberian languages, linguistic typology, Native North and South American languages, Eleme (Nigeria) and Burushaski, having worked with speakers of indigenous, endangered, threatened and minority language communities around the globe (including numerous Munda languages). Dr Anderson is also the co-developer of the Language Hotspots theory.

Arun Ghosh is Professor of Linguistics at Burdwan University in West Bengal, India. He is an expert in Bangla and Munda languages. He has worked with speakers of a number of Munda languages, both North Munda and South Munda. In addition to his expertise in Santali, Professor Ghosh is currently working on issues in the grammar of Bonda (Remo) and Didayi (Gta?). He has published extensively, not only on Munda linguistics but also on broader Austroasiatic issues.

Arlo Griffiths is Professor of Sanskrit at Leiden University in the Netherlands. He is an internationally renowned Vedic scholar having produced the first translation of the Paipillad rescension of the Atharvaveda. His interest in Gutob, including several field visits, began a decade ago before his rise to becoming the youngest Chair in Sanskrit at Leiden.

K. David Harrison is Assistant Professor of Linguistics at Swarthmore College, Pennsylvania, USA and Director of Research at the Living Tongues Institute for Endangered Languages. He is the leading recognized international expert on endangered languages. Dr Harrison has worked with speakers of half a dozen Munda languages, both North and South Munda, during field trips in Orissa. Dr Harrison is also the co-developer of the Language Hotspots theory.

Masato Kobayashi is Assistant/Associate Professor of Linguistics at Hakuoh University. Dr Kobayashi has an unparalleled breadth of knowledge of South Asian linguistics. His extreme erudition is attested to not only by his work in Munda linguistics, but also by the fact that he has published extensively in comparative Indo-Aryan linguistics, and in a host of varied topics in Dravidian linguistics.

Ganesh Murmu is Associate Professor and Senior Researcher at the Department of Tribal Languages and Literatures at Ranchi University, Jharkhand, India. A Santal by ethnicity and speech, Dr Murmu brings a unique insight in Kherwarian linguistic analysis to this volume and his many excellent publications continue the great tradition of Munda-speaking linguists, including J. Kullu, P. Kerketta, and Dr R.D. Munda.

Toshiki Osada is Professor at the Research Institute for Humanity and Nature located in Kyoto, Japan. He is the recognized expert on the Kherwarian group of languages,

being an excellent speaker of Mundari. His diverse list of publications in Munda linguistics include lexical materials, field notes on undescribed varieties and theoretically oriented papers on word classes in linguistic analysis, among them works in both Japanese and English. Professor Osada is also a director of the Japanese-led comprehensive project on the Harrappa and Indus civilizations in Pakistan.

Manideepa Patnaik is the recognized expert in both the analysis and description of the Juang language, which she has researched for the past two decades, as well as in the syntax of a number of Munda languages including Sora. Dr Patnaik is currently the Director of the Institute of Indian Languages Literatures and Cultures, Orissa. Her publications and presentations in Munda linguistics have focused primarily on the structure of complex clauses or sentences in languages of South Asia and in Munda languages such as Juang specifically.

John Peterson is Professor of Linguistics at Osnabrück University, having finished his monumental *Habilitationsschrift* on Kharia in 2006. This now sets Kharia apart as the single best described Munda language, in particular in terms of its syntactic features which will be very important to understanding the (pre-)history of this important and somewhat enigmatic Munda language. Professor Peterson's work on Munda has been focused on descriptive and theoretical aspects of Kharia syntax and grammar more widely. He has also published extensively on northern Indo-Aryan languages.

Felix Rau is a PhD student at Leiden University. He is a promising young scholar who already has considerable expertise in Gorum, Desia, Telugu and a range of other languages. His contribution to this volume suggests a sophisticated and nuanced appreciation of Munda language structure that is advanced for someone of any age. Munda specialists are eagerly awaiting his field-based description of Gorum.

Norman H. Zide is Professor Emeritus at the University of Chicago and is the founder of contemporary Munda linguistics as a scholarly discipline in the West (in India it must be Ramamurti). He has overseen the development of the field from its embryonic state, and initiated the first systematic and only non-Indian (to date) Munda Languages Survey. He has published or written papers in circulation among *cognoscenti* on almost every imaginable topic in Munda linguistics, ranging from Gutob baby talk to comparative Munda phonology, cultural history, echo formation in Gta?, the morphosyntax of verb agreement, and Korcu tonal processes, etc.

PREFACE

The Munda Languages volume herein is the result of nearly a decade of work. Although nine years have passed since the chapters were first commissioned from their original intended authors, this was deemed too short a time for some. As originally conceived, the volume was very different. For a variety of reasons the volume has morphed into what it is now, twelve rather different sketches of the major languages and subgroups of the Munda language family, the westernmost representative of the far-flung Austroasiatic linguistic phylum as well as a brief discussion of Nihali/Nahali which may be related to the Munda languages (or may be an isolate language).

It has not been a labour of love for the editor but it has indeed been a labour. Several would-be contributors were unable to offer their work and others were prevented for a wide range of personal reasons from contributing all that they had intended. One chapter changed its author[s] three times, others twice. Indeed, never before has such a collection of works on Munda languages been attempted on this scale (nor is likely to be again for a long time, if ever). Some of the languages in this volume receive here their first comprehensive description, e.g. Gta? or Kera? Mundari.

All in all, this book represents the state of Munda linguistics at the beginning of the twenty-first century. Contributors are taken from all over the world: three from India (including one native Munda-speaking scholar), two from Japan, three from central Europe (Germany/Holland) and three from the United States. To be sure, this includes contributions from practically every researcher currently active in the field of Munda linguistics with only a few exceptions. Each approaches the analysis of the Munda languages from a different background and perspective. Naturally this includes some variation in both notation and terminology, as well as certain points of analysis. For example, palatal nasals may be represented by \tilde{n} or by η , etc. Other conventions, on the other hand, are followed across the chapters including the use of c and j for the palatal obstruents (stops/affricates) and the IPA symbols for the retroflex (or post-alveolar) series, viz. $ɖ$, $ʈ$, $ʂ$, etc. Note that variation in analysis is seen within a single article, where all parties in co-authored papers may not agree with each other or share the same interpretation of the data. Overall it is believed that the interested reader will derive a wealth of benefits from the materials presented in this volume.

The chapters all follow the same general outline to make comparison easier across them. Some individual chapters deviate from the common pattern of numbered sections slightly, but all sub-headings are clearly marked and the volume is well-indexed so these are mainly minor cosmetic issues.

The editor would like to thank the following people for helping to bring this volume into existence. First, thanks to all contributors for the fine content of their chapters. An especially large thank you must be given to Don Reneau, who as

always has offered me sage advice and help on technical formatting issues, including literally running to the rescue of a contributor who was having insurmountable technical difficulties. Manideepa Patnaik facilitated arrangements for the field trips to India. Mark Eglinton offered his photographic expertise on the 2005 field trip. Sara Foerster helped with a small but gratefully appreciated fraction of the re-keying of contributions by those who did not use Word and/or the SIL IPA fonts. Special thanks are here offered to Ironbound Films who made the 2007 field trip a possibility and the Living Tongues Institute for Endangered Languages for allowing me the opportunity to do this research. I would also like to thank my wife Mary and youngest son Oliver for bearing me during the final months of preparation when these issues were literally driving me insane. I would also like to thank my primary language consultants who shared with me their knowledge and time and patience and without whom the final product would have been impossible and indeed pointless. These include K.C. Naik and C.M. Haibru (Ho), C.M. Singh (Bhumij), S. Dangada-Majhi (Remo), and Opino Gomango and Oruncho Gomango (Sora), among others.

Gregory D.S. Anderson
Salem, Oregon, USA
December 2007

ABBREVIATIONS

-	Affix boundary	COM	Comitative
.I	First conjugation	COMP	Complementizer
.II	Second conjugation	COMPL	Completion
=	Clitic boundary	CONAT	Conative
1	First person	COND	Conditional
2	Second person	CONJ	Conjunctive
3	Third person	CONT	Continuative
A	Active	CONTNG	Contingent action mood
A:TEL	Anticipatory telic	COP	Copula
ABL	Ablative	CR	Correlative
ACC	Accusative	CUST	Customary
ACT	Active	CV	Converb
ADD	Additive	DAT	Dative
ADESS	Adessive	DEF	Definite
ADJ	Adjective	DEIC	Deictic
ADV	Adverbial	DEM	Demonstrative
AFF	Affective/Affected	DEP	Dependent
AKT	Aktionsart	DESID	Desiderative
ALL	Allative	DIR	Directional
AMB	Ambulative	DISC	Discourse element
AN	Animate	DIST	Distal
ANAPH	Anaphoric Pronominal	DL	Dual
ANT	Anterior	DPT	Departive
ART	Article	DS	Different subject
ATTR	Attributive	DUBIT	Dubiative
AUGM	Augment	DUR	Durative
AUTOPOES	Auotpoesis	EM	<i>Encyclopedia Mundarica</i>
AUX	Auxiliary	EMOT	Emotive 'light verb'
BEN	Benefactive	EMPH	Emphatic
C:TEL	Culminatory telic	EPEN	Epenthetic
CAP	Capabilitive	EX	Exclusive
CAUS	Causative	EXCES	Excessive
CF	Combining form	EXCLAM	Exclamatory
CLOC	Cislocative	EXPR	Expressive
CLSSFR	Classifier	FEM	Feminine
CMPLT	Completive	FIN	Finite
CNCTV	Connective	FOC	(Restrictive) Focus
CNTR	Contrastive focus	FUT	Future
COLL	Collective	GEN	Genitive

HON	Honorific	PERS	Person
HPA	Homo-organic post-alveolar	PFV	Perfective
HUM	Human	PFX	Prefix
IMP	Imperative	PL	Plural
IMPF	Imperfect	PLUP	Pluperfect
INAN	Inanimate	POL	Politeness
INC	Inclusive	POSS	Possessive
IND	Indicative	PP	Postposition
INDEF	Indefinite	PR	Pronominal
INESS	Inessive	PRF	Perfect
INF	Infinitive	PROG	Progressive
INFER	Inferential	PROHIB	Prohibitive
INGR	Ingressive	PRON	Pronominal
INS	Instrumental	PROX	Proximal
INTENS	Intensifier	PRS	Present
INTERJ	Interjection	PRTCL	Particle
INTNSV	Intensive	PRTCPL	Participle
IPFV	Imperfective	PST	Past
IRR	Irrealis	PURP	Purposive
ITER	Iterative	Q	Interrogative
ITR	Intransitive	QF	Quantifier
LEX	Lexical Verb	QUANT	Quantifying
LOC	Locative	QUOT	Quotative
LV	Loan Verb	RECIP	Reciprocal
M	Middle (Kharia, Santali)	REDPL	Reduplication
MASC	Masculine	REP	Repetition (full copy of word)
MID	Middle (Gutob)	REL	Relative
MID.1	Middle deictic 1	RFLXV	Reflexive
MID.2	Middle deictic 2	RHET	Rhetorical
MOD	Modal	SBJNCT	Subjunctive
MODFR	Modifier	S:ITER	Semel-iterative
N.SFX	<i>n</i> -suffix	SEQ	Sequential converb
NARR	Narrative	SG	Singular
NEG	Negative	SIMULT	Simultaneous
NF	Non-Finite	SM	South Munda
NHUM	Non-Human	SOC	Sociative
NM	North Munda	SS	Same Subject
NML	Nominalizer	SUB	Subessive
NMLZ	Nominalizer	SUBJ	Subject
NONHUM	Non-Human	SUBORD	Subordinate
NPST	Non-past	SUD	Sudden action
NP	Noun-Phrase	SUPERESS	Superessive
NUM	Number	SUPERLAT	Superlative
OBJ	Object[ive]	T/A	Tense/Aspect
OBLQ	Oblique	TAG	Tag-word
OPT	Optative	TAM	Tense-Aspect-Mood
PASS	Passive		

TEL	Telicity	VBLZ	Verbalizer
TLOC	Translocative	VOC	Vocative
TNS	Tense	w/o	Without
TOTAL	Totality	x1/y/x2	Y = Infix
TR	Transitive	x1<y> x2	Y = Infix

This page intentionally left blank

INTRODUCTION TO THE *MUNDA LANGUAGES**

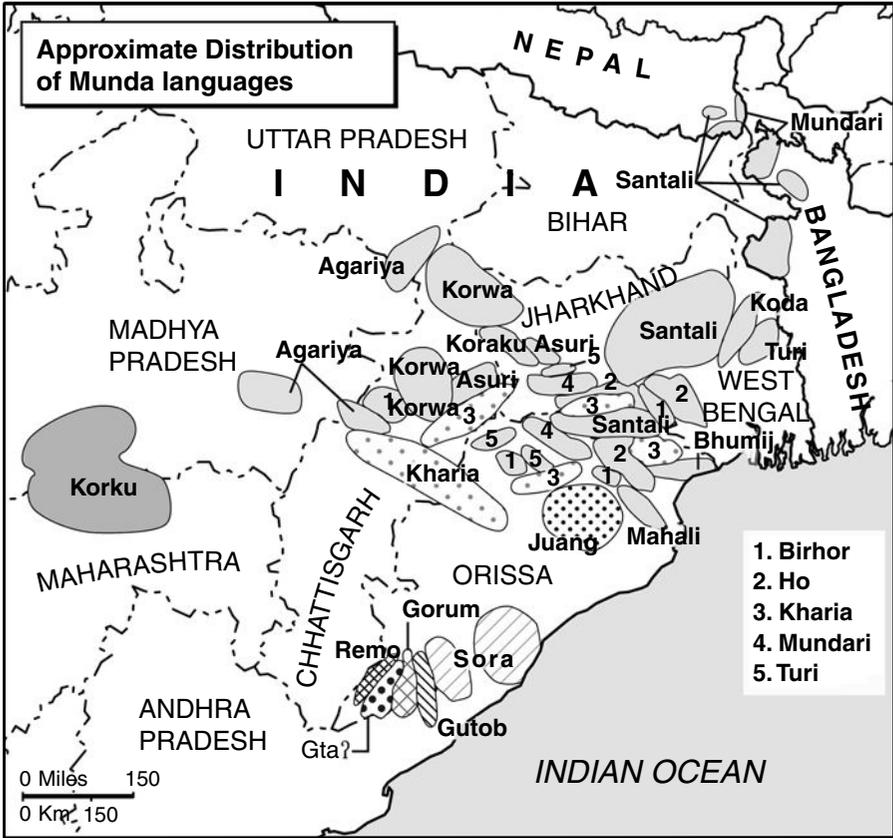
Gregory D.S. Anderson

1 OVERVIEW

The Munda languages are a group of Austroasiatic languages spoken across portions of central and eastern India by perhaps as many as ten million people total. The Munda peoples are generally believed to represent the autochthonous populations over much of their current areas of inhabitation. Originally, Munda-speaking peoples probably extended over a somewhat larger area before being marginalized into the relatively remote hill country and (formerly) forested areas primarily in the states of Orissa and the newly constituted Jharkhand; significant Munda-speaking groups are also to be found in Madhya Pradesh, and throughout remote areas of Chhattisgarh, West Bengal, Uttar Pradesh, Andhra Pradesh, and Maharashtra (see Map 1.1). Of course much of this territory was settled or colonized by the Indo-Aryan speakers and, at an earlier period, by the Dravidian-speakers as well, which concluded about 2,500 years ago or so.

The pre-history of the Munda languages remains obscure. Munda languages constitute the westernmost representatives of the far-flung Austroasiatic linguistic phylum. Two other Austroasiatic groups are found in the present-day territory of India, the Khasi of Meghalaya and the Nicobarese-speaking groups of the Nicobar Islands. The other subgroups of Austroasiatic are all found outside of India, and it is generally believed that the Austroasiatic ancestral language was not to be found in India but rather further to the east. Thus, at some point the ancestors of the Munda-speaking peoples must have migrated westward into the subcontinent. When, how, and by what path they entered India remains a subject of considerable debate. Indeed, it is not even clear that there was a single migration of pre-Munda speakers, but there may have been two or more such movements.

Consensus has not yet been reached on the internal relationships of the Munda languages, but several subgroups have been proposed and some of these appear to be sound. It is hoped that further work in comparative Munda grammar and lexicon may shed light on this issue. The northern-, eastern- and westernmost groups of Munda languages are clearly related and appear to fall into two broad groupings. The first of these is the westernmost Munda language Korku (Zide, this volume-a) which appears to be a sister to the remainder of this subgroup, conventionally labelled Kherwarian. This latter unit is really a large and complex dialect/language chain, the better-known varieties of which are known as Santali (Ghosh, this volume) – by far the largest of the Munda groups – and its close sister languages Mundari (Osada, this volume) and Ho (Anderson *et al.*, this volume). Kherwarian also includes a number of minor varieties as well, for example, Turi, Asuri, Birhor, Bhumij, Korwa, etc. as well as a special quasi-creolized variety of Mundari used by former Kurukh speakers, that is,



MAP 1.1 APPROXIMATE DISTRIBUTION OF MUNDA LANGUAGES

Kera? (Kobayashi and Murmu, this volume). Korku and Kherwarian together are conventionally known as North Munda; this is a secure subgroup within Munda.

The remaining Munda languages are almost only found in the state of Orissa (some Kharia speakers are found in Jharkhand, West Bengal, and Chhattisgarh as well), which appears to be the epicentre of diversity of the family (see Map 1.2). How each of these non-North Munda languages or subgroups (logically known as South Munda in contrast to North Munda) are related to each other remains a topic of considerable debate. Some languages clearly form subgroups, such as Sora (Anderson and Harrison, this volume) and Gorum/Parenga (Anderson and Rau, this volume) in the Sora-Gorum subgroup or Gutob (Griffiths, this volume) and Remo/Bonda (Anderson and Harrison, this volume) in the Gutob-Remo subgroup. The classification of the remaining three languages remains an open question, as does how exactly the non-North Munda languages diversified and developed from the common Proto-Munda ancestral language. These three languages are Kharia (Peterson, this volume), Juang (Patnaik, this volume), and Gta?/Didayi (Anderson, this volume).

The ‘traditional’ classification of Munda is to be found in Zide (1969) and Zide and Stampe (1968), and may be represented graphically as follows (see Figure 1.1).



MAP 1.2 APPROXIMATE DISTRIBUTION OF MUNDA LANGUAGES OF ORISSA

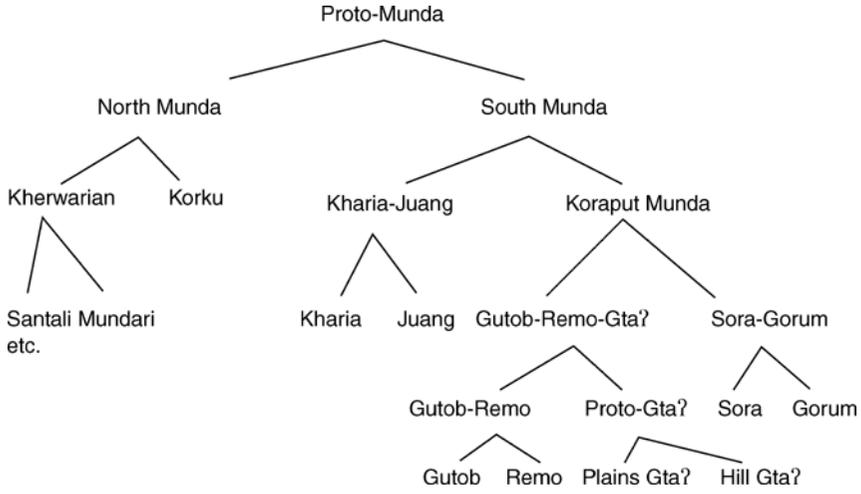


FIGURE 1.1 TREE DIAGRAM DETAILING 'TRADITIONAL' CLASSIFICATION OF MUNDA LANGUAGES

Anderson (1999) is a first recent attempt to rethink this traditional classification. He modified the tree as seen in Figure 1.1 only slightly by eliminating the Koraput Munda node and having South Munda branch directly into three daughter groups, still keeping Kharia-Juang and Gutob-Remo-Gta? intact (see Figure 1.2).

A more radical revision was proposed by Anderson (2001), which adopted a slightly different approach, whereby only the clearly obvious subgroups within South Munda were recognized as such (see Figure 1.3). He proposed an early dialect continuum in which various languages shared certain features (some retentions and some innovations) but did not form a taxonomic unit *per se*. Such a chain has been called by Malcolm Ross (e.g. 1996) ‘a linkage’ in the literature on Austronesian and Papuan languages. This revision is largely motivated by the curious, non-archaic, and multifaceted parallels shared between Kharia and Proto-Gutob-Remo which the traditional understanding of South Munda could not explain.

Whether and/or how Munda is related to the enigmatic Nihali/Nahali language of Madhya Pradesh is a complicated question, but for an introduction to the issues, see Zide (this volume-b).

It is surprising that nothing in the way of quotations from a Munda language turned up in (the hundreds and hundreds of) Sanskrit or middle-Indic texts. There

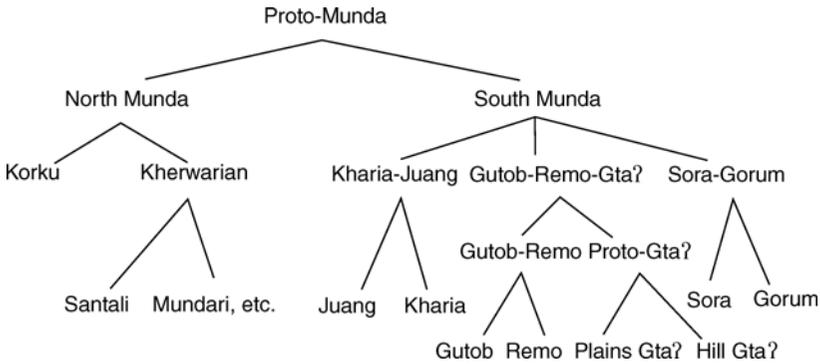


FIGURE 1.2 TREE DIAGRAM DETAILING REVISED TRADITIONAL CLASSIFICATION OF MUNDA LANGUAGES (ANDERSON 1999a)

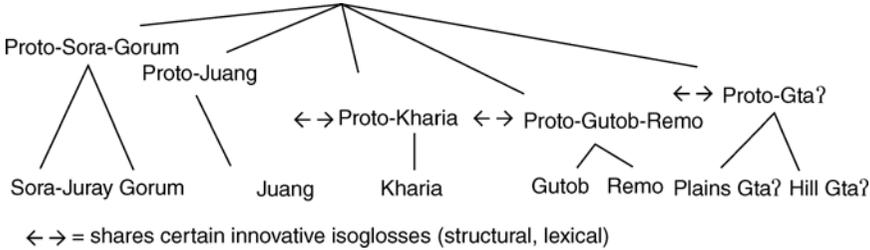


FIGURE 1.3 TREE DIAGRAM OF CLASSIFICATION OF PROTO-SOUTH MUNDA LANGUAGES AS REVISED BY ANDERSON (2001)

is also a surprising lack of borrowing of names of plant/animal/bird, etc. into Sanskrit (Zide and Zide 1976). Much of what has been proposed for Munda words in older Indic (e.g. Kuiper 1948) has been rejected by careful analysis. Some possible Munda tribal names have been proposed, for example, *Savara* (Sora) or *Khara*, but ethnonymy is notoriously messy for the identification of language groups, and a single ethnonym may be adopted and used for linguistically rather different or entirely unrelated groups.

Despite being spoken in a country with a written tradition over two millennia old, the Munda languages remained in total obscurity until the middle of the nineteenth century. Throughout the twentieth century, with a few notable exceptions, the minor Munda languages were quietly disappearing while the intellectual community (both in India and abroad) showed no concern. Because the so-called Kol(arian) languages, as this group was frequently known in the past, were spoken by ‘backward tribals’, they received little attention from early scholars. Various Christian missionaries have provided us with many of the early works on the Munda languages.

The first known reference to data from a Munda language may be Voysey (1821, published 1844). Before Cust (1878, 1884), most of the mention of the Munda languages comes from vocabularies and early primers, for example, Tickell (1840a/b), Hodgson (1848), Phillips (1845), Phillips (1852, Santali grammar), Campbell (1866, Santali vocabulary), Puxley (1868, Santali vocabulary), Das (1871, Mundari grammar), Skrefsrud (1873, Santali grammar), Coates (1875, Chota Nagpur vocabulary) or Cole (1879, Santali article), Pendercast (1881, Sora vocabulary). During the 1890s Rev. de Smet’s *Rudiments of Mundari Grammar* (1891), Crooke’s (1892) Korwa vocabulary, Banerjee’s *Kharia Grammar* (1894), Hoffmann’s early work (1893), and part 1 of Campbell’s *Santal Dictionary* (1899) were published.

The first decade of the twentieth century witnessed a relative boom in Munda linguistics; for example, the second and third parts of Campbell’s dictionary appeared (1900, 1903). Other important publications of this era were Nottrott’s *Kol Grammar* (1905), Hoffman’s Mundari works (1903, 1909) and Cole’s Santali study (1906). This period had also witnessed Grierson’s *Linguistic Survey of India* (volume 4) and various other works and Schmidt’s groundbreaking study establishing the Austroasiatic connection of the Munda languages (1906). Burrows’ *Ho Grammar* (1915), various studies by Roy, Przyluski and especially P. Boddington’s Santal materials were highlights of the 1910s and 1920s, as well as the ethnographically interesting ‘Tea District’ handbooks, with their bizarre phrase-book qualities that were produced for Sora, Kharia, and Mundari, in the wake of an increasing Munda-speaking migrant labour force mobilized to work in the tea plantations of Assam and Darjeeling.¹ The period from 1930s through 1940s featured the works of the French Austroasiaticists Haudricourt and de Hevesy, and the Dutch Indologist F.B.J. Kuiper; however, two studies of this era stand far above the others in their impact on Munda linguistics – Hoffmann’s seminal *Encyclopedia Mundarica* (1930–1950) and Ramamurti’s landmark *Manual of the Sora (Savara) Language* (1931). The 1950s produced works by MacPhail and Biligiri, H. Maspero in the French Austroasiaticist tradition and Pinnow’s Kharia and Juang materials.

The 1960s and 1970s witnessed another boom in Munda linguistics. Pinnow (1966) is the first comprehensive study on the verb in the Munda languages. Written in 1960, it lacks knowledge of the existence of Gta?. Norman Zide, the pre-eminent Western scholar in the field of the emergent discipline of Munda linguistics, led a group of linguists (including A. Zide, D. Stampe, K.C. Bahl, and R.D. Munda at

the University of Chicago and other linguists from a range of institutions, including F. Fernandez, S. Starosta, D. Matson, R. King, etc.) under the auspices of the Munda Language Project, which produced a number of dissertations and smaller works in this period. Indian linguists involved in Munda studies in this period include B. Das, K. and B.P. Mahapatra and S. Bhattacharya. In the next two decades, in addition to a variety of publications by N. Zide, A. Zide and S. Starosta, P. Donegan and D. Stampe, Munda languages have occasionally garnered interest from linguists pursuing a range of typologically or theoretically oriented lines of research, including Masica (1976; various South Asian areal features), and Mithun (1984; with brief discussions of incorporation in some South Munda languages).

In the 1990s and early twenty-first century, a new generation of linguists has undertaken an intensive study of the Munda languages. In addition to Norman Zide, still active in this period, the names of Toshiki Osada, John Peterson, Arun Ghosh, Greg Anderson, Manideepa Patnaik, Ganesh Murmu, K.S. Nagaraja, Felix Rau, Masato Kobayashi, and N. Ramaswami have appeared in connection with a range of descriptive, typological, and comparative historical studies of the Munda language family. Linguists approaching Munda data from a theoretical or South Asian areal perspective from the last two decades include Abbi (1992, reduplication), Odden (1987, theoretical aspects of Gta? phonology), Hook (1991, on complex verb structure), and Sadock (1991, syntactic/theoretical aspects of Gta? incorporation).

A number of Munda-speaking scholars have actively participated in the linguistic analysis and study of the Munda languages such that an indigenous scholarly tradition may be spoken of. These include the Santals Ganesh Murmu and R. Murmu, D. Sahu, Kharia J. Kullu, and P. Kerketta to name but a few.

Most Munda languages remain unwritten or have only fledgling literary varieties. Given the pressure to have a script of one's own to be considered a 'real' language prevalent in the South Asian area (Zide 1996, 2000), it comes as no surprise that there are three or four of the Munda languages which have had indigenous script traditions in the twentieth century. Of these, only the Warang Chiti script of the Ho appears to have any chance of gaining acceptance (see Anderson *et al.* [this volume] on Warang Chiti and Ghosh [this volume] for a discussion of the seeming failure of the Ol Ciki script of the Santals, also Pinnow 1972).

What will researchers find interesting about the Munda languages and what remain the outstanding issues in the descriptive analysis of these languages? Little data on the Munda languages are in wide circulation among linguists who may be inclined to know about them. This volume will serve as a first step in the direction of bringing to light some of the noteworthy features of the individual Munda languages and the family as a whole. All chapters in this volume more or less follow the same general outline, and are largely weighted to morphological structures in the languages. Nevertheless, researchers on such diverse topics as reduplication, noun incorporation, articulatory phonetics, numeral systems, agreement morpho-syntax, complex predicate structure, or nasalization will all find some aspect of some Munda language intriguing, stimulating or even challenging analysis.

Among the most interesting of the linguistic phenomena to be found in Munda languages may be included the highly elaborated systems of demonstratives found in many Munda languages (see, for example, the discussion of these systems in the chapters on Santali, Mundari, or Gorum in this volume). Munda vowel and consonant systems can be quite complex, with different register and secondary articulatory features, many of which are still now in need of description. Another topic of

considerable interest are the elaborate and intersecting systems of voice/valence/transitivity, person-marking, and tense/aspect that characterize Kherwarian verbal systems. Further, the highly elaborated system of noun incorporation found in Sora push the limits of our understanding of such constructions from a theoretical perspective. The highly developed systems of reduplication and expressive formation that characterize most Munda languages also bear mention here. Finally, the interaction of tense/aspect marking and negative operators in negative formations in South Munda Gutob stand out among the most complex of such systems known.

As for what topics remain for future analysis, it can be said without reservation, that almost all aspects of every Munda language require more analysis before we have an adequate consensus understanding of even their basic features. That said, it is clear from a comparison of the various chapters that syntactic issues and phonetic analysis are in desperate need of further systematic investigation. Studies on topics in the semantics and discourse of Munda languages are practically non-existent. Comprehensive comparative study has not been really possible up to this point either on the lexicon or the grammar (though many excellent preliminary attempts have been made, particularly by Pinnow, N. Zide, A. Zide, Stampe, and Anderson), so a more thorough and comprehensive investigation into most historical linguistic issues in Munda also remains a goal for the future. This will aid considerably in the comparison of Munda with its Austroasiatic sister languages, as well as begin to gain a better understanding of the (linguistic) pre-history of South and Southeast Asia.

I will not give a brief introduction to the content of each chapter pointing out their strengths and weaknesses here as is typical in collected volumes such as this. I leave it to the interested readers to draw their own conclusions. Suffice it to say, that while this volume represents a monumental leap forward in the advancement of Munda linguistics, there is still much ground to cover before an adequate description of the languages will have been achieved. However, as the reader will see, these languages have much to contribute to our understanding of such diverse topics ranging from noun incorporation, transitivity, or the very nature of human language itself to the pre-history and contemporary sociolinguistics of India.

NOTES

- * Thanks to Norman Zide for some suggestions for this chapter. All errors remain my responsibility.
- 1 The sample sentences in the Tea District publications include such useful phrases as Sora *barasui badimaranji raptiurungtaaji pa?* 'do you have the chance to get other coolies' *pantagooleengan asoongdoong!* 'you must not defecate on the road!' and *unta badimaran unggadong* 'do not abuse that coolie!' (38–39, 89).

REFERENCES

- Abbi, Anvita (1992) *Reduplication in South Asian Languages*, Delhi: Allied Publishers.
- Anderson, Gregory D.S. (1999) 'A new classification of the Munda languages: Evidence from comparative verb morphology', Paper presented at 209th meeting of the American Oriental Society, Baltimore, MD.
- (2001) 'A new classification of Munda: Evidence from comparative verb morphology', *Indian Linguistics*, 62:27–42.

- Anderson, Gregory D.S. (2001) (this volume, Chapter 13) 'Gta?'.
 Anderson, Gregory D.S. and K.D. Harrison (this volume-a, Chapter 11) 'Remo (Bonda)'.
 — (this volume-b, Chapter 7) 'Sora'.
 Anderson, Gregory D.S., T. Osada, and K.D. Harrison (this volume, Chapter 5) 'Ho and the other Kherwarian languages'.
 Anderson, Gregory D.S. and F. Rau (this volume, Chapter 8) 'Gorum'.
 Banerjee, G.C. (1894) *An Introduction to the Kharia Language*, Calcutta: Bengal Secretariat Press.
 Burrows, Lionel (1915/1980) *Ho Grammar with Vocabulary: An Eastern Himalayan Dialect*, New Delhi: Cosmo Publication (originally in Calcutta: Catholic Orphan Press).
 Campbell, Alexander (1899, 1900, 1903) *A Santali-English Dictionary*, 3 parts, Pokhuria: Santal Mission Press.
 — (1905) *English-Santali; Index to Santali-English Dictionary*, Pokhuria: Santal Mission Press.
 Campbell, G. (1866) 'The ethnography of India (Appendix 2: Santal vocabulary)', *Journal of the Asiatic Society of Bengal*, 35(2):1–52.
 Coates, J.M. (1875) *Vocabulary of Seven Languages and Dialects of Chotanagpore*, Calcutta: Bengal Secretariat Press.
 Cole, Frederick T. (1879) 'Lists of words and phrases with their Santali equivalents', *Indian Antiquary*, 8:194–196.
 — (1906) *Santali Primer*. Benegaria: Santal Mission Press.
 Croke, William (1892) 'Vocabulary of the Korwa language', *Journal of the Royal Asiatic Society*, 61(1):125–128.
 Cust, R.N. (1878) *A Sketch of the Languages of the East Indies*, London: Trubner.
 — (1884) 'Grammatical note and vocabulary of the language of the Korku: A Kolarian tribe in Central India', *Journal of the Royal Asiatic Society of Great Britain and Ireland (new series)*, 16:164–179.
 Das, Rakal (1871) 'An introduction to the Mundari language', *Journal of the Asiatic Society of Bengal*, 40:46–66.
 De Smet, Rev. (1891) *Rudiments of a Mundari Grammar*, Calcutta: Catholic Orphan Press.
 Donegan, Patricia J. (1993) 'Rhythm and vocalic drift in Munda and Mon-Khmer', *Linguistics in the Tibeto-Burman Area*, 16(1):1–43.
 Donegan, Patricia J. and David Stampe (1983) 'Rhythm and the holistic organization of language structure', in *The Interplay of Phonology, Morphology, and Syntax*, Chicago: Chicago Linguistic Society: 337–353.
 Fernandez, Frank (1968) *A Grammatical Sketch of Remo: A Munda Language*, PhD dissertation, University of North Carolina.
 Ghosh, A. (this volume, Chapter 2) 'Santali'.
 Grierson, George A. (1906) *Linguistic Survey of India. Vol IV: Munda and Dravidian*, Calcutta: Office of the Superintendent of Government Printing.
 Griffiths, A. (this volume, Chapter 12) 'Gutob'.
 Hodgson, Brian H. (1848) 'The aborigines of Central India', *Journal of the Asiatic Society of Bengal*, 17(2):550–558.
 Hoffmann, Johan (1893) *A Mundari First Reader*, Calcutta: Bengal Secretariat Press.
 — (1903) *Mundari Grammar*, Calcutta: Bengal Secretariat Press.
 — (1909) *A Mundari Grammar with Exercises*, Calcutta: Catholic Orphan Press.
 — (1930–1950) *Encyclopedia Mundarica*, 15 volumes, Patna: Government Superintendent Printing.
 Hook, Peter. E. (1991) 'The compound verb in Munda: An areal overview', *Language Sciences*, 13(2):181–195.

- Kobayashi, M. and G. Murmu (this volume) 'Kera? Mundari'.
- Kuiper, F.B.J. (1948) *Proto-Munda Words in Sanskrit*, Amsterdam: Noord-Hollandsche Maatschappij.
- Masica, Colin (1976) *Defining a Linguistic Area*, Chicago: University of Chicago Press.
- Mithun, Marianne (1984) 'The evolution of noun incorporation', *Language*, 4(60): 847–894.
- Nottrott, Alfred (1905) *Kol Grammar*, (translation by A.G.F. Fitzgerald and Paul Wagner), Ranchi: G.E.L. Mission Press (original edition 1882, *Grammatik der Kolh-Sprache*, Gütersloh: C. Bertelsmann).
- Odden, David (1987) 'Arguments against the vowel plane in Gta?', *Linguistic Inquiry*, 18(3):523–529.
- Osada, T. (this volume, Chapter 3) 'Mundari'.
- Patnaik, M. (this volume, Chapter 9) 'Juang'.
- Pendercast, M.H. (1881) 'Short vocabulary of the Savara language', *Journal of the Royal Asiatic Society*, 13:425–428.
- Peterson, J. (this volume, Chapter 10) 'Kharia'.
- Phillips, Jeremiah (1845) *A Santali Primer*, Calcutta: School Book Society.
- (1852) *An Introduction to the Santali Language*, Calcutta: School Book Society.
- Pinnow, Heinz-Jürgen (1959) *Versuch einer historischen Lautlehre der Kharia-Sprache*, Wiesbaden: Otto Harrassowitz.
- (1960) *Beiträge zur Kenntnis der Juang-Sprache*, Unpublished manuscript, Berlin.
- (1966) 'A comparative study of the verb in the Munda languages', in N. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics* (Indo-Iranian Monographs vol. 5), The Hague: Mouton, 96–193.
- (1972) 'Schrift und Sprache Lako Bodras', *Anthropos*, 67:822–857.
- Pinnow, Heinz-Jürgen, J. Kullu and P. Kerketta (1965) 'Der Wertiger und andere Texte', *Indo-Iranian Journal*, 9(1):32–68.
- Puxley, Edward L. (1868) *A Vocabulary of the Santali Language*, London: W.M. Watts.
- Ramamurti, G.V. (1931) *A Manual of the So:ra: (or Savara) Language*, Madras: Government Press.
- Ramaswami, N. (1992) *Bhumij Grammar*, Mysore: Central Institute of Indian Languages.
- Ross, Malcolm (1996) 'On the genetic affiliations of the Oceanic Languages of Irian Jaya', *Oceanic Linguistics*, 35(2):258–271.
- Sadock, Jerrold (1991) *Autolexical Syntax*, Chicago: University of Chicago Press.
- Schmidt, Pater P.W. (1906) 'Die Mon-Khmer Völker: Ein Bindeglied zwischen Völkern Zentralsiens und Austronesiens', *Archiv der Anthropologie, Braunschweiger*, 33[n.s. 5]:59–109.
- Skrefsrud, Lars Olsen (1873) *A Grammar of the Santhali Language*, Benares: Medical Hall Press.
- Starosta, Stanley (1967) *Sora Syntax: A Generative Approach to a Munda language*, PhD dissertation, University of Wisconsin.
- Tickell, S.R. (1840a) 'Grammatical Construction of the Ho language', *Journal of the Asiatic Society of Bengal*, 9:997–1007.
- (1840b) 'Vocabulary of the Ho Language', *Journal of the Asiatic Society of Bengal*, 9:1063–1090.
- Voysey, H.W. (1844) 'Vocabulary of Goand and Cole words', *Journal of the Asiatic Society of Bengal*, 13(1): 19–23.
- Zide, N.H. (1969) 'Munda and non-Munda Austroasiatic languages', in Thomas Sebeok (ed.) *Current Trends in Linguistics*, The Hague: Mouton, 5:411–430.

- Zide, N.H. (1996) 'Scripts for Munda languages', in Peter Daniel and William Bright (eds) *The World's Writing Systems*, New York: Oxford University Press, 612–618.
- (2000) 'Three Munda Scripts', *Linguistics of the Tibeto-Burman Area*, 22:199–232.
- (this volume-a, Chapter 6) 'Korku'.
- (this volume-b, Chapter 14) 'On Nihali'.
- Zide N.H. and Stampe, D.L. (1968) 'The position of Kharia-Juang in the Munda family' in B. Krishnamurti (ed.) *Studies in Indian Linguistics* (Professor M.B. Emeneau's śaṣṭipurṭi volume), Poona: Centre of Advanced Study in Linguistics, Deccan College, 370–377.
- Zide, N.H. and A.R.K. Zide (1976) 'Proto-Munda cultural vocabulary: Evidence or early agriculture' in P.N. Jenner, C. Thompson, and Stanley Starosta (eds) *Austroasiatic Studies, Part II*, Honolulu: University of Hawaii Press, 1295–1334.

CHAPTER TWO

SANTALI

Arun Ghosh

1 THE LANGUAGE AND ITS SPEAKERS

1.1 Linguistic type

Santali is a Munda language with a suffixing, agglutinating, and a basic SOV structure. The most notable characteristic of the language is the weak distinction between noun and verb, the addition of verbal suffixes that will turn any lexeme into a verb, and case-markers, enclitic definitives, and number markers into a noun. In the present study, therefore, the word classes have been postulated on the basis of morpho-syntactic criteria rather than lexical criteria alone.

The consonant inventory consists of five categories of plosives distinguished with regard to five places of articulation (labial, alveolar, post-alveolar, palatal, and velar), a nasal corresponding to each, one trill, one flap, one lateral, two fricatives, and two semivowels. There are eight vowels, without any significant length distinction. Words are predominantly dissyllabic with stress on the second syllable in dissyllabic and multisyllabic words, and on the root vowel if it is monosyllabic. Diphthongs do occur with rising and falling varieties. Consonant clusters are rare. As to the syllable structure the language has a predilection for CV structure closely followed by CVC. One notable feature of the language is that nasals can form the nucleus, found nowhere else among Indian languages apart from Munda.

Grammatically speaking there are classes of nominals (noun and pronoun), verb, demonstrative, adjective, adverb, and particle. The language has a gender agreement system. Nouns and demonstratives show a two-way gender distinction – animate and inanimate. Furthermore, some nouns show an overt masculine–feminine distinction in their forms. Nouns are declined in three numbers – singular, dual, and plural. Pronouns show forms for all the three numbers in all three persons with an inclusive–exclusive distinction in the first person. The third personal pronoun is actually derived from the distal demonstrative root. Use of the anaphoric third personal form is not very common. The demonstrative is divided into three classes – simple, interrogative, and indefinite. Each class shows an animate–inanimate distinction. Case is affected by the use of suffixes and postpositions. Whereas sociative and ablative are marked by postpositions, the genitive, comitative, instrumental-locative, allative, and locative are marked by suffixes. Nominative is marked on the verb as transitive subject, intransitive subject, and transitive object. Case is divided into core and peripheral. Whereas the core agrees with the verb in the form of pronominal argument (subject, direct object, and indirect object), the peripheral is marked on the nominals and does not have any agreement with the verb. The root morphemes, derived forms, and phrases which serve as attributes in endocentric attributive constructions and as predicate complements in subject–predicate complement sentences are grouped as adjectives. Adjectives do not agree in number or in gender, except

in a few borrowed items, and in the case of qualified nouns. Adjectives are not inflected to show degrees of comparison. To obtain degrees of comparison postpositions are added to the word with which something is to be compared. Numerals are generally found as quantifiers combined with classifiers. The quantifiers are used singly when they are used in enumerating human beings in the indefinite, whereas in counting human beings in the definite and non-human beings and inanimate objects, the quantifiers are used combined with classifiers. Adverbs are modifiers irrespective of the position they occupy in the sentence.

The verb is defined by the fact that it takes TAM suffixes with or without the markers for the active and neutral, personal terminations, the marker for finiteness of action, and the gerundial suffix. Words which denote actions, events, or conditions are not the only ones treated as verbs, but any word is here treated as such provided that it takes verbal suffixes. For finiteness of action it requires the finite marker /*al*/.

The TAM suffixes have two forms depending on whether it is active or middle – one used on the verb in transitive constructions with the other being employed on verbs in intransitive constructions. The ambitransitive roots take on active as well as middle suffixes. Verbal roots fall into three classes: intransitive only, transitive only, and ambitransitive.

Verbal stems are divided into causative, reciprocal, benefactive, passive/reflexive, mediopassive, iterative/intensive, and compound, and morphological processes are employed to form stems including prefixation, infixation, suffixation, reduplication, and compounding. Except the particles /*kanl*/ and /*taēkanl*/ for the present and past tenses respectively there are no separate tense markers. There is a concept of time dimension in the present, past, and future tenses and the aspect suffixes along with their own functions denote time as well.

On the syntactic level, generally the head follows its determiner. In the simple sentence then word order is SOV. In complex sentences, conjunctions, coordinate clauses, and the subordinate clauses perform nominal or attributive function. Sentence modification is achieved through a set of particles. Polar questions can be shown by a marked intonation pattern.

1.2 Name of the language and the tribe

The name Santali, the language, is derived from the ethnic name *Sāotal*, the Anglicized version being Santal. From *Sāotal* the neighbouring non-Santals use the name *Sāotali* for their language. The Santals themselves call their language *hɔɽ* or *hɔɽ rɔɽ*. Being more analytical they attach *rɔɽ* ‘language, speech’ to *hɔɽ* ‘Santal man’, hence ‘language of the Santals’. The language is also sometimes returned under the name *mājhi bhasa* ‘language of the Majhis’. In North Bengal, in the districts of Murshidabad, Malda, Dinajpur, etc. the language is known as *janli* or *pahaṛia*. In South Bengal, particularly in the districts of 24-Parganas (north and south) and Howrah, and in Orissa the language is known as *ṭhar*. In Bihar it is known as *parsi* ‘foreign’. The name *Sāotal*, is derived, in turn, from *Sāmanta-pāla*¹ ‘dweller of the frontier’ and is used to name the tribe by the non-Santals, particularly the Bengalees. L.O. Skrefsrud² tried to derive the name from *Sāot*, a place in Midnapur in West Bengal where the Santals were supposed to have been settled in remote antiquity. There is still a place called *Sāot* in Midnapur district where Santal habitation is common. The Santals call themselves *hɔɽ* ‘man’ and the name *Santal* is only used by those who are Christianized. From the term *majjhi* or *mājhi* ‘village headman’ they

also call themselves *mañji* or *māñji* when asked about their caste. Being the oldest ethnic stock in India they are also known as *ādivāsi* ‘those who have been living in the land from the beginning.’ S.K. Chatterji³ attaches great importance to this term for them remembering the contribution of the Santals in the evolution of Indian life and culture.

1.3 Genetic affiliation of Santali

Santali belongs to the Kherwari group of the North Munda sub-family of the Munda family which is, in turn, a section of the eastern group of the great Austroasiatic family of languages. Pinnow (1959)⁴ offered a classification of the Austroasiatic languages into two main branches – West Obergruppe, consisting of the Munda languages and Naha-li⁵ and Ost Obergruppe including the rest of the Austroasiatic languages. Norman H. Zide divides the Munda languages into two main groups – South Munda (SM) and North Munda (NM). The eastern (i.e. Kherwari) and the western (i.e. Korku) form the North Munda branch. Santali is separated off as a special branch of Kherwari. South Munda, on the other hand, consists of Koraput Munda (KM) and Central Munda (CM). Koraput Munda consists of Sora-Gorum (SG) group and Gutob-Remo-Gta? (GRG). Central Munda consists of Kharia and Juang.⁶ Bhattacharya⁷ conforming to the same geographical classification puts forward another classification based on morphological criteria. He proposes a two-way division – Lower Munda, consisting of the three extreme southern Munda languages, *ḍiḍeyi* (Gta?), Bonda (Remo), and Gutob spoken in Koraput and Malkangiri districts of Orissa, bordering Andhra Pradesh with Upper Munda covering the rest. As the three southernmost languages included in Lower Munda do not show pronominal incorporation in the verb, differing from the rest of the Munda languages in respect to genitive marker and by not having dual number, they are branched off from the rest of the Munda languages.⁸ But whatever the situation is, the position of Santali remains the same. It has been branched off as a special member of the northeastern group possibly because it has reached a higher stage of development than any other language of the same group and also preserves the peculiar linguistic features of Munda more faithfully than the rest.

1.4 Number of speakers and locale of concentration

The total population of the Santals is 6,050,000⁹ in India, Bangladesh, Bhutan, and Nepal, of which 5,959,000 are in India, 157,000 in Bangladesh, and 33,332 in Nepal. The exact figure for Bhutan is not available. The main concentration of the Santals is in India with scattered settlements in Bangladesh, Nepal, and Bhutan. They are also found all over the world wherever they have found job opportunities. The classified data of 2001 census is for all the states of India, except for West Bengal, which is not available yet. In West Bengal it is estimated to be 2,280,540.

The distribution of the Santals in different states of India, according to the 1981 census, is shown in Table 2.1.

As it appears from the census data the Santal population is spread over a large strip of land covering almost the whole of India. The most compact area of concentration is the western part of West Bengal, the southern portion of erstwhile Bihar, now Jharkhand, the areas of Bihar adjacent to Jharkhand, the northeastern districts of

TABLE 2.1: DISTRIBUTION OF SANTALI SPEAKERS IN INDIA

State	Total	Male	Female
Andhra Pradesh	50	30	20
Andaman & Nicobar Islands	7	3	4
Arunachal Pradesh	484	282	202
Bihar (including Jharkhand)	2,161,032	1,087,820	1,073,212
Chandigarh	10	8	2
Delhi	126	69	57
Goa, Daman & Diu	1	1	—
Gujarat	40	30	10
Haryana	10	5	5
Himachal Pradesh	7	7	—
Jammu & Kashmir	28	26	2
Karnataka	45	30	15
Madhya Pradesh	715	445	270
Maharashtra	110	64	46
Manipur	351	348	3
Meghalaya	212	199	13
Mizoram	2,301	2,278	23
Nagaland	1,100	1,080	20
Orissa	529,574	264,854	264,720
Punjab	10	10	—
Rajasthan	52	37	15
Sikkim	16	14	2
Tripura	3,518	1,812	1,706
Uttar Pradesh	282	257	25
West Bengal	1,632,440	822,973	809,467

Orissa and Assam. The data of 1981 did not give figures for Assam, as census could not be conducted due to political turmoil. According to the 1971 census the figure goes to 86,303.¹⁰ So far as the Jharkhand is concerned the figure is not available, as the state was not formed during the last census. The figure for Bihar in Table 2.1 can be read as that of Jharkhand, as the state of Jharkhand was constituted of the southern portion of Bihar which houses almost the whole of the tribes. If the data are represented in a map the areas just described will form a compact area, comprising western parts of the districts of Birbhum and Burdwan, the Sadar subdivision¹¹ of Bankura, Jhargram subdivision of West Midnapur, Purulia in West Bengal, the whole of Jharkhand, especially the Santal Parganas, Hazaribagh, Singbhum and Dhalbhum districts, south of Bhagalpur and Munghyr in Bihar, and Balasore, Mayurbhanj and Keonjhar in Orissa. In Assam the Santal settlements are confined to the tea gardens only, where they were brought as labourers. The process of shifting is now transformed into the process of migration for job opportunity in the tea gardens. They are also sparsely distributed in the northern districts of West Bengal,¹² in the districts of South 24-Parganas, Hooghly, Howrah and so on in south Bengal where they are employed as day-labourers. The Santals settled at Rajsahi, Rangpur, and Chattagram in Bangladesh are said to be immigrants, migrated from the other part of the river Padma, that is, India at different periods of history. But considering their concentration in Nepal and Bhutan, and going by the different theories of their origin and migration it may be conjectured that they are sons of the soil and that

others found further west in India migrated for better living.¹³ The hypothesis may be strengthened by their presence in Nepal and Bhutan.

The Santals are the most numerous among the tribes who speak Munda. In the western fringe of West Bengal, north Orissa and Jharkhand, normally, they muster very strong. This helps them maintain group solidarity and preserve their language and culture much better than elsewhere. The Santals now remaining in other places are nothing but scattered masses floating here and there, and in the process they are all but melted with other dominant cultures in the region. The greater part of their substance has already commingled in the fluid around them, the remainder is saturated with it, and it is only in the very kernel and inner centre of the largest lumps that something like the pure original substance is to be sought.

1.5 Bilingualism, processes of Aryanization and nativization

The proportion of the population retaining the language decreases as the tribe spreads over to industrial areas, nearer to towns and cities. It is also affected by the spread of education among the younger generation. The more they are educated, the more they are oblivious to their own language and drawn nearer to neighbouring languages of the Indo-Aryan group, say, Bengali in West Bengal, Hindi in Jharkhand and Bihar, Oriya in Orissa, and Assamese in Assam. There are, of course, two reasons behind this attitude; first, living by the side of the majority communities and in mixed localities they have to be bilinguals, as there is no other alternative. Over a period of time they are more with the majority language than their own. Being constantly in an alien system they automatically adapt. Second, retention of a language also has a socio-economic background. In the job market the language has less potential than the neighbouring majority languages. So far as the prestige factor is concerned, for them it has none, since until recently the language was not recognized in the eighth schedule of the Constitution of India. So there were and still are ample reasons for leaving their own language and welcoming the neighbouring one. In West Bengal, as a whole, 55% of the Santals speak their own language but in the district of 24-Parganas the proportion falls to 10%. The total number of bilinguals among the Santals is 1,501,638, being 34.66% of the total.¹⁴ Whereas the national average of bilingualism is 13.34%, it is more than 30% in the case of tribal population. Obviously, necessity for knowing the common language of the area is very great in the case of tribes who speak a language which is not a majority one. As already said, being a minority-language community, they cannot help speaking the dominant language of the region where they stay. In every-day interaction, they have to come across a group of people who belong to a different speech community, namely the Bengali-speech community in West Bengal, the Mundari and the Hindi-speech community in Jharkhand, the Hindi-speech community in Bihar, the Oriya-speech community in Orissa, and the Assamese-speech community in Assam. It is very common then that the local languages will influence their own language in some way or another. This way Santali is, to some extent, modified in phonology, morphology, and lexicon. Syntax is not much altered. The lexical elements of the dominant languages are taken into Santali either through a process of nativization or just as they are. The connotation of some lexical elements is changed under the influence of the neighbouring languages.

New postpositions and particles are borrowed and nativized following the phonotactic rules of the language. The extent of influence on Santali grammar and lexicon may be clarified with certain examples.

Being in contact with the neighbouring world they start adopting its language system along with many words being Santali-equivalent; for example, *lədu* 'sweet-meat', *mīthəi* 'sweet', *dal* 'beans', *caole* 'rice', *pənəhi* 'shoes' have been incorporated into their system. As professions such as shoemaker, blacksmith, and carter do not belong to the Santal society, the corresponding terms like *muci* 'shoemaker', *kamar* 'blacksmith', and *gaḍvan* 'carter' are borrowed from the neighbouring language system. Similarly, for trades not belonging to their system, they have borrowed trade-related terms from the neighbouring language system; for example, *mal* 'goods', *cij* 'sample', *asbab* 'furniture', *bajar* 'market', *khərca* 'expenditure', and so on. In time calculation the Aryan influence is noticeable; thus *ghəri* 'a while', *din* 'day', *bəchər* 'year', *cirəkāl* 'a long time'. Even though they have their own village administration terms like *maṅjhi* 'village headman' (also *mājhi* 'id.')

(cf. Skt. *madhya*), and *pançayat* 'village council' are borrowed by them from the neighbouring Indo-Aryan languages to accommodate themselves in the government-controlled Panchayat system. Even in the household Aryan influence is noticeable. The inner apartment is *bhitər* or *bhitri*, door is *duər*. Terms like *pukhri* 'pond', *bande* 'id.' are of Aryan origin. Normally the kinship terms and terms for body parts are kept intact. So are the numerals. But here, the Aryan influence is so strong that all three areas are affected to some extent, giving way to foreign elements. Thus, wife is *bəhu*, nephew *bhəgnə*, niece *bhəgni*, *bhəi* is 'brother', hand is *hat*, one is *ək*. The present generation can only utter their numerals from one through five or six, then counting goes on in Aryan numerals. Terms for sentimental feelings like *maya* 'affection', *daya* 'pity', *laj* 'shame' are also borrowed. The first three ordinal numbers *pəhillpoilo* 'first', *dəsar* 'second', and *tesar* 'third' are also borrowed from Indo-Aryan.

On the grammatical level, too, Aryan influence is noticeable. The distinction between masculine and feminine, though attested to in a few examples, is framed on the analogy of the Aryan system; the masculine noun ending in /-a/ and feminine in /-i/: thus *kala* 'deaf' masculine, and *kəli* 'id' feminine, *koṅka* 'mad' masc. *kuṅki* 'id', feminine, *koṅa* 'boy': *kuṅi* 'girl'. A good number of postpositions like *ləgit* 'for', *səṅge* 'alongwith', *upər* 'above', *bhitər/bhitri* 'in', *sathə* 'along with' are borrowed from either Bengali or Hindi. A good number of particles is also borrowed from the same source: *jodiljudi* 'if', *jəmən* 'so that', *tahle* 'then/for that', *ar* 'and', *təkhən* 'then', *təbe* 'then' and so on. In conjugation also some roots like *cal* 'go', *bujh* 'understand', etc. are borrowed from the Aryan languages like Bengali, Hindi, or Oriya, but with modifications: /-aol/ or /-əul/ is added to the root resulting in *calao* or *bujhəu*. When these roots are conjugated the indigenous suffixes are used. TAM suffixes are never borrowed. Aryan nouns are also used as verbs but are so nativized that they fit well in the Santali system. In *bəhu-ad-e-a-ṅ* 'I gave him a wife', *bidə-ka-e-a-ko* 'they sent him off', the Santali suffixes are pitted against Aryan nouns. The non-native elements making their way into the vocabulary are readily nativized by giving a phonetic twist, and thus they are well incorporated in the language. In the non-native elements, following the native pattern the vowel /a/ is neutralized if it is preceded or followed by /i/ or /u/: thus *bhitər* for *bhitər* or *bhetər* 'inside', *kəli* for *kali* 'goddess Kali', *kəmi* for *kam* 'work', *juən* for *juan* or *joan* 'youth', *bəhu* for *bahu* 'wife', *pəchim* for *pachim* 'west', *cəukidar* for *caukidar* 'village guard', *upər* for *upar*

‘on the top, above’, *sutəm* for *suta* ‘thread’, *məndir* for *mandir* ‘temple’, *ləḍu* for *laḍḍu* ‘sweetmeat’, *səḍi* for *saḍi* ‘cloth’, *rəni* for *rani* ‘queen’, *ṭhāi* for *ṭhāi* ‘place’, *utər* for *uttar* ‘north’, *rəi* for *rai* ‘mustard’, *ləgit* for *lagi* ‘for’, *məmi* for *mami* ‘maternal aunt’, etc. Borrowed verb roots which are nativized by adding /-aol/ or /-əul/ like *paḍao* ‘to fall’, *calao* ‘go’, *jəḍao* ‘to link’, *lagao* ‘cultivate’ the borrowed words or grammatical elements ending in a vowel also add a checked consonant at the end: thus – *ləgit* for *lagi*, *ṭak* for *ṭa* definite article.

1.6 Dialectal differences

Santali being scattered in different places is subject to differences in phonology, morphology, and lexicon. Considering that there has been no settled standard, R.N. Cust mentioned four dialects which, according to him, was not improbable.¹⁵ In some reports the number of dialects is even more.¹⁶ Although four or six is a bit ambitious, a line of demarcation is emerging gradually between the variety spoken in the districts of East and West Midnapur, Purulia and the southern portion of Bankura (comprising Khatra, Ranibandh, Raipur, Taldangra, and Simlapal blocks) in West Bengal, Balasore, Mayurbhanj, and Keonjhar in Orissa and that spoken in the northern portion of the district of Bankura (comprising Chatna and Saltora blocks), in Birbhum, Malda, Dinajpur north and south, Murshidabad, Cochbehar, and Jalpaiguri in West Bengal, Santal Parganas, Dumka, Singbhum, and Dhalbhum in Jharkhand, and Munghyr and Bhagalpur in Bihar. Campbell¹⁷ mentioned two varieties like Northern and Southern, although he did not enter into the details. During our field trip we checked the varieties in detail and found two varieties in the areas named. Therefore Campbell’s distinction of Southern versus Northern holds.¹⁸ Differences between the two dialects – Northern (henceforth NS) and Southern (henceforth SS) are as follows.

1.6.1 Phonology

In the SS the pronunciation of [ə] is changing and gradually being replaced by [a], especially in the pronunciation of the younger generation, whereas in the NS it remains intact.

Loss of nasal and its compensation by nasalizing the preceding vowel is the rule of the SS, whereas in the NS the nasal remains intact.

Compare:

SS	NS	
<i>pōṭ</i>	<i>ponḍ</i>	‘white’
<i>mājhi</i>	<i>maṅjhi</i>	‘village headman’
<i>āṛia</i>	<i>əṅḍiə</i>	‘male cow’
<i>bhīdar</i>	<i>bhindər</i>	‘to fall’
<i>aḍoṅ</i>	<i>andok</i>	‘to come out’
<i>cādo</i>	<i>cando</i>	‘sun’
<i>nṅṛe</i>	<i>nəṅḍe</i>	‘here’
<i>ṅṛe</i>	<i>əṅḍe</i>	‘there’
<i>hṅṛe</i>	<i>hande</i>	‘there yonder’

Pronunciation of post-alveolar stop as flap is a common feature of the SS as opposed to the NS where it is pronounced as a stop.

Compare:

SS	NS	
<i>põɽ</i>	<i>ponɽ</i>	‘white’
<i>baɽa</i>	<i>baɽa</i>	‘to know’
<i>huɽiŋ</i>	<i>huɽiŋ</i>	‘younger’
<i>geɽɛ</i>	<i>geɽɛ</i>	‘swan’
<i>bheɽa</i>	<i>bheɽa</i>	‘ram’

There is preference for nasals to checked consonants in case of genitive suffixes for inanimate, enclitic definitive and allative suffix /-sec'/ in the SS as opposed to the NS which preserves the checked ones.

Compare:

SS	NS	
<i>/-aŋ/</i>	<i>/-ak'/</i>	‘genitive suffix for inanimate’
<i>/-reaŋ/</i>	<i>/-reak'/</i>	‘genitive suffix for inanimate’
<i>/ɽaŋ/</i>	<i>/ɽak'/</i>	enclitic definitive
<i>/sen/</i>	<i>/sec'/</i>	‘towards’

In the SS, sometimes there is a tendency, also shared by the younger generation of the NS to pronounce /e/ as [i] and /o/ as [u]. In certain examples like *abin* ‘you two’ (cf. *aben* in NS), *unku* ‘they two’ (cf. *onko* in NS), *ni* ‘this’ (cf. *ne* in NS) the pronunciation has been established. Among the elders and the enlightened, in the NS /e/ and /o/ are pronounced as half-close front and back vowels respectively.

1.6.2 Morphology

In morphology SS has certain features which the NS does not possess.

In the genitive case in the SS two suffixes are found when the governed noun is animate, one for the singular and the other for plural. /-ic'/ is used when the noun governed is in singular and /-ren/ when it is plural. In the NS /-ren/ is for both singular and plural.

Example:

SS		NS
<i>ɽi-ic' mɛɽɔm</i> (SG.)	‘my goat’	<i>ɽi-ren mɛɽɔm</i> ‘my goat/ goats’(SG. and PL.)
<i>ɽi-ren mɛɽɔm</i>	‘my goats’	

The causative stem forming suffix is /-hɔcɔ/ in the SS and /-ocol/ in the NS, although this is more phonological than morphological.

1.6.3 Vocabulary

In the lexicon SS and NS are somewhat different, initiated by borrowing from the neighbouring languages. The local borrowings in the two dialects are so high that sometimes one appears to be unintelligible to the other. In certain cases the usage

is also different. Terms like *eyga* ‘mother’, *era* ‘wife’, *herel* ‘husband’, etc. are treated as vulgar in the SS and are almost obsolete, whereas in the NS the terms are regularly used without any pejorative sense. In the SS the corresponding terms for the relations are *ayo*, *bāhu*, and *jāvāy* are used for the kin relations respectively. Other examples are, for ‘cloth’ SS has *lugri* and NS has *kicric*, shirt is *dotɔ* in the SS and *jama* in the NS.

1.7 Orthography

Santali was first written down in Roman script by European missionaries. It was used then for translating the Bible into Santali, for writing grammars, and also for folk-tales and past history of the people. At that time many scripts, at least four, were being used for writing – like Devanagari, Bengali, Oriya and Roman. From the middle of the nineteenth century through the third quarter of the twentieth century the language was mostly written in Roman script. Even the people themselves used to write in Roman only for obvious reasons. Regional or Devanagari scripts came to be used much later. After Pandit Raghunath Murmu developed a script of their own, namely Ol Ciki, movements started for its recognition. In the 1970s many schools were started to impart training in Ol Ciki with initiative of a group of the Santals. A group of writers started writing in Ol Ciki. Magazines, journals, and newspapers began to be published. To the present date a consensus has not been reached as to which script is to be adopted.

A good many writings, like *Kherwal Bāṅso Dhāṛom Pūthi* by Ramdas Tudū, *Ol Dāḥ Onārhē* and *Liṭā Godet* by Ramchand Murmu, *Darega Dhon* and *Bidhu Chandan* by Raghunath Murmu, *Bhurka Ipil*, *Bidak Bela*, etc. by Saradaprasad Kisku have been published in Bengali script. Even Ramdas Tudū and Raghunath Murmu who had developed scripts for the language published their own books in the language using Bengali script. A good number of books were also written in Devanagari and Roman scripts – Stephen H. Murmu’s *Hāṛ Bapla Puthi*, G.C. Tudū’s *Chandmala* (a collection of poems) and *Bakhra* are written in Roman. Since 1965 The Santali Literary and Cultural Society has been rendering remarkable service for the development of the Santali language and literature by means of publishing Santali books and journals in Roman script. Narayan Soren, Balkishore Baskey, Bhagbat Murmu, Babulal Murmu, and Manikchand Hansda are well-known writers of Santali literature. Many of their writings are published in Devanagari script. However, Manikchand Hansda and S.D. Besra wrote in Roman script. Some of the most important Santali magazines are *Pera Hāṛ* (in Roman) since 1922, *Marshal Tabon* (in Roman) since 1946, *Hāṛ Sombad* (in Devanagari) since 1947, *Pachim Bangla* (in Bengali script) since 1956, *Jug Sirijol* (in Roman) since 1971. Apart from these many other periodic journals are being published from West Bengal, Jharkhand, Bihar, Orissa, and Assam. *Tetre* from West Bengal and *Upel* from Assam need mention. *Tetre* is published in Bengali script and *Upel* in Roman. Two monthly magazines are published from Bangladesh – *Aboak’ kurumuTureak’ Kurai* and *GoDet*, from Rongpur and Dhaka respectively. Both are in Roman script and serve as a link among those of Bangladesh, India, and Nepal.

Santali has recently been recognized by the Government of India as an official language, and it has been included in the Eighth Schedule of the Constitution of India. The Ol Ciki script has also been recognized by the Government of West Bengal as a medium of publication of Santali textbooks. It has already been included

in the curricula of the Universities. Whereas two of the Universities, Burdwan and Vidyasagar, are flexible in their approach (i.e. teaching, framing of question papers, and writing the same, will be in Roman, Bengali, or Ol Ciki until a particular script is accepted on consensus), the Vishva Bharati University imparts through Roman script. Some of the Universities of Bihar and Jharkhand have introduced Santali as a modern Indian language and literature in their syllabi, liberally accepting both Roman and Devanagari. It is yet to be seen which script is accepted finally, as a good percentage of population is in favour of regional scripts, Roman, or Devanagari. They are opposing introduction of Ol Ciki. Seminars and conferences are being organized in the government and on the organizational level to resolve the issue. But the debate continues.

2 PHONOLOGY

Santali is a Munda language with an elaborate system of vowels and consonants when compared with other languages of the same group (cf. Bhattacharya 1954, 1975, Stampe 1963, Matson 1964, Biligiri 1965, Zide 1965, Fernandez 1967, Rao 1982, Osada 1992, Nagaraja 1999, Ghosh 2003). There are eight vowel phonemes with nasalized counterparts (except /e/ and /o/). But the nasal ones are used rarely as compared with their oral counterparts. Length is not phonemic, although according to Bodding (1922, 1929) all vowels may be short or long. It has post-alveolar consonants and aspirated stops. It also has checked consonants occurring finally. Although most of the aspirates are found in loanwords, in some words of the native system the aspirates do occur but their origin is doubtful. In some cases aspirates originate through syllable reduction like *aji+hanhar* = *ajhmar* ('husband/wife's elder sister'). Although aspiration, in Pinnow's version (1959), was absent in Proto-Munda, it has spread into the phonemic system of Santali. The status of the checked consonants is clear – these are final allophones of the corresponding stop consonants except for post-alveolar, which has none. The checked plosives (labial, alveolar, palatal, and velar) involve a closure in the glottal cavity checking release of the air. A weak release is audible, however, if it is immediately followed by the corresponding unchecked consonant. Except post-alveolar all plosives have nasals in the phonemic level, and occurrence of the post-alveolar nasal is predictable. There is post-alveolar non-aspirated flap without having its aspirated counterpart, as mentioned in some literature (Neukom 2001:5).

2.1 Vowels

Table 2.2 lists Santali vowels.

TABLE 2.2: SANTALI VOWELS

	Front		Central		Back	
High	<i>i</i>	<i>ĩ</i>			<i>u</i>	<i>ũ</i>
Mid-high	<i>e</i>		<i>ə</i>	<i>ẽ</i>	<i>o</i>	
Mid-low	<i>ɛ</i>	<i>ẽ</i>			<i>ɔ</i>	<i>õ</i>
Low			<i>a</i>	<i>ã</i>		

2.1.1 Vocalic allophony

Note that /ə/ is found mostly in the environment preceded or followed by /i/ or /u/: for example, ədʒi 'very much', əgu 'bring', niə 'this very', inə 'that very', gəi 'cow', dəi 'elder sister', gədʒi 'cart', tuə 'orphan', where the occurrence is predictable. In some cases there is no such predictable environment, like əd 'take possession of', əs int. (bullock or buffalo), bəd 'high lying rice-field', bəd 'flood', rəʃ 'a tune', dʒər 'branch of tree', səbrə 'tasteless', tʃəyə 'leave alone', dəblə 'too broad', dətrə 'a certain plant', gəhlə 'low pitched', etc. In a number of examples like ən 'law' (cf. Bengali/Hindi, ain 'id.'), əd 'origin' (cf. Bengali/Hindi, adi 'original'), əg 'fire' (cf. Sanskrit agni, Bengali agun, Hindi ag 'id.'), əglə 'forthcoming' (cf. Hindi agla~agila 'id.'), ək 'sugar-cane' (cf. Skt. ikṣu 'id.'), əs 'scales of fish' (cf. Hindi āis 'id.'), rəs 'heap' (cf. Bengali raʃi 'id.'), rət 'night' (cf. Hindi, rait 'night'), rəskə 'joy' (cf. Hindi, rasika 'id.'), dət 'run' (cf. Hindi dauṛ, Bengali dauṛ 'id.'), jət 'caste' (cf. Bengali jati 'id.'), pəc' 'five' (cf. Skt pañca 'id.'), dəl 'pulse' (cf. Hindi dail 'id.'), dən 'witch' (cf. Hindi dain, Bengali daini 'id.'), ghət 'fault, sin' (cf. Hindi ghaṭi 'id.'), bhəgnə 'nephew' (cf. Bengali bhagina~bhagna 'id.'), bəs 'age' (cf. Hindi bais, Bengali bɔyes 'id.');

although there is no predictable environment on the surface level, the occurrence of /ə/ can be justified by comparing it with the Indo-Aryan words where there is [ɪ] or [u] in the neighbouring syllable. In some of the examples like pəc', bəs, etc. /ə/ occurs because of the neighbouring palatal consonant or semivowel, [ç] and [y], respectively. Bodding (1922, 1929) referred to at least three 'resultant' vowels: ə, e, and o. Besides ə (Bodding's ə) no other variety, either e or o, is attested to in our data.

2.1.2 Distribution chart

Distribution chart for the vowels is provided as follows:

Vowel	Initial	Medial	Final
/i/	+	+	+
/e/	+	+	+
/ɛ/	+	+	+
/a/	+	+	+
/ə/	+	+	+
/ɔ/	+	+	+
/o/	+	+	+
/u/	+	+	+

2.1.3 Contrast pairs

Minimal and subminimal pairs are provided for illustration.

Vowels:

/i/:/u/:

il	'feather'	ul	'mango'
birij	'unpaid'	burum	'lie down'
si	'plough'	su	'hiss'

<i>li:/el/:</i>			
<i>ir</i>	'reap'	<i>er</i>	INTR.(call)
<i>bir</i>	'forest'	<i>ber</i>	'sun'
<i>ji</i>	'smell'	<i>je</i>	CONJ.part.
<i>el:/el/:</i>			
<i>egar</i>	'undertake'	<i>eger</i>	'scold'
<i>hel</i>	'time'	<i>hel</i>	'bind together'
<i>aɽe</i>	'edge'	<i>aɽe</i>	'vicinity'
<i>el:/ol/:</i>			
<i>eɽak'</i>	'other'	<i>oɽak'</i>	'uncover'
<i>jel</i>	'deer'	<i>jol</i>	'fire'
<i>je</i>	INDEF.PART.	<i>jo</i>	'taste'
<i>el:/al/:</i>			
<i>em</i>	'give'	<i>am</i>	2 SG. PR.
<i>met'</i>	'eye'	<i>mat'</i>	'bamboo'
<i>baɽe</i>	'please'	<i>baɽa</i>	'frequently'
<i>el:/ol/:</i>			
<i>er</i>	'sow'	<i>ɔr</i>	'pull'
<i>hel</i>	'bind together'	<i>hol</i>	'quickly'
<i>sɛ</i>	'or'	<i>sɔ</i>	'smell'
<i>al:/ol/:</i>			
<i>ak'</i>	'bow'	<i>ɔk'</i>	'to smoke'
<i>jak'</i>	'touch slightly'	<i>jɔk'</i>	'sweep'
<i>hɔra</i>	'away,up'	<i>hɔɔ</i>	'tortoise'
<i>ol:/ol/:</i>			
<i>oco</i>	causative suff.	<i>ɔcɔk'</i>	'remove'
<i>hon</i>	demonstrative	<i>hɔn</i>	'son'
<i>lo</i>	'draw water'	<i>lɔ</i>	'burn'
<i>al:/al/:</i>			
<i>acar</i>	'pickles'	<i>ɔcur</i>	'turn'
<i>bal</i>	'burn a hole'	<i>bəl</i>	'influence of Bonga'
<i>paera</i>	'swim'	<i>pəurə</i>	'distilled liquor'
<i>ol:/ul/:</i>			
<i>ot'</i>	'mushroom'	<i>ut'</i>	'swallow'
<i>kol</i>	'send'	<i>kul</i>	'tiger'
<i>ato</i>	'village'	<i>ətu</i>	'flow'
<i>el:/al/:</i>			
<i>eto</i>	'broken in'	<i>ətu</i>	'flow'
<i>ber</i>	'time'	<i>bəɽ</i>	'a certain tree'
<i>hela</i>	'cut down abundantly'	<i>hələ</i>	'bravo'
<i>ol:/al/:</i>			
<i>oco</i>	causative suff.	<i>əcu</i>	'set to do'
<i>hola</i>	'yesterday'	<i>hələ</i>	'bravo'
<i>cirə</i>	'a piece of land'	<i>ciro</i>	'the sun grass'
<i>al:/ā/:</i>			
<i>ak'</i>	'bow'	<i>āk'</i>	'bellow'
<i>jak'</i>	'touch slightly'	<i>jāk'</i>	'keep in order'
<i>aya</i>	'fem.attendant'	<i>āyā</i>	'true'

<i>/i:/ĩ/:</i>	<i>miru</i> ‘parrot’	<i>mīrū</i> ‘rimless’
<i>/ɛ:/ɛ̃/:</i>	<i>hēc</i> ‘come’	<i>hēc̃</i> ‘heavy’
<i>/ɔ:/ɔ̃/:</i>	<i>hɔ</i> INTERJ.	<i>hɔ̃</i> ‘also’
<i>/ə/:</i> <i>/ə̃/:</i>	<i>kər</i> ‘sore on the leg’	<i>kər̃</i> ‘thistle’
<i>/u /:</i> <i>/ū /:</i>	<i>ut</i> ‘swallow’	<i>ūt̃</i> ‘camel’
	<i>miru</i> ‘parrot’	<i>mīrū</i> ‘rimless’

Bodding (1922) gives several values of each of the vowels, that is, each vowel may be ‘narrow’ or ‘wide’ or ‘mid-mixed’, in his own words, depending on the environment in which it occurs. So far as our data goes, such distinctions are non-phonemic. Moreover, Bodding’s (1922) identification of six ‘modified’ vowels like *a, e, ɛ, i, o* and *u* could not be identified in our data except *ə/*, that is, Bodding’s *a* and in line with Pinnow (1959:35) they can be discarded as non-phonemic.

2.1.4 Length and nasalization

Vowel length is not phonemic. A vowel can be long or short depending on whether it occurs in an open (long) or closed syllable (short). Nasalization is phonemic. Nasalized forms of all the vowels are attested in our data in a limited number of examples. Because of the paucity of data contrastive pairs could not be given.

2.2 Diphthongs

Bodding (1922, 1929) lists three kinds of diphthongs – descending (after glide), ascending (before glide), and level (both vowels are equally strong). He arranged the following diphthongs into the above three categories. He also suggests most of the Santali diphthongs belong to the first category.

Descending:	<i>ae, ao, ai, au, eo, eo, ei, eo, iu, oe, oe, oi, and ui</i>
Ascending:	<i>ea, iq, oa, uq, and ui</i>
Level:	<i>ea, iq, io, iu, oa, uq</i>

For the ascending diphthongs he had the words ‘may be counted as’, and for the level ones he noted that ‘the last class is not, however, always pronounced as one diphthong; they are frequently, some of them generally, dissolved into two distinct syllables with a euphonic semivowel between them’ (1922:6). In our data, descending diphthongs are the most common ones, although with differences from those of Bodding, with a few ascending diphthongs. Level ones are found nowhere in our data.

Descending:	<i>ae, ao, eo, iu, ui, ai, əu, ei, əe, eo, oe, and oi</i>
Ascending:	<i>ea, oa, uə, iə, and io</i>

Table 2.3 details Santali vowel combinations.

TABLE 2.3: SANTALI VOWEL COMBINATIONS

V ¹ /V ²	i	e	ɛ	a	ə	ɔ	o	u
i					iə*		io*	iu
e	ei			ea*			eo	
ɛ							ɛo	
a		ae					ao	
ə	əi							əu
ɔ		ɔe						
o	oi	oe		oa*				
u	ui					uə*		

Note that wherever there is /i/ or /u/ in the combination /a/ becomes [ə]. It is also to be noted that in Bodding’s (1922, 1929) data, diphthongs like ‘ei’ were not attested to. He had ‘eɨ’, ‘eɔ’, and ‘oɨ’, all of which are not found in our data. We found ‘oɨ’ instead. ‘ui’ always behaves like a descending one as opposed to that in Bodding where it is both ascending and descending. The so-called level ones of Bodding are always found dissolved into two syllables. For most of the combinations, except /iə, io, ea, uə/, the coda is empty. Except these four all other combinations may be considered as a nucleus followed by glide. Examples:

<i>iu</i>	<i>jiu</i>	‘spirit’	<i>iə</i>	<i>niə</i>	‘this very’
				<i>tiək’</i>	‘lead’
	<i>n̄uri</i>	‘forest tree’		<i>ənɟiə</i>	‘male’
				<i>bəɽiə</i>	‘with difficulty’
				<i>iəte</i>	‘because of’
<i>ei</i>	<i>meila</i>	‘fair’	<i>io</i>	<i>tiok</i>	‘to reach’
	<i>meilaŋ</i>	INT. ‘come’	<i>ea</i>	<i>barea</i>	‘two’
				<i>dea</i>	‘back’
				<i>reaɽ</i>	‘cold’
				<i>sea</i>	‘to rot’
<i>eo</i>	<i>eoɽa</i>	‘to wind thread round the spindle’	<i>oa</i>	<i>noa</i>	‘this’
	<i>heoa</i>	‘to accustom’		<i>toa</i>	‘milk’
	<i>leoħa</i>	‘mix with a liquid’			
			<i>uə</i>	<i>ruə</i>	‘fever’
<i>eo</i>	<i>ħeo</i>	‘to carry on heap’		<i>ruəɽ</i>	‘to return’
	<i>n̄eota</i>	‘to invite’		<i>tuəɽ</i>	‘orphan’
	<i>leoħe</i>	‘stick to’			
<i>ae</i>	<i>sedae</i>	‘in the past’			
	<i>aema</i>	‘many’			
	<i>bəɽħae</i>	‘round about’			
	<i>mæjju</i>	‘woman’			
	<i>naeke</i>	‘priest’			
<i>ao</i>	<i>s̄ao</i>	‘in association’			
	<i>barħao</i>	‘to increase’			
	<i>batao</i>	‘to obey’			
	<i>benao</i>	‘to make’			

<i>əi</i>	<i>ləi</i>	‘to tell’
	<i>əikəu</i>	‘to feel’
	<i>kəi</i>	‘sin’
	<i>dəi</i>	‘elder sister’
<i>əu</i>	<i>əuri</i>	‘soon’
	<i>bəisəu</i>	‘to establish’
	<i>hiləu</i>	‘to shake’
	<i>biqəu</i>	‘to test’
<i>ɔe</i>	<i>hɔedak’</i>	‘thunder shower’
	<i>bɔeɦa</i>	‘brother’
	<i>kɔe</i>	‘to ask for’
	<i>ɔkɔe</i>	‘who’
<i>oi</i>	<i>moidoŋ</i>	‘bald’
	<i>noi</i>	‘listen, my girl’
	<i>oi</i>	INT. (‘yes’)
	<i>poi</i>	‘retribution’
	<i>qoi</i>	‘wooden spoon’
<i>oe</i>	<i>hoe</i>	‘to be’
	<i>hoeo</i>	‘to shave’
	<i>poesa</i>	‘pice’
	<i>podoe</i>	‘puff’
<i>ui</i>	<i>uihar</i>	‘fond remembrance’
	<i>rui</i>	‘cotton’
	<i>turui</i>	‘six’
	<i>muiguc’</i>	‘dirty’

2.3 Vowel harmony

So far as the harmonic sequence is concerned, there appear to be certain restrictions that need mention here:

- (i) In the same stress group, if *li* or *lu* occurs, *lə* but not *la* will occur; *niə* ‘this very’, *inə* ‘that very’, *ruə* ‘fever’, *busək* ‘give birth to’, *butəl* ‘power’, *bidə* ‘to dismiss’.
- (ii) *la* but not *lə* co-occurs with *le o e ɔ*; *aben* ‘you’ DL, *abon* ‘we’ INCL, *ale* ‘we’ EXCL, *alo* prohibitive PART, *nahel* ‘plough’, *tahen* ‘to stay’, *aten* ‘to listen’, *asen* ‘to lead’, *əŋjɔm* ‘to hear’, *sadɔm* ‘horse’, *adɔm* ‘many’, *asɔl* ‘chief’, *benao* ‘to make’, *boŋga* ‘evil spirit’, *pea* ‘three’, *dela* ‘invite to come’.
- (iii) If there is *le* or *lɔ* in the first syllable of a stress unit having more than one syllable there must always be *le* or *lɔ* in the following syllables; *eger* ‘to scold’, *ehɔp’* ‘to begin’, *etɔm* ‘right’, *enec* ‘to dance’, *gejer* ‘to break’, *gendrec* ‘a rag’, *gɔɾɔ* ‘to help’, *gɔɾen* ‘part.’, although diphthongs like *ɔe* and *ee* are found in a limited number of examples.
- (iv) Although sequence like *e-i* and *o-i* are found high vowels seldom co-occur with mid-high or mid-low vowels (exceptions being: *bedin* ‘pagan’, *begari* ‘forced labour’, *beperi* ‘trader’. These are all loanwords.).
- (v) *le* and *lo* never co-occur with *lu* in the same stress unit.

2.4 Consonant inventory

Aspirates in Table 2.4 are given in parentheses, as they seem to be borrowed, not native. As the language has borrowed extensively from the neighbouring Indo-Aryan languages, two sub-systems had developed; I presume one to be native and the other borrowed. In the native subsystem the contrast between voiced and voiceless plosives, except for retroflex, is neutralized in final position where they are replaced by corresponding checked sounds. Within the borrowed sub-system voiced and voiceless plosives contrast in all three positions. The borrowed sub-system has aspirated plosives, both voiced and voiceless. The native sub-system seldom uses aspirates except for a few words like *dhiri* ‘stone’, *ajhmar* (*əji* + *hanhar*) ‘wife’s or husband’s elder sister’. Ghosh (1994) did not mention the aspirated plosives, as the words’ aspirates are mostly, if not all, of Indo-Aryan origin. Taking together both the native and borrowed words, Neukom (2001:5), however, posits aspirates including an aspirated flap /ɽh/ in the phonemic system of Santali. Moreover he has, though hesitatingly, identified checked [p’ t’ c’ k’] as phonemes, which may not be so. The distribution of the consonants is shown in two charts, obviously with the idea of showing the differences between the two.

2.4.1 Consonants

Table 2.4 lists Santali consonants.

Note that the borrowed sub-system is different from the native sub-system in that it has aspirated plosives in all the five places of articulation. In addition, it has final contrast of voiceless and voiced unaspirated plosives, which is not there in the native subsystem. Comparative charts showing distribution of consonants as seen in Tables 2.5 and 2.6 will illustrate the point.

A closer scrutiny of the Tables reveals that

- (i) Whereas, within the native sub-system, the contrast between voiced and voiceless unaspirated plosives, except alveolar in the final position, is neutralized, it exists in the borrowed sub-system. There are, however, a few instances (e.g. *oḍok* ‘take out’, *əbuk* ‘wash’, *ṅut* ‘dark’, *ət* ‘earth’) where non-checked plosives occur finally, but those are just too few to lead to a conclusion. Neukom (2001:5) considers, perhaps in a reserved way, checked [p’ t’ c’ k’] as phonemes, but in our data they are actually allophones of /p t c k/.
- (ii) In the native sub-system [p’ t’ c’ k’] are specified when before voiceless plosives, and nasals, when they occur finally, optionally appear as voiced [b d j g] before vowels. [t’] is also realized as [ṭ’] before [k].

TABLE 2.4: SANTALI CONSONANTS

	Bilabial	Alveolar	Post-alveolar	Palatal	Velar	Glottal
Plosives:vl	<i>p</i>	<i>t</i>	<i>ʈ</i>	<i>c</i>	<i>k</i>	
VL.ASP.	(<i>ph</i>)	(<i>th</i>)	(<i>ʈh</i>)	(<i>ch</i>)	(<i>kh</i>)	
VD.	<i>b</i>	<i>d</i>	<i>ɖ</i>	<i>j</i>	<i>g</i>	
VD.ASP.	(<i>bh</i>)	(<i>dh</i>)	(<i>ɖh</i>)	(<i>jh</i>)	(<i>gh</i>)	
Nasal	<i>m</i>	<i>n</i>		<i>ɲ</i>	<i>ŋ</i>	
Fricative		<i>s</i>				<i>h</i>
Trill		<i>r</i>				
Flap			<i>ɽ</i>			
Lateral		<i>l</i>				
Glide	<i>w</i>			<i>y</i>		

TABLE 2.5: NATIVE SUB-SYSTEM

	-V	V-V	V- #
/p/ ¹⁹	[p]	[p]	[pʰ]/--#
/b/	[b]	[b]	—
/t/	[t]	[t]	[tʰ]/--# [tʰ]/--[k]
/d/	[d]	[d]	—
/tʃ/	[tʃ]	[tʃ]	[tʃ]
/dʃ/	[dʃ]	[dʃ]	[dʃ]
/k/	[k]	[k]	[kʰ]/--#
/g/	[g]	[g]	—
/c/	[c]	[c]	[cʰ]/--#
/j/	[j]	[j]	—
/m/	[m]	[m]	[m]
/n/	[n]	[n]	[n]
		[ɲ]/--hpA	
/ɲ/	[ɲ]	[ɲ]	[ɲ]
/ŋ/	—	[ŋ]	[ŋ]
/s/	[s]	[s]	—
/h/	[h]	[h]	[h] ²⁰
/r/	[r]	[r]	[r]
/ɽ/	—	[ɽ]	[ɽ]
/l/	[l]	[l]	[l]
/w/	—	[w]	—
/y/	—	[y]	—

TABLE 2.6: BORROWED SUB-SYSTEM

	-V	V-V	V- / #
/p/	[p]	[p]	[p]
/ph/	[ph]	[ph]	[ph]
/b/	[b]	[b]	[b]
/bh/	[bh]	[bh]	[bh]
/t/	[t]	[t]	[t]
/th/	[th]	[th]	[th]
/d/	[d]	[d]	[d]
/dh/	[dh]	[dh]	[dh]
/c/	[c]	[c]	[c]
/ch/	[ch]	[ch]	[ch]
/j/	[j]	[j]	[j]
/jh/	[jh]	[jh]	[jh]
/k/	[k]	[k]	[k]
/kh/	[kh]	[kh]	[kh]
/g/	[g]	[g]	[g]
/gh/	[gh]	[gh]	[gh]
/m/	[m]	[m]	[m]
/n/	[n]	[n]	[n]
		[ɲ]/--hpA	—
/s/	[s]	[s]	[s]
/h/	[h]	[h]	[h]
/r/	[r]	[r]	[r]
/ɽ/	—	[ɽ]	[ɽ]
/l/	[l]	[l]	[l]
/w/	—	[w]	—
/y/	—	[y]	—

Examples:

- (1) *sap'-ed-a-e* 'He is catching.' vs. *sab-ed-a-e*
- (2) *sab-a-e* 'He catches/will catch.' vs. *sap'-a-e*
- (3) *sap'-ket'-ko-a-e* 'He caught them.'

- (iii) Aspirate plosives, both voiced and voiceless, are part of the borrowed sub-system. Examples like *dhiri* 'stone' and *qhumbək'* 'crumple', as given by Neukom (2001:5), are either due to contraction of a vowel (*dihiri*>*dhiri*) or are borrowed from Indo-Aryan through a process of nativization (*k'* added to *qhumbə*).
- (iv) Velar nasal occurs medially (mostly with homorganic nasals) and only finally, never initially in the native sub-system, whereas in the borrowed sub-system it does not occur at all.
- (v) Palatal nasals do not occur in the borrowed sub-system.
- (vi) /tʃ/ in both sub-systems does not occur initially. Aspirated /tʃh/, postulated by Neukom (2001:5), without any minimal pair is not attested to in our data. Bodding (1932) also does not give any.
- (vii) Glides do occur in both, but only medially. [v] is not attested in our data. Bodding (1936) mentions this as a labiodental open voiced sound.

2.4.1.1 Consonant phoneme oppositions

(i) Voiceless vs. Voiced:

<i>/p/ : /b/ :</i>			
<i>pon</i>	‘four’	<i>-bon</i>	1PR PL INCL
<i>ape</i>	2PR PL	<i>aben</i>	2PR DL
<i>/t/ : /d/ :</i>			
<i>tala</i>	‘middle’	<i>dal</i>	‘thrash’
<i>ato</i>	‘village’	<i>adɔ</i>	‘then’
<i>/c/ : /j/ :</i>			
<i>celeɲ</i>	‘who’	<i>jeleɲ</i>	‘length’
<i>/k/ : /g/ :</i>			
<i>kalɔt</i>	‘pullet’	<i>galɔt</i>	‘agree upon’
<i>aka</i>	verbal suff.	<i>aga</i>	‘exaggerated’
<i>/tʃ/ : /dʒ/ :</i>			
<i>talao</i>	‘disobey’	<i>dʒalao</i>	‘make over’
<i>sufur</i>	‘crunch’	<i>sudʒur</i>	‘pour down’
<i>lat</i>	‘take possession of’	<i>ledʒ</i>	‘excrement of horse’

(ii) Unaspirate vs. Aspirate (In these pairs borrowed words are marked with ‘L’):

<i>/p/ : /ph/ :</i>			
<i>padao</i>	‘break wind’	<i>phada</i>	‘open’(L)
<i>paɾak</i>	‘split’	<i>pharak</i>	‘distance’(L)
<i>ape</i>	2 pr.pl.	<i>aphɔr</i>	‘sow seed’(L)
<i>/b/ : /bh/ :</i>			
<i>bar</i>	‘two’	<i>bhar</i>	‘load’
<i>lab</i>	‘profit’		
<i>/t/ : /th/ :</i>			
<i>tala</i>	‘middle’	<i>thapa</i>	‘slap’(L)
<i>tis</i>	‘when’	<i>thir</i>	‘still’(L)
<i>latar</i>	‘below’	<i>lathak</i>	‘humpy’(L)
<i>kitəb</i>	‘book’(L)	<i>kɔthne</i>	‘feign not to know’
<i>/d/ : /dh/ :</i>			
<i>dam</i>	‘value’(L)	<i>dhama</i>	‘basket’(L)
<i>diri</i>	‘delay’(L)	<i>dhiri</i>	‘stone’
<i>bada</i>	‘loan of seed’	<i>badha</i>	‘obstruction’(L)
		<i>godhɾa</i>	‘a felled sapling’(L)
<i>/c/ : /ch/ :</i>			
<i>cal</i>	‘go’(L)	<i>chal</i>	‘bark’(L)
<i>bickom</i>	‘rather’	<i>bichnəu</i>	‘disentangle’(L)
<i>/j/ : /jh/ :</i>			
<i>jal</i>	‘lick’	<i>jhəl</i>	‘long’(L)
<i>ajmao</i>	‘test’	<i>ajhmar</i>	‘elder sister’
<i>/tʃ/ : /tʃh/ :</i>			
<i>taka</i>	‘rupee’(L)	<i>tʃhək</i>	‘deceive’(L)
<i>kitli</i>	‘kettle’(L)	<i>kuʃhri</i>	‘room’(L)
<i>/dʒ/ : /dʒh/ :</i>			
<i>dajra</i>	‘bullock’	<i>dʒhaj</i>	‘stick’(L)
<i>canɟbɔl</i>	‘tail’	<i>jadʒhna</i>	‘phlegmatic’

<i>/k/:/kh/:</i>			
<i>kan</i>	verb copula	<i>khan</i>	‘if’(L)
<i>bakɕa</i>	‘perverse’	<i>bakhra</i>	‘share’(L)
<i>/g/:/gh/:</i>			
<i>gan</i>	‘about’	<i>ghanɕa</i>	‘bell’(L)
<i>bagra</i>	‘mixed’	<i>aghrao</i>	‘extend’(L)

(iii) Alveolar vs. Post-alveolar:

<i>/t/:/tʃ/:</i>			
<i>tala</i>	‘half’	<i>ɕalao</i>	‘transgress’
<i>at</i>	‘lose’	<i>ãt</i>	‘greatly’
<i>/d/:/dʒ/:</i>			
<i>dale</i>	‘be heaped’	<i>dala</i>	‘scales of balance’
<i>ad</i>	‘half’(L)	<i>aɕ</i>	‘cover’
<i>/r/:/rʃ/:</i>			
<i>ara</i>	‘saw’	<i>aɕa</i>	‘sort, kind’
<i>bar</i>	‘two’	<i>baɕ</i>	‘pool’

(iv) Trill vs. Lateral:

<i>/r/:/l/:</i>			
<i>rəi</i>	‘mustered’(L)	<i>ləi</i>	‘tell’
<i>arə</i>	‘nine’	<i>alə</i>	1PR.PL.EXCL.
<i>ber</i>	‘time’	<i>bel</i>	‘spread a mat’

(v) Nasals:

<i>/m/:/n/:</i>			
<i>ma</i>	optative particle	<i>na</i>	postposition
<i>eman</i>	‘do such’	<i>enaŋ</i>	‘a while ago’
<i>dəm</i>	‘delay’	<i>dən</i>	‘leap’
<i>/n/:/ɲ/:</i>			
<i>naram</i>	‘soft’	<i>ɲaɕam</i>	‘suck in’
<i>im</i>	‘lever’	<i>ɲ</i>	1PR SG.
<i>/n/:/ŋ/:</i>			
<i>enec</i>	‘until’	<i>eŋga</i>	‘mother’
<i>mun</i>	‘seer’	<i>muŋ</i>	‘kind of pulse’

(vi) Fricatives:

<i>/s/:/h/:</i>			
<i>sakam</i>	‘leaf’	<i>hako</i>	‘fish’
<i>busup</i>	‘straw’	<i>baha</i>	‘flower’
<i>as</i>	‘hope’(L)	<i>ah</i>	INTRJ of pleasure

(vii) Glides:

<i>/w/:/y/:</i>			
<i>dawa</i>	‘claim’(L)	<i>daya</i>	‘mercy’(L)

2.5 Suprasegmental phenomena

Santali has no phonemic tones or registers. Boddling (1929:6) also denies the existence of tone in the language. But level, rising, or falling intonation is significant in the

sentence. Stress is always on the second syllable of a word unless it is monosyllabic, irrespective of whether it is an open or a closed syllable.

2.6 Syllable structure

Santali words have the syllable structure of $(C^1)V(C^2)(C^1)$, where C^1 stands for all kinds of plosives, nasals (except $/ŋ/$), trills, laterals, and fricatives. C^2 stands for homorganic nasals. V is a vowel or nasal. In addition there are certain general constraints.

- (i) Words can begin with any vowel, diphthong, and consonant other than $/ɾ/$, $/ŋ/$, $/w/$, and $/y/$.
 - (a) The nucleus may be a vowel or nasal. Both $/ŋ/$ and $/n/$ are attested in this function. Forms like *Imahnderl* ‘day before yesterday’ [ma.hn.der], *Imahŋgal* ‘dear’ [ma.hŋ.ga], *Imahndal* ‘direction’ [ma.hŋ.ɖa] have a nasal at the nucleus of the second syllable.
 - (b) The opposition [\pm voice] is neutralized in the coda in the native sub-system.
 - (c) No native word has either $/s/$ or $/h/$ in the coda. Where these are found they are invariably Indo-Aryan loans, e.g. *as* ‘wish’, *bes* ‘good’, *bah* ‘excellent’.
 - (d) Non-native words with CC at the onset simplify CC either by vowel insertion or by dropping a C; for example, *prabhu* > *purbhu* ‘lord’, *skul* > *iskul* ‘school’. With CC at the coda the same process is employed; for example, *bench* > *benci*, *forest* > *phores*, *bottle* > *bɔtɔl*.
- (ii) An instance of CC at the coda is invariably a combination of C^2 and C^1 ; for example, *sendra* ‘hunt’, *mɔŋɖ* ‘tail’, *mɔŋj* ‘beautiful’.
- (iii) CCC in the middle of a word is realized as CC.C; for example, *sendra* [send.ra] and C at the onset of the next syllable is invariably alveolar trill.

Santali has the types of syllables illustrated below:

V	<i>ɛ.hɔp</i>	‘begin’			
VV	<i>ae.ma</i>	‘many’			
VC	<i>ac</i>	‘self’			
CVC	<i>sen</i>	‘go’	<i>dal</i>	‘beat’	<i>rɔɾ</i> ‘speak’
CV	<i>ə.gu</i>	‘bring’	<i>ma.hŋ.ga</i>	‘dear’	
CVV	<i>ba.ɖae</i>	‘know’			
CVCC	<i>candɖ.bɔl</i>	‘tail’	<i>gend.rec</i>	‘rag’	
VVC	<i>ead</i>	‘memorize’(L)			
CVVC	<i>tuər</i>	‘orphan’			

It is to be noted that syllables consisting of V, VV, or VVC, are rare in Santali. The language has a predilection for CV structure, which Bodding (1922,1929) has described as ‘open syllable’ and Neukom (2001) as ‘light syllable’, closely followed by CVC, Bodding’s ‘close’ and Neukom’s ‘heavy’ syllable.

In a frequency counting, out of 1,858 syllables (taken from Macphail (1983) distributed over 1,081 words), CV is 885 (i.e. 47.631%), CVC 641 (i.e. 34.499%), V 104 (i.e. 5.597%), VC 75 (i.e. 4.036%), VV 10 (i.e. 0.538%), and CVV 122 (i.e. 6.566%). Other types are negligible.

Santali also has predominantly dissyllabic word structure. In the same counting it was observed that 773 (i.e. 71.507%) are dissyllabic; 269 (i.e. 24.884%) monosyllabic, and

31 (i.e. 2.867%) are trisyllabic (most of which are borrowed). Due to the predominance of dissyllabic words, native words which were supposed to have been trisyllabic have been reduced to dissyllabic by dropping a vowel off of the second syllable; for example, *a + kirij* > *əkriɲ* ‘sell’, *haɾam* + <*p*> = *hapɾam* < *haparɾam* ‘ancestors’.

So far as the syllable structure of the stems is concerned, the following sequences are observed:

CV.CV	<i>ma.re</i>	‘ancient’
CV.CVV	<i>ma.nao</i>	‘honour’
CV.CVC	<i>da.pal</i>	‘fight’
CVC.CVC	<i>dal.paŋ</i>	‘half-naked’
CVCC.CV	<i>ɖaŋg.ra</i>	‘bullock’
V.CV	<i>ə.ɖi</i>	‘very’
V.CVV	<i>e.koi</i>	‘same’
V.CVC	<i>ɔ.pɔr</i>	‘pull mutually’
VC.CV	<i>ək.te</i>	‘time’
VC.CVV	<i>ul.təu</i>	‘reverse’
VC.CVC	<i>ɔn.tɔr</i>	‘mind’
VV.CV	<i>ae.ma</i>	‘many’
VV.CVV	<i>əi.kəu</i>	‘feel’

Of these sequences, CV.CVC is predominantly used. Tri- and more syllabic words are rare. In words with more than one syllable the second syllable gets the stress.

2.7 Morphophonology

- (i) Checked consonants of the verb stems optionally become voiced before finite *-a/* and marker for the imperative mood.

(4)	<i>dak'</i>	‘rain’	<i>dag-a.</i>	‘It rains/will rain.’
			also <i>dak'-a</i>	
(5)	<i>sap'</i>	‘catch’	<i>sab-a-m</i>	‘You catch.’
			also <i>sap'-a-m</i>	
(6)	<i>rəput'</i>	‘break’	<i>rəpud-a-e</i>	‘He breaks/will break.’
			also <i>rəput'-a-e</i>	
(7)	<i>perec'</i>	‘fill’	<i>perej-a-e</i>	‘He fills/will fill.’
			also <i>perec'-a-e</i>	

- (ii) Checked consonants of the TAM suffixes in the active (*t'*) are obligatorily voiced before finite *-a/* and vowel-initial pronominal suffixes (*ɪn* ‘I’ and *-e* 3PR SG.).

- (8) *jəm-ked-a-e*
eat-PST:A-FIN-3SG:SUBJ
‘He ate.’
- (9) *dal-ked-e-a-ko*
beat-PST:A-3SG:OBJ-3PL:SUBJ
‘They beat him.’
- (10) *dal-keɽ'-ko-a-e*
beat-PST:a-3PL:OBJ-FIN-3SG:SUBJ
‘He beat them.’

- (11) *sap'-led-e-a-kin*
 catch-PLUP-3SG:OBJ-FIN-3DL:SUBJ
 'They two caught him.'

This way *et'* > *ed* in present imperfect, *ket'* > *ked* in past, *let'* > *led* in pluperfect.

(iii) Checked consonants in verb stems become voiced when followed by middle voice suffix */-ok'/*. Thus,

- (12) *hij-uk'-a-e* (<*hec'* 'come')
 come-M-FIN-3SG:SUBJ
 'He comes/will come.'

- (13) *dej-ok'-mε* (<*dec'* 'climb')
 climb-M-2SG:IMP
 'Climb.'

(iv) *le/* of TAM suffixes */-et'/*, */-ket'/* and */let'/* is raised when followed by 1PR SG. suffix. Thus,

- (14) *tɔl-kid-ɨn-a-e*
 tie-PST:A-1SG:OBJ-FIN-3SG:SUBJ
 'He tied me.'

(v) Verb stems ending in a vowel (especially */o/*) get an increment *[-n]* in the imperative.

- (15) *teygon-mε* (<*teygo* 'stand')
 stand-2SG:IMP
 'Rise up.'

- (16) *topon-mε* (<*topo* 'bathe')
 bathe-2SG:IMP
 'Take bath.'

3 MORPHOLOGY

3.1 Nominal morphology

The morphemes that take case affixes and postpositions and that show agreement in the verb are classed as nominals with subcategories of nouns and pronouns. Nouns and pronouns are grouped under the same class, as they have the same functional load. The major difference between the two is that whereas on the paradigmatic axis nouns show opposition in number, (\pm gender), and case, pronouns show opposition in person, number, and case.

3.1.1 Number

Santali has three grammatical numbers, both on nominals and predicates – singular (unmarked), dual (marked by *-kin*), and plural (marked by *-ko*). Number marking is obligatory in the case of animate nouns. Pronouns have separate forms for all the three numbers in the first and second person. The third person pronominal form takes the dual and plural suffixes to form their dual and plural forms. In the third person animate there is a marker */-i/*; in the interrogative and the indefinite, */-e/* in the singular, but those are more gender markers (animate and inanimate).

<i>seta-ø</i> 'dog'	<i>seta-kin</i> 'two dogs'	<i>seta-ko</i> 'dogs'
dog-SG.	dog-DL.	dog-PL.

<i>nu-i-ø</i> ‘this person’ this-ANIM.-SG.	<i>nu-kin</i> ‘these two persons’ this-ANIM.-DL.(ANIM.)	<i>no-ko</i> ‘these persons’ this-ANIM.-PL (ANIM.)
<i>no-a-ø</i> ‘this thing’ this-INAN.-SG.	<i>no-a-kin</i> ‘these two thing’ this-INAN.-DL.	<i>no-a-ko</i> ‘these things’ this-INAN.-PL.
<i>ɔkɔ-e-ø</i> ‘who’ INTER.-ANIM.-SG.	<i>ɔkɔ-e-kin</i> ‘who two’ INTER.-ANIM.-DL.	<i>ɔkɔ-e-ko</i> ‘who all’ INTER.-ANIM.-PL.
<i>ok-a-ø</i> ‘which’ INTER.-INAN.-SG.	<i>ok-a-kin</i> ‘which two’ INTER.-INAN.-DL.	<i>ok-a-ko</i> ‘which all’ INTER.-INAN.-PL.
<i>jāhã-e-ø</i> ‘anyone’ INDEF.-ANIM.-SG.	<i>jāhã-e-kin</i> ‘any two persons’ INDEF.-ANIM.-DL.	<i>jāhã-e-ko</i> ‘anybody’ INDEF.-ANIM.-PL.
<i>jāhã-ø</i> ‘anything’ INDEF.(INAN.)-SG.	<i>jāhã-kin</i> ‘any two things’ INDEF.(INAN.)-DL.	<i>jāhã-ko</i> ‘any all’ INDEF.(INAN.)-PL.

Note that, in case of demonstratives the gender suffix for the animate in the singular merges with the number suffixes in dual and plural, but the inanimate marker remains intact. This may indicate that */-i/* is the animate gender marker which is dropped when dual and plural suffixes are present. This may also indicate that */-kin/* and */-ko/* are originally dual and plural markers for the animate, and their application to inanimate nominals is rather secondary. In case of interrogatives and indefinites, however, the animate gender suffix is consistently present in singular, dual, and plural.

The dual forms of the first and second personal pronouns are also used to denote singularity when these are used among certain kin-relations. Parents-in-law and children-in-law use second person dual in addressing each other, and use the exclusive form of the first personal pronoun in talking to each other and referring to themselves, when just one person is meant. This forms a sub-system of honorific usage.

- (17) *ceka-en-a-ben* *bəhu*
how-PST:M-FIN-2DL:SUBJ daughter
‘How are you, daughter-in-law?’
- (18) *ruə-k’-kan-a-ljɪn*
fever-M-COP-FIN-1DL:EX
‘I am getting fever.’

There is no honorific pronoun in Santali. But nowadays there is a tendency among the educated Santals to use the dual form of the second personal pronoun to address and talk to a respected, senior, or unfamiliar person(s).

The plural also functions as an expression of singularity among certain kin-relations. Co-parents-in-law in addressing each other and talking among themselves use the inclusive form of the first person when just one person is meant.

- (19) *henda ho sumdhi, cet’leka menak’-bon-a* ‘O co-parent-in-law, how are you ?
(20) *ədi muskil-re-bon paɾao-aka-n-a* ‘I am in great trouble’

Sometimes */-ko/* with the addition of */ta-/*, hence */-takol/*, functions as the plural suffix.

- tuɖu* ‘a Santal sept’ > *tuɖu-tako* ‘men of the Tudu sept’
hɔpɔnera-ŋ ‘my daughter’ > *hɔpɔnera-ŋ-tako* ‘my daughters’

In the absence of more examples it is difficult to say anything definite concerning this, that is, whether it is plural or has something to do with belongingness.

3.1.2 Case

Case markers attach directly to the bare nominal. Santali cases can be divided into following categories:

- Core (subject and object): unmarked.
- Peripheral (comitative, genitive, instrumental, sociative, allative, ablative, and locative): marked on the nominal.

Core cases are unmarked on the nominal but are marked in the verb in the form of incorporated pronouns (if subject or object is animate) as transitive subject, intransitive subject, and transitive object. In the case of verbs taking two objects the so-called indirect object (dative?) is marked in the verb with the applicative prefix /a-/. If there are two objects and both are animate, only the indirect is marked in the verb, that is, the indirect object is raised to object with the applicative /a-/. Genitive is also marked in the verb when it denotes inalienable possession, in which case the prefix /t-/ is attached to the applicative forms of the pronouns; otherwise it is marked in the noun phrase and functions as an attribute. Consider the following examples:

- (21) *gidrə rak'-ed-a-e* (/e/ is marked for the transitive SUBJ. *gidrə*)
 child cry-IMPREF-FIN-3SG:SUBJ
 'The child is crying.'
- (22) *phulmoni sən-ək'-a-e* (/e/ marked for intransitive SUBJ.)
 Phulmoni go-m-fin-3sg:subj
 'Phulmoni will go/goes.'
- (23) *dal-ked-e-a-e* (/e-/ marked for transitive OBJ.)
 beat-PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'He beat him.'
- (24) *ləi-a-ko-a-e* (/ako-/ marked for transitive OBJ. with applicative /a/)
 tell-APPL-3PL:OBJ-FIN-3SG:SUBJ
 'He will tell them.'
- (25) *uni də ac'-ak' janga rəput'-akat'-t-ae-a*
 3SG TOP 3S-GEN leg break-PRF:A-POSS-3SGPOSS-FIN
 'He has broken his leg.' (/tae-/ marked for possessor/undergoer of the action)

Table 2.7 lists case markers.

Genitive. There are two sets of suffixes for genitive – one for animate the other for inanimate. The suffix for the animate /ren/ is used when the governed noun is animate and /ak'/ is used when the governed noun is inanimate. For the inanimate there are other suffixes too, such as /reak'/, /reaŋ/, /renak'/, and /renaŋ/. The suffixes marked as 'other' are normally attached to inanimate nouns and demonstratives. In one example like *hapram-ko-reak' katha* 'story of the ancestors',

TABLE 2.7: CASE MARKERS

Case	Marker	Example	Syntactic function
<i>Core</i>			
Nominative	∅	<i>bəeha</i>	Transitive SUBJ Intransitive SUBJ Transitive OBJ
<i>Peripheral</i>			
Genitive	<i>-ren</i> (anim) <i>-ak',reak'</i> (inan)	<i>ɪn-ren bəeha</i> 'my boy' <i>ɪn-ak' ti</i> 'my hand' <i>dare-reak' jə</i> 'fruit of tree'	possessor
Comitative	<i>-thenl-then'</i>	<i>bəeha-then</i>	goal, place
Instrumental-	<i>-te</i>	<i>ɖay-te</i>	instrument,
Locative		<i>iskul-te</i>	cause, motion
Sociative	<i>sāo*</i>	<i>bəeha-sāo</i>	association
Allative	<i>-senl-sec'</i>	<i>oʔak'-sec'</i>	direction
Ablative	<i>khənlkhəc'</i>	<i>dare-khən</i>	source, origin
Locative	<i>-re</i>	<i>dak'-re</i>	spatio-temporal location

Note

Peripheral cases are marked on the nominals.

/-reak'/ is found to be attached with animate noun. Examples showing occurrences of genitive suffixes:

<i>ɪn-ren merəm</i>	'my goat'	<i>am-ren həpən</i>	'thy son'
<i>ac'-ak' jaŋga</i>	'his leg'	<i>abo-ak' oʔak'</i>	'our house'
<i>oʔak'-reak' duər</i>	'door of the house'	<i>dare-reak' jə</i>	'fruit of the tree'

Here genitive suffixes are used to express possessive relationship or belongingness to somebody or something.

Sometimes genitive suffixes also mark the topic of discussion:

<i>noa-reak'</i>	<i>mit'jaŋ</i>	<i>kəhəni</i>	<i>ləi-ad-ɪn-a-e</i>
this-GEN	one-CLSSFR	story	tell-APPL:PST:A-1SG:OBJ-FIN-3SG:SUBJ
'he told me a story about this' – denoting something about or concerning.			

So far as formation of */-renl* and */-reak'/* is concerned, Neukom (2001:29) tries to derive these from locative suffix */-re/* with the 'additional element *-n*' (in case of */-ren l*) and inanimate */ak'/*. Ghosh (1994) tries to derive */-reak'/* in the same way, but he could not derive */-renl* from the same locative, as the meaning of the final *-n* (in */-renl*) was not clear. Neukom, too, has the same problem as he submits 'which has no clearly definable meaning' (2001:29). L. Burrows (1915:17) called this type of genitive a locative-genitive, as this type of genitive often denotes a belongingness of something to a place.

Comitative. The suffix for this case is */-thenl* (with the variant *-then'*). Bhattacharya (1975:148–149) assigned this suffix a dative role in saying that 'when the verb has two objects, if both the objects are animate, the indirect object which precedes the direct object takes a dative suffix (or postposition) whereas the direct object is in

the accusative case, i.e. it is represented by an object particle in the verb...'. His examples are:

maeju-then gidrə-n em-ked-e-a
 woman-DAT child-1SG:SUBJ give-PST:A-3SG:OBJ-FIN
 'I gave the child to the woman.'

gidrə-then maeju-n em-ked-e-a
 child-DAT woman-1SG:SUBJ give-PST:A-3SG:OBJ-FIN
 'I gave the woman to the child.'

In my data */-then/* is not found in the dative role and if there are two animate objects of the verb, the indirect object with the applicative marker is always shown in the verb rather than the direct object. Neukom (2001: 29) assigns the suffix dative status. But the examples he has cited (except examples 10 and 11) appear more comitative than dative. Data containing */-then/* in my field notes are more comitative than anything. Consider the following:

(26) *in-then mena-k'-a*
 1SG-with exist-M-FIN
 'It is with me.'

(27) *ona alo-then-e seter-en-khan-ge uni bijn dɔ*
 that light-near-3SG:SUBJ reach-PST:M-if-FOC 3SG snake TOP
hətər-te-e gɔc'-en-a
 weapon-INS-3SG:SUBJ kill-PST:M-FIN
 'Having been reached near the light that snake was killed by the weapon.'

(28) *uni budhi-then cɔɔ-bon*
 that old woman-near go-1PL.INC
 'Let us go to that old woman.'

(29) *in am-then noa katha ləi-ləgit'-in həc'-len-a*
 1SG:SUBJ 2-near this word tell-for-1SG:SUBJ come-PLUP:M-F
 'I had come to tell you this.'

(30) *uni ato mənjhi-then sɔpɔhət'-e nam-a*
 3SG village headman-with help-3SG:SUBJ get-FIN
 'He will get help with the village headman.'

Instrumental-Locative. The suffix */-te/* is used to indicate two functions – instrumental and locative. With the locative function it indicates the place (as well as the time), the destination (reached) as against the allative */-sen/* which indicates 'movement towards, direction', etc.

(31) *oka-te-m cal-ak'-kan-a*
 where-LOC-2SG:SUBJ go-M-COP-FIN
 'Where are you going?'

(32) *iskul-te cal-ak'-mɛ*
 school-LOC go-M-2SG:IMP
 'Go to the school.'

(33) *nəi-te-ŋ cal-ak'-a*
 river-LOC-1SG:SUBJ go-M-FIN
 'I go/shall go to the river.'

(34) *tin soməe-te-m cal-ak'-a*
 when time-LOC-2SG:SUBJ go-M-FIN
 'At what time will you go?'

Functions as instrumental are:

(i) Instrument:

(35) *ɖaŋ-te tiok'-mə*
 stick-by pull down-2SG:IMP
 'Pull down by the stick.'

(36) *uni tuŋi-te cẽŋe tuŋi-led-e-a-e*
 3SG arrow-INS bird pierce-PLUP:A-3SG:OBJ-FIN-3SG:SUBJ
 'He had pierced the bird with the arrow.'

(ii) Cause:

(37) *uni bɔtɔr-te dɔŋ-akad-a-e*
 3SG fear-INS run away-PRF:A-3SG:SUBJ
 'He has run away out of fear.'

(iii) Manner:

(38) *khusi-te-bon rəskə- k'-kan-a*
 pleasure-INS-1PL.INC enjoy-M-COP-FIN
 'We are enjoying with pleasure.'

(iv) Through, by means of:

(39) *tinək' gɔnɔŋ-te-m hatao-a*
 how price-INS-2SG:SUBJ take-FIN
 'At what price will you take?'

(40) *noa kapaŋ-te ca-lak'-mə*
 this door-INS go-M-2SG:IMP
 'Go through this door.'

(41) *noa duər-te cal-ak'-mə*
 this corridor-INS go-M-2SG:IMP
 'Go through this door.'

Sometimes /-te/ is also found with the subject in passive construction:

(42) *ɨŋ-te-ŋ mak'-akan-a*
 1SG-INS-1SG:SUBJ cut-PRF-M-FIN
 'It is cut by me.'

Sociative. Sociative is marked by a postposition *sāo* 'with, in association with'.

(43) *alɛ-sāo hij-uk'-mə*
 1PL.EX-SOC come-M-2SG:IMP
 'Come with us.'

- (44) *pətəb-sāo selet hɔ̃-e əgu-i-a*
 book-SOC slate too-3SG:SUBJ bring-y-FIN
 ‘He will bring slate with books too.’

sāo is also found to occur with suffix */-te/*, thus *sāote*:

- (45) *həpən-sāo-te bana eŋga hɔ̃-e gɔc'-led-e-a*
 cub-SOC-INS bear-FEM. too-3SG:SUBJ kill-PLUP:A-3SG:OBJ-FIN
 ‘He killed a she-bear along with her cub.’

Allative. The allative suffix */-sen/* (variant */-sec'/*) expresses direction towards destination.

- (46) *gɔta-sen-ge dhar hena-k'-a*
 all-ALL-FOC sharpness exist:M-FIN
 ‘There is sharpness in all directions.’

bir-sen ‘towards the forest’ *buru-sen* ‘towards the mountain’

Ablative. The postposition for the ablative is *khən* (with variant *khəc'*) and is used to express ‘from, away from’:

- (47) *dare-khən ārgon-me*
 tree-ABL come down-2SG.:IMP
 ‘Come down from the tree.’

- (48) *sərim-khən jɛl-ək'-kan-a*
 roof-ABL see-m.-COP-FIN
 ‘It is being seen from the roof.’

khən is also found to occur with inanimate suffix */-ak'/*:

- (49) *raja oɾak'-khən-ak'-e dəŋ-ked-a*
 king house-ABL-INAN-3SG:SUBJ run away-PST:A-FIN
 ‘The king fled away from the house.’

Locative. The suffix for the locative is */-re/*, indicating location in spatiotemporal axes. It specifies destination or place reached. So, if */-te/* indicates locative of motion, */-re/* indicates locative of rest.

khaŋaŋ-re ‘in the ditch’ *siŋgəŋ ɔktɔ-re* ‘at night fall’
ato-re ‘in the village’ *juətət'-re* ‘in the dark’ *bhitri-re* ‘inside’

3.1.3 Person

Santali distinguishes between alienable and inalienable possession. In the case of alienable possession, the possessor appears in the genitive before the head noun. In case of inalienable possession, instead of a genitive attribute the head noun itself is marked for the possessive relationship by any one of the suffixes for the three persons, namely *-ŋ* for first person, *-m* for second person, and *-t* for third person, irrespective of singular, dual or plural.

The system applies to kinship relational terms only. As in:

<i>bəkɔ-t</i>	‘his/their younger brother’	<i>bəkɔ-ŋ</i>	‘my/our younger brother’
<i>apa-t</i>	‘his/their father’	<i>eŋga-t</i>	‘his/their mother’
<i>eŋga-m</i>	‘thy/your mother’	<i>kimin-me</i>	‘thy/your wife’
<i>həpən-iŋ</i>	‘my/our son’		

The full form of the first person is *ij̃* instead of *-j̃*, and the second person imperative form *m̃e* for *-m̃* is used when the term for the kinship ends in a consonant.

3.1.4 Definiteness

Santali has one marker *-t̃eɬ'* for nouns to mark definiteness and another *-t̃ak'* for pronouns. To mark unspecified objects Santali has none; the numeral for 'one' may be used to mark nondefinite referential entity. The definite marker is used to emphasize, especially identity.

- (50) *gidr̃ə-t̃eɬ'-ko-e* *em-at'-ko-a,* *ale d̃ə bay*
 child-DEF-3PL:OBJ-3SG:SUBJ give-APPL:PST:A-3PL:OBJ-F 1PL.EX TOP NEG
 'He gave only to the children, not to us.'
- (51) *q̃əɾ-t̃eɬ'* *əgu-i-m̃e,* *dare-t̃eɬ'* *ik̃ə-k̃ə-k'-m̃e*
 branch-DEF bring-y-2SG:IMP tree-DEF be-MOD-M-2SG:IMP
 'Bring the branch, let the tree be.'

3.1.5 Nominal class/gender

Santali has two types of gender distinction, one grammatical and the other lexical. The grammatical gender distinguishes, on the one hand, between [\pm animate], which is a native system, and between [\pm male] with the inflections */-a/* and */-i/*, respectively, which is borrowed from Indo-Aryan. The lexical gender is, in Bhattacharya's (1976:195) term, 'compounded sex-based gender', that is, sex-linked words are attached to the sexually indeterminate nouns to mark male or female.

Santali distinguishes between two classes of nouns – animate and inanimate.

From verb-roots and stems, derived demonstratives as well as lexical attributes, indefinite pronouns, and derived adverbials, animate nouns are derived with suffix */-ic'/* and inanimate nouns with */-ak'/*.

<i>r̃əh̃ɔy-ic'</i>	'sower'	<i>r̃əh̃ɔy-ak'</i>	'that which is sown' < <i>r̃əh̃ɔy</i> 'to sow'
<i>dadal-ic'</i>	'man who beats'	<i>dadal-ak'</i>	'that which is beaten' < <i>dal</i> 'to beat'
<i>nui-ic'</i>	'his one'	<i>nui-ak'</i>	'his thing' < <i>nui</i> 'this one'
<i>ñɔt̃e-n-ic'</i>	'one of this side'	<i>ñɔt̃e-n-ak'</i>	'thing of this side' (< <i>ñɔt̃en</i> 'belonging to this direction', 'in this direction')
<i>pond̃-ic'</i>	'white one'	<i>pond̃-ak'</i>	'white thing' < <i>pond̃</i> 'white'
<i>arak'-ic'</i>	'red one'	<i>arak'-ak'</i>	'red thing' < <i>arak'</i> 'red'
<i>eɬak'-ic'</i>	'any one'	<i>eɬak'-ak'</i>	'anything' < <i>eɬak'</i> 'any'
<i>j̃āh̃ān-ic'</i>	'somebody'	<i>j̃āh̃ān-ak'</i>	'something' < <i>j̃āh̃ā</i> 'any, some'
<i>noŋka-n-ic'</i>	'man like this'	<i>noŋka-n-ak'</i>	'thing like this' < <i>noŋka</i> 'in this manner'

The verb root with TAM suffixes can also take these suffixes, and then the entire verb construction becomes a noun, for example:

dal-ked-e-y-ic' 'one who struck him'

Besides living beings the following objects are also considered animate. Celestial bodies, that is, the sun, moon, and stars are considered animate. Consider the

following examples where the above are marked in the verb in the form of third personal pronominal marker:

- (52) *siŋcādo rakap'-kan-a-e*
 sun rise -COP-FIN-3SG:SUBJ
 'The sun is rising.'
- (53) *niŋdɔcādo dubuc'-en-a-e*
 moon set-PST:M-FIN-3SG:SUBJ
 'The moon set.'
- (54) *hola-ren ipil-ij ŋel-ket'-ko-a*
 yesterday-GEN star-1SG:SUBJ see-PST:A-3PL:OBJ-FIN
 'I saw the stars of yesterday.'

Spiritual beings are regarded as animate. Consider the following examples:

- (55) *uni mit'-taŋ kəli-boŋga benao-akad-e-a-e*
 he one-CLSSFR Kali-idol make-PRF:A-3SG:OBJ-FIN-3SG:SUBJ
 'He has made a Kali idol.'

The words 'puff-ball', 'ear-wax', 'complete shells of snails', 'thorn when pricked' are regarded as animate.

- (56) *puŋkə-ko halay-ket'-ko-a*
 mushroom-3PL:SUBJ collect-PST:A-3PL:OBJ-FIN
 'They collected mushrooms.'
- (57) *ērgɔt' mena-k'-ko-a*
 ear-wax have-M-3PL:OBJ-FIN
 'There are earwaxes.'
- (58) *ij ij-ren jhinuk-ij rapak'-ket'-ko-a*
 I 1s-GEN shells-1SG:SUBJ burn-PST:A-3PL:OBJ-FIN
 'I burnt my shells.'
- (59) *jənum-ij tot'-ked-e-a*
 thorn-1SG:SUBJ extract-PST:A-3SG:OBJ-FIN
 'I extracted the thorn.'

Dead human beings and animals are considered animate if the name is used or the person or animal referred to.

- (60) *gɔc'hɔŋ-ko əgu-ked-e-a*
 deadman-3PL:SUBJ bring-PST:A-3SG:OBJ-FIN
 'They brought the dead man.'
- (61) *gɔc' ənɟiə-ko topa-ked-e-a*
 dead bull-3PL:SUBJ bury-PST:A-3SG:OBJ-FIN
 'They buried the dead bull.'

The animate-inanimate distinction is also marked in the choice of genitive suffixes *-l-ren/* for animate and *-ak'/* with variants for inanimate.

Inflected gender. Sometimes the distinction between \pm male is expressed morphologically. Certain nouns ending in *-a* form their feminine by replacing the final *-a* by *-i*. This is obviously an Indo-Aryan trait borrowed by Santali.

koṛa ‘boy’: *kuṛi* ‘girl’; *bhola* ‘dog’: *bhuli* ‘bitch’
bhedā ‘ram’: *bhid̥i* ‘sheep’ *mama* ‘maternal uncle’: *māmi* ‘maternal aunt’
kala ‘deaf’: *kāli* ‘deaf.FEM.’ *koṅka* ‘foolish’: *kuṅki* ‘foolish.FEM.’
koḍa ‘dumb’: *kuḍi* ‘dumb.FEM.’ *caḍra* ‘bald headed’: *cāḍri* ‘bald.FEM.’

Sex-based gender. This may be purely lexical where separate words for male and female are used to indicate different categories.

āṇḍiā ‘ox’: *gāi* ‘cow’ *bōeha* ‘brother’: *misera* ‘sister’
jāwāy ‘husband’: *bāhu* ‘wife’ *herel* ‘husband’: *maejiu* ‘wife’
apat ‘father’: *eṅgat* ‘mother’ *kaḍa* ‘he-buffalo’: *bitkil* ‘she-buffalo’

Compounded sex-based gender. In this type the distinction between ±male is made clear by combining sex-based words like *āṇḍiā*, *saṇḍi*, *pēṭhar*, and *kuḍu* for male and *eṅga*, *bāchi*, and *pāṭhi* for female with sexually indeterminate words.

āṇḍiā pusi ‘male cat’: *eṅga pusi* ‘cat’ FEM *saṇḍi sim* ‘cock’: *eṅga sim* ‘hen’
pēṭhar mihū ‘male calf’: *bāchi mihū* ‘do’ FEM *kuḍu sukri* ‘boar’: *pāṭhi sukri* ‘pig’

3.1.6 Pronouns (personal, interrogative, etc.)

The free-forms for the personal pronouns are given in Table 2.8. These pronouns may not be overtly marked in the utterance if the identity of the referent is not focused.

Note that there is an inclusive–exclusive difference in first person dual and plural. These forms are used in different social settings among certain kin-relations. They are not used indiscriminately. A more formal description would be that inclusive is [+speaker +hearer] where as exclusive is [+speaker –hearer], that is, the inclusive is marked by the presence of the addressee in the discourse while in the case of the exclusive it is not. There is also a distinction in the third person, that of anaphoric vs. unmarked.

Personal pronouns are not overtly marked if they are not focused.

- (62) *khan ona aṅjəm-kate goṭa bəd bōehar-kin si-caba-ked-a.*
 then that hear-CV whole high low land-2DL:SUBJ plough-finish-PST-FIN
 ‘Then having heard that they two ploughed up the whole land.’
- (63) *am - gē lai-me tōbe cikə-kate-ṅ gəj-e-a*
 thou- FOC tell-2SG:IMP then how-CV-1SG:SUBJ kill-3SG:OBJ-FIN
 ‘Then thou tell how shall I kill him.’

TABLE 2.8: PERSONAL PRONOUNS

Person	Singular	Dual		Plural	
		INCL	EXCL	INCL	EXCL
First	<i>ij̄n</i>	<i>alaj̄</i>	<i>əl̄ij̄n</i>	<i>abo</i>	<i>alē</i>
Second	<i>am</i>		<i>aben</i>		<i>ape</i>
Third	<i>ac̄</i>		<i>əkin</i>		<i>ako</i>
	<i>uni</i>		<i>unkin</i>		<i>onko</i>

In (62) the full form of the subject pronoun is not used as it is not focused, it is only marked in the predicate. In (63) the full form of 2SG. PR subject is used with the focus marker *ge* as it is focused.

Use of the dual or plural when one person is addressed. The first person exclusive dual *əlɪŋ* is used by a single speaker when s/he talks about or refers to him/herself to his/her father/mother-in-law, and first person plural inclusive is used by a single speaker when s/he talks about or refers to him/herself to his/her co-parent-in-law. The second person dual is used by parent-in-law in addressing son/daughter-in-law when just one person is meant.

- (64) *henda ho sumdhi cet'-leka mena-k'-bon-a*
 o co-parent-in-law how-like be-M-1PL.INC-FIN
 'O co-parent-in-law how are you (we)?'
- (65) *əɟi muskil-rɛ-bon paɾao-akan-a*
 very trouble-LOC-1PL:SUBJ fall-PRF:M-FIN
 'I am in great trouble.'
- (66) *ceka-en-a-ben bəhu*
 how-PST:M-FIN-2DL:SUBJ daughter-in-law
 'How are you, daughter-in-law?'
- (67) *ruə-k'-kan-a-lɪŋ*
 fever-M-COP-FIN.-1DL:SUBJ
 'I am getting fever.'

Anaphoric vs. Unmarked. There are two sets of forms for the third person, one based on *ac'* and the other on demonstrative *uni* which is unmarked. The forms based on *ac'* are anaphoric while those based on the demonstrative *uni*, 'that person', are both anaphoric and deictic. *ac'* (with *əkin* and *ako* in dual and plural) is used for human referent only and refers back to the topic; otherwise *uni* (*unkin* and *onko* in dual and plural) is used.

- (68) *tuər kuɾi gidrə. ac' baba bən-uk'-e-a*
 orphan girl child. she father NEG-M-3SG:OBJ-FIN
 'The orphan girl. She has no father.'

It can also be used in the same sentence when it is co-referential of the subject.

- (69) *ɪŋ əgu gəi dɔ ac'-ak' goɾa-rɛ-y-e ader-akad-e-a*
 1SG bring cow TOP he-GEN cowshed-LOC-y-3SG:SUBJ bring-PRF:A-3SG:OBJ-FIN
 'He has brought the cow which I brought to his cowshed.'

Use of *uni* is anaphoric and deictic both. It should also be mentioned that frequency-wise use of *uni* is more common.

- (70) *uni təkħɔn əndɛ-ge ərsi-te ɲel-ək'-kan-tahēkan-a-e*
 She then there-FOC mirror-LOC see-M-COP-PST-FIN-3SG:SUBJ
 'Thereupon she was looking at the mirror.'

Forms of the interrogative pronouns are given in Table 2.9.

There are two types of interrogatives with \pm animate distinction—referential and non-referential. The referential is used when the identity of the person or object is

TABLE 2.9: INTERROGATIVE PRONOUNS

	Animate	Inanimate
Referential	<i>ɔkɔe</i>	<i>oka</i>
Non-referential	<i>cele</i>	<i>cet'</i>

known to the addressor, and the non-referential when it is uncertain. Both the types form their dual and plural with *-kin* and *-ko*.

- (71) *ɔkɔe-kan-a-m*
 which person COP-FIN.-2SG:SUBJ
 'Who are you?'
- (72) *oka-ʔak' dɔ am-ak'*
 which-DEF TOP thou-GEN
 'Which (thing) is yours?'
- (73) *cele kan-a-e*
 who COP-FIN-3SG:SUBJ
 'Who is there?'
- (74) *cet'-t-am jutum*
 what-POSS-2SG:SUBJ name
 'What is thy name?'

Note that *ɔkɔe* can be used for human only, that is, it is + human, for non-human animate to be specified the definitive suffix *-ʔak'* is added.

- (75) *oka-ʔak' dɔ am-ak'*
 which-DEF TOP 2S-GEN
 'Which one is yours?'

Other forms of Interrogative are

tis, 'when'; *tinək'* 'how much'; *cedak'* 'why'; *ceka / cikə* 'how'
tis is also found to occur with locative suffix *-re*.

The most common indefinite root is *jāhā* 'any' which itself is used for the inanimate and it takes person marker *-e* for the animate. Both take *-kin* and *-ko* for dual and plural. Forms of the indefinite pronouns are given in Table 2.10.

- (76) *jāhāe-ge met-a-e-me*
 anybody-FOC tell-APPL-3SG:OBJ-2SG:IMP
 'Tell anybody.'
- (77) *jāhā-te cal-ak' ləi-oʔo-a-ŋ-me*
 any-LOC go-M tell-away-APPL-1SG:OBJ-2SG:IMP
 'Tell me when thou go anywhere.'
- (78) *jāhā disəm-te-bon dəʔ-a*
 any country-LOC-1PL.EX:SUBJ go away-FIN
 'We will go away to any country.'

TABLE 2.10: INDEFINITE PRONOUNS

Animate	Inanimate
<i>jāhāe</i> ‘anyone’	<i>jāhā</i> ‘anything’

Sometimes the inanimate form takes the definitive suffix *-tak’* to denote an inanimate object.

- (79) *jāhā-tak’-gε hatao-me*
 any-DEF-FOC take-2SG:IMP
 ‘Take any (of these).’

adəm, ‘some’, is also found to be used as an indefinite for animates, and for inanimates it takes the inanimate suffix *-ak’* (i.e. *adəmak’*), though more often it is used as an attribute rather than a pronoun. *εtak’*, ‘another’, is also used as an indefinite pronoun with *-ic’* for the animate and *-ak’* for the inanimate. Paradigms are listed in Table 2.11:

TABLE 2.11: OTHER INDEFINITES

	Animate	Inanimate
Singular	<i>adəm</i> ‘somebody’	<i>adəmak’</i> ‘something’
Dual	<i>adəmkɪn</i>	<i>adəmak’kɪn</i>
Plural	<i>adəmkɔ</i>	<i>adəmak’kɔ</i>
Singular	<i>εtak’ic’</i> ‘another person’	<i>εtak’ak’</i> ‘another thing’
Dual	<i>εtak’kɪn</i>	<i>εtak’ak’kɪn</i>
Plural	<i>εtak’kɔ</i>	<i>εtak’ak’kɔ</i>

3.1.7 Demonstratives

Santali has two types of demonstratives with three distinctions – proximate, distal, and remote. While the proximate is in relation to the speaker, the distal and remote are in relation to the addressee. Two types of demonstratives are **simple**, represented by the roots *no* ‘this’, *on* ‘that’, and *han* ‘yonder’; and **particularized**, represented by *ne* (also *ni*) ‘just this’, *en* ‘just that’, and *hen* ‘just that yonder’. Both types have ±animate distinction and intensified forms as well. They may be represented in tabular form as shown in Table 2.12.

The intensified demonstratives are derived from both simple and particularized demonstratives for the proximate by infixing <*k’*>; thus *nuk’ui* ‘this very one’, *nək’ɔy* ‘this very thing’ from the simple type, and *nik’i* ‘just this very one’ from the particularized type. Note that only two forms, one for the animate and the other for the inanimate, are derived from the simple type without dual or plural, and only one from the particularized type for animate only. It is to be noted that when someone or something is to be specified, it is to be done with the proximate, and not with the distal or remote.

There are also two other types of demonstratives, one referring to what is seen and the other referring to what is heard, detailed in Table 2.13. In the case of the former there are two distinctions, one for the distant and the other for the remote. But in the case of the latter, there is only distant. While the demonstratives for what is seen are derived from *on* ‘that’ and *han* ‘that yonder’, the demonstrative of sound is derived

TABLE 2.12: DEMONSTRATIVES

Type I: Simple Demonstrative	Animate	Inanimate
Proximate	<i>nui</i> ‘this person’	<i>noa</i> ‘this thing’
DL	<i>nukin</i>	<i>noakin</i>
PL	<i>noko</i>	<i>noako</i>
Distal	<i>uni</i> ‘that person’	<i>ona</i> ‘that thing’
DL	<i>unkin</i>	<i>onakin</i>
PL	<i>onko</i>	<i>onako</i>
Remote	<i>həni</i> ‘that person yonder’	<i>hana</i> ‘that thing yonder’
DL	<i>hankin</i>	<i>hanakin</i>
PL	<i>hanko</i>	<i>hanako</i>
Type II: Particularized Demonstrative	Animate	Inanimate
Proximate	<i>nii</i> ‘just this person’	<i>niə</i> ‘just this thing’
DL	<i>nikin</i>	<i>niəkin</i>
PL	<i>neko</i>	<i>niəko</i>
Distal	<i>ini</i> ‘just that person’	<i>inə</i> ‘just that thing’
DL	<i>inkin</i>	<i>inəkin</i>
PL	<i>enko</i>	<i>inəko</i>
Remote	<i>hini</i> ‘just that one yonder’	<i>hinə</i> ‘just that thing yonder’
DL	<i>hinkin</i>	<i>hinəkin</i>
PL	<i>henko</i>	<i>hinəko</i>

TABLE 2.13: DEMONSTRATIVES REFERRING TO SIGHT AND SOUND

	Demonstrative ref. to sight		Demonstrative ref. to sound
	Distal	Remote	Distal
Singular	<i>əne</i> ‘that seen over there’	<i>hane</i> ‘that seen over yonder’	<i>ətə</i> ‘that heard over there’
Dual	<i>ənekin</i>	<i>hanekin</i>	<i>ətəkin</i>
Plural	<i>əneko</i>	<i>haneko</i>	<i>ətəko</i>

from *ət* (possibly *ot*). As both the forms end in *-ə*, it may be that it is a suffix pointing to some direction. The sets have no \pm animate distinction.

3.1.8 Numerals

Numerals in Santali are generally found as quantifiers combined with classifiers. The actual number is denoted by the quantifiers used without classifiers in enumerating human beings in the indefinite, and with classifiers in enumerating human beings with definite and non-human animate and inanimate objects.

3.1.8.1 Cardinal

There are ten basic numerals in Santali, from 1 to 10, and further numerals are derived from these basic numerals. There are also two other numerals for ‘twenty’ and ‘hundred’, which are also used as bases for further numerals, but they are borrowed from Indo-Aryan.

Basic native cardinals	Borrowed
<i>mit</i> ‘one’ <i>turui</i> ‘six’	<i>isi</i> ‘twenty’
<i>bar</i> ‘two’ <i>eae</i> ‘seven’	<i>sae</i> ‘hundred’

<i>pε</i>	'three'	<i>irəl</i>	'eight'
<i>pon</i>	'four'	<i>arε</i>	'nine'
<i>mōɽē</i>	'five'	<i>gɛl</i>	'ten'

There are two systems of counting, one based on 'ten', that is, decimal; and the other by scores, that is, vigesimal. From 'eleven' through 'nineteen' numerals are formed by adding numerals from 'one' to 'nine' to the base that is 'ten'. From 'twenty' onwards decades are formed by the process of multiplication and the numerals between the decades are the result of multiplication and addition. In multiplication the base is a multiplicand and that which multiplies is a multiplier. In addition the base is an augend and that which is added is an addend. The formation of the numerals may be given in a formulaic shape.

Numerals from 'eleven' to 'nineteen': AUGEND + ADDEND

$gɛl + mit' = gɛlmit'$ 'eleven'

$gɛl + irəl = gɛlirəl$ 'nineteen'

Numerals for decades : MULTIPLIER × MULTIPLICAND

$bar \times gɛl = bargɛl$ 'twenty'

$pe \times gɛl = pegɛl$ 'thirty'

Intermediate numerals: MULTIPLIER × MULTIPLICAND + ADDEND

$bar \times gɛl + mit' =$ 'twenty-one'

$pon \times gɛl + eae =$ 'forty-seven'

All the numerals are derived in this way until 'ninety nine'. For 'hundred' Santali does not have any native numeral. It is *sae* (cp. Skt. *śatam* 'hundred', Bangla *ś* 'id.'), borrowed from Indo-Aryan.

In the vigesimal system, that is, counting by scores, the same process operates, only the base for 'twenty' is *isi*. Even decades are multiples of 'twenty' and uneven decades are a combination of 'twenty' and 'ten'.

$mit' + isi = mit'isi$ 'twenty'

$bar + isi = barisi$ 'forty'

$barisi + gɛl + mit' =$ 'fifty-one'

3.1.8.2 Classifiers

There are three sets of classifiers in Santali.

Set I: *ʃɛn* (with variant *ʃɛc*): This classifier is used with the numeral for 'one' and signifies non-living human beings, non-human living beings, non-human non-living objects and the inanimate. It is also used for human beings when specified.

$mit'ʃɛn gɔc' hɔɽ$ 'one dead man' $mit'ʃɛc' uric$ 'one bullock'

$mit'ʃɛn rakkhos$ 'one man-eater' $mit'ʃɛn hətɪər$ 'one weapon'

$mit'ʃɛn hɔɽ$ 'one man'

Set II: *ea*: This classifier is used with the numerals from 'two' to 'four' and for 'twenty', as well as the same classes of nouns as in Set I.

barea boŋga ‘two ghosts’ *pea sim* ‘three cocks’
ponea jinis ‘four things’ *ponea gate* ‘four friends’

Set III: *gɔʔen* (with variant *gɔʔec*): This classifier is used with the numerals from ‘five’ to ‘ten’ and with the distributive numerals. It is also used rarely with the numeral ‘one’. Occurrences with ‘one’ are, however, found in songs and presumably demanded for metrical purpose. The classifier is used with the same classes of nouns as in Sets I and II.

mɔ̃ʒẽ-gɔʔen kaɖa ‘five buffaloes’ *turui-gɔʔen boŋga* ‘six ghosts’
eae-gɔʔen putul ‘seven dolls’

3.1.8.3 Distributive numeral

Distributive numerals in Santali are formed by reduplication of the initial consonant along with the vowel. Bodding (1929:60) has given examples of such formations, from ‘one’ through ‘ten’. In my data only the first five numerals form their distributive counterpart in this process, and any other distributives are formed with *kate*. Thus:

mit ‘one’ > *mimit* ‘one each’
bar ‘two’ > *babar* ‘two each’
pe ‘three’ > *pepe* ‘three each’
pon ‘four’ > *popon* ‘four each’
mɔ̃ʒẽ ‘five’ > *mɔ̃mɔ̃ʒ* ‘five each’
turui-kate ‘six each’

3.1.8.4 Inclusive numeral

Santali possesses three inclusive numerals – *banar*, ‘both’; *pene*, ‘all three’; and *ponon*, ‘all four’ – derived by infixing <n> to the base numerals for ‘two’ to ‘four’. They can be used as a subject or object of a sentence and as an attribute in endocentric attributive constructions.

- (80) *banar-ge tuʔi-kin-me*
 both-EMPH pierce-3DL:OBJ-2SG:IMP
 ‘Pierce the both.’
- (81) *pene əgu-akat’-ko-a-ŋ*
 all three bring-PRF:A-3PL:OBJ-FIN-1SG:SUBJ
 ‘I have brought all three.’
- (82) *ponon kombɔo-ko gɔc’-en-a*
 all four thief-3PL:SUBJ die-PST:M-FIN
 ‘All four thieves died.’

3.1.8.5 Ordinal

Santali has three ordinal numbers – *pəhil* ‘first’, *dɔsar* ‘second’, and *tesar* ‘third’ – all borrowed from Indo-Aryan. Sometimes the ±animate suffixes *-ic* and *-ak* are also added. Each of them can occur as attribute or adverb.

Santali, being a neighbour of the eastern Indo-Aryan languages like Bangla, Hindi, Oriya, and Assamese, has direct contact with speakers of all these languages. Among other linguistic features, Santali also uses Indo-Aryan numerals. As a result, the first six numerals are used uniformly by all sections of society. The younger generation more often uses Indo-Aryan numerals from seven onwards while the older generation preserves the original numerals.

3.1.9 Adpositions

Santali has a large number of postpositions. The postpositions are added either to the bare nominals or to the number suffixes and the definitive marker. Some of them require the genitive case. Some are used after infinitives and others after a complete phrase – even after a sentence. Some postpositions are complex in the sense that they are composed of one postposition and a suffix. Some postpositions can take derivative suffixes for adjective and \pm animate nouns. Here are some of the common postpositions:

lagat' llagit'. To indicate purpose or intention after all nominals and infinitives. It may also be combined with the suffix *-te* without any change in meaning.

- (83) *uni raj hɔpɔn-lagit' dɔ bis ləɖu əgu-ad-e-a-e*
 3SG king son-for TOP poison sweet bring-APPL:PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'He brought poisonous sweet for the prince.'

- (84) *am-lagit'-ijñ hec'-aka-n-a*
 2SG-for-1SG:SUBJ come-PRF-M-FIN
 'I have come for thou.'

- (85) *uduk'-lagit'-e calao-aka-n-a*
 show-for-3SG:SUBJ go-PRF-M-FIN
 'He has come for showing.'

modre. 'among'. Bangla word with *-re*, i.e. *moddhere* is also found to occur in the same sense.

- (86) *noko-modre kombɔ mena-e-a*
 these-among thief have-3SG:OBJ-FIN.
 'There is a thief among these persons.'

- (87) *ape-moddhere jãhãe-ge lɔgɔn hij-uk'-me*
 2PL-among anyone-EMPH quickly come-M-2SG:IMP
 'Any one of you come quickly.'

dhəbic'. Convey the sense 'till, until, up to'. Used after nominals indicating space and after infinitives indicating time.

- (88) *cuɽa-dhəbic' ãɽgo-en-a-e*
 peak-upto climb-PST:M-FIN-3SG:SUBJ
 'He climbed up to the peak.'

- (89) *uni əuri hij-uk'-dhəbic' okate-hɔ alo-m cala-k'-a*
 3SG before come-M-until where-too PROHIB-2SG:SUBJ go-M-FIN.
 'Do not go anywhere until he comes.'

bhitrire. ‘inside’ with locative suffix *-re*, hence ‘within’ too.

- (90) *oçak'-bhitrire mena-e-a*
house-inside have-3SG:OBJ-FIN.
‘He is there in the house.’

talare. locative of *tala* ‘middle’ with locative suffix *-re*.

- (91) *oçak'-talare duçup'-aka-n-a-e*
house-middle in sit-PRF-M-FIN-3SG:SUBJ
‘He is sitting in the middle of the house.’

latarre. locative of *latar* ‘below’ with suffix *-re*, used only with noun.

- (92) *dhiri-latarre*
stone-under
‘Under (a) stone.’

cetanre. ‘above, top’, loc of *cetan* ‘top’, with suffix *-re*.

- (93) *çuyri-cetanre mit'-çen dare tahëkan-a*
hill-under one-CLSSFR tree PST-FIN
‘There was a tree under the hill.’

leka. Adjectival and adverbial postposition meaning ‘like’. It is also found to occur with locative suffix of motion *-te* in the sense of ‘by any means’.

- (94) *nui-leka hçr kan-a-e*
DEM-like man be-FIN-3SG:SUBJ
‘He is a man like this person.’
- (95) *oka-lekate hõ noa dõ alo-m em-çk'-a*
which-by means too this TOP PROHIB-2SG:SUBJ give-M-FIN
‘Thou would not give it by any means.’

ate. ‘taking alongwith’.

- (96) *uni læhi-ate taçam-ed-a-e*
3SG:SUBJ stick-take along walk-PRS:A-FIN-3SG:SUBJ
‘He is walking taking a stick along.’

hõteç'te. ‘for, by, due to’, indicating instrumentality.

- (97) *dak' bõde-hõteç'te bõ-n topo-le-n-a*
water dirty-due to NEG-1SG:SUBJ bathe-PLUP-M-FIN
‘I did not take bath due to dirty water.’

tuluc'. ‘being with, association with’, used with nominals and verb stems.

- (98) *ijn-tuluc'-e hec'-akan-a*
1s-along with-3SG:SUBJ come-PRF-M-FIN
‘He has come with me.’

- (99) *sen-ək'-tuluc'-e* *jəpit'-idi-ed-a*
 walk-M –while-3SG:SUBJ sleep-simultaneously-IMPRF:A-FIN
 'While walking he sleeps.'

kate. Gerundial postposition, meaning 'having done'. It is also used after nominals and adjectives in an adverbial sense.

- (100) *ruəṛ- hēc'-kate* *uni* *noa* *katha* *ləi-ked-a-e*
 return -come-CV he this story tell-PST:A-FIN-3SG:SUBJ
 'Having returned he told this story.'
- (101) *phəsiara-kate* *hatao-ked-a-e*
 deceitful-CV get-PST:A-FIN-3SG:SUBJ
 'He got possession of it deceitfully.'

mente. Lit. 'by saying' (*men* 'say'+ *te*), 'for the purpose of'.

- (102) *ṭheṅga-mente* *mak'-əgu-akad-a-e*
 stick-for the purpose of cut-bring-PRF:A-FIN-3SG:SUBJ
 'He cut and brought it for the purpose of making stick.'

iəte. Used after nominals and infinitives, with a sense of 'owing to, due to, on account of'.

- (103) *əḍi* *həṛ-iəte* *ba-e* *sor-len-a*
 many people-owing to NEG-3SG:SUBJ come close-PLUP:M-FIN
 'He did not come closer on account of multitude of people.'

3.1.10 Derivation

There are three processes to derive nominals in Santali: prefixation, infixation, and suffixation. The processes are employed by verbs, lexical and derived adjectives, and nouns to derive the nominals.

3.1.10.1 Suffixation

The suffixes *-ic'* for animate and *-ak'* for inanimate may be described as nominalizers, too (Neukom 2001:57–58). They are used to form referential nominals from:

Reduplicated verb stems and verb roots

<i>dadalic'</i>	'one who engaged in beating'	< <i>dal</i>	'beat'
<i>jəmak'</i>	'food'	< <i>jəm</i>	'eat'
<i>gəc'ic'</i>	'dead one'	< <i>gəc'</i>	'die'
<i>əllic'</i>	'writer'	< <i>əl</i>	'write'
<i>əlak'</i>	'that which is written'		
<i>rəḥəyak'</i>	'that which is sown'	< <i>rəḥəy</i>	'sow'
<i>enec'ic'</i>	'dance'	< <i>enec'</i>	'dance'
<i>kirijnak'</i>	'that which is bought'	< <i>kirij</i>	'buy'

Simple and derived adjectives

<i>ponḍic'</i>	'white one',	<i>ponḍak'</i>	'white thing'	<i>ponḍ</i>	'white'
<i>bogeic'</i>	'good one',	<i>bogeak'</i>	'good thing'	<i>boge</i>	'good'

nɔtenic ‘one of this side’, *nɔtenak* ‘thing of this side’ <*nɔten* ‘belonging to this side’ <*nɔte* ‘this side’
nɔndɛnic ‘one of this place’, *nɔndɛnak* ‘thing of this place’ <*nɔndɛn* ‘belonging to this place’ <*nɔndɛ* ‘here’
noŋkanic ‘one like this one’, *noŋkanak* ‘thing like this’ <*noŋkan* ‘like this’ <*noŋka* ‘in this manner’

Postposition

sāotenic ‘companion’ <*sāoten* ‘accompanying’ <*sāo* + *te*

Suffix

ɔlteak ‘that with which is written(pen)’ <*ɔl* ‘write’ *te* ‘instrumental suffix’

3.1.10.2 Infixation

The most productive process of nominal derivation in Santali is infixation. There are at least five infixes – <*tV*>, <*nV*>, <*mV*>, <*ɽV*>, and <*pV*> – of which <*tV*>, <*nV*> and <*pV*> are the most productive. The other two, <*mV*> and <*ɽV*> are used rarely. They are inserted into verbs, nouns, and lexical adjectives to derive nominals. The vowel of the infix (normally the first vowel if the root or stem is bisyllabic) is that of the root to which it is inserted, the exception being *ɛhɔp* ‘begin’, which repeats the second vowel.

Examples:

<*tV*>: *bɔtɔr* ‘fear’, <*bɔr* ‘to fear’; *rɔtɔk* ‘seam’, <*rɔk* ‘sew’; *jutum* ‘name’, <*jum* ‘to name’; *ɔtɔr* ‘warp of a web’, <*ɔr* ‘draw/pull’; *ɛtɔhɔp* ‘beginning’, <*ɛhɔp* ‘begin’; *gɔtɔɔ* ‘help’, <*gɔɔ* ‘to help’; *ɔtɔmɔn* ‘origin’, <*ɔmɔn* ‘germinate’; *satahet* ‘breath’, <*saht* ‘breathe’; *jelen* ‘long’, <*jetelen*, length’.

<*nV*>: *ɔnɔl* ‘written piece’, <*ɔl* ‘write’; *jɔnɔk* ‘broom’, <*jɔk* ‘sweep’; *banak* ‘hook’, <*bak* ‘to hook’; *benet* ‘stopper/lid’ <*bet* ‘cover with palm’; *gɔnɔŋ* ‘bride price’, <*gɔŋ* ‘give to marriage’; *tɔnɔl* ‘knot’, <*tɔl* ‘bind’; *ranakap* ‘up, development’, <*rakap* ‘rise’; *ɔɽ* ‘begin’ > *ɔnɔɽ* ‘preface’; *ɔnɔsar* ‘breadth’, <*ɔsar* ‘broad’.

<*pV*>: *hɔpɔn* ‘children’, <*hɔn* ‘child’; *rapaj* ‘king and his retinue’, <*raj* ‘king’; *məpəŋji* ‘village chiefs’, <*məŋji* ‘village chief’; *kipisɔɽ* ‘zamindar and his retinue’, <*kisɔɽ* ‘zamindar’. The infix <*pV*> is often employed to derive plural nouns.

<*mV*>: *lamak* ‘scraper’, <*lak* ‘peal, scrap’; *cemet* ‘teaching aid’, <*cet* ‘teach’; *semlet* ‘association’, <*set* ‘associate’; *hɔmɔn* ‘nephew’, <*hɔn* ‘child’.

<*ɽV*>: *gɔɽɔm* ‘grand old, namesake’, <*gɔm* ‘to name’; *cɛɽɛ* ‘bird’, <*cɛ* ‘squeak’.

3.1.10.3 Prefixation

There is only one unproductive prefix in Santali, *ma-*, which serves as a nominalizer in converting verbs to nouns. It is very restricted in use, and only three verb roots

take this prefix for nominalization. In some cases it carries the meaning of result and in others that of the active agent.

marsal ‘light’, <*arsal* ‘to light’; *marak* ‘peacock’, <*rak* ‘cry’; *macet* ‘teacher’, <*cet* ‘teach’.

3.1.11 Adjectives

There is little justification for positing a separate adjectival class except for a few derived adjectival words, which function as attributes in endocentric attributive constructions and as predicate complements. There are, of course, a handful of borrowed words that are adjectives at the source and thus may be treated as typical adjectives. The borrowed words show ± sex distinction, otherwise there is no gender distinction as such. Borrowed pairs like *khepa:khepi* ‘mad’, *kala:kəli* ‘deaf’, *koŋka:kuŋki* ‘mad’, *kāṛā:kāṛi* ‘blind’, *totra:tutri* ‘stammering’, *kōda:kūdi* ‘dumb’, and *phogra:phugri* ‘toothless’ are declined for gender with *-a* in the masculine and *-i* in the feminine.

-an is the suffix for deriving adjectives from nouns and demonstrative adverbs of location, manner, and direction: *daṛean koṛa* ‘strong boy’, <*daṛe* ‘strength’; *kaḍawan hoṛ* ‘buffalo-having man’, <*kaḍa* ‘buffalo’; *dayawan kuṛi* ‘kind girl’, <*daya* ‘kindness’; *nōndən macet* ‘teacher of this place’, <*nōndə* ‘here’; *nōten daŋgra* ‘bullock of this side’ <*nōte* ‘this side’; *noŋkan kəmi* ‘work like this’ <*noŋka* ‘like this’.

The following are some examples of their use as a predicate complement:

- (104) *uni dō dayawan-kan-a-e*
 3SG. TOP kind-COP-FIN-3SG:SUBJ
 ‘He is kind.’
- (105) *ona ənəl dō noŋkan-ge-a*
 that writing TOP like this-FOC-FIN
 ‘That essay is like this.’

The word itself is not inflected to show degrees of comparison. To obtain comparative and superlative degrees, postpositions are used with the words with which something is to be compared. *khən* is used for the comparative and words of multitude with the same postposition for superlative.

- (106) *hana dare noa dare-ko khən dō sēṛa-ge-a*
 that yonder tree this tree-PL-ABL TOP big-FOC-FIN
 ‘That tree is bigger than this tree.’

3.1.12 Adverbials

There are three types of adverbs: simple, derivative, and words with locatives suffixes.

usəra ‘quickly’ as in (107) *usəra kəmi-me*
 quickly do-2SG:IMP
 ‘Do quickly.’

dəme ‘very much’ as in (108) *dəme-ko rak’-ed-a*
 very much-3PL:SUBJ cry-IMPFRF:A-FIN
 ‘They are crying very much.’

logɔn ‘hurriedly’ as in (109) *logɔn hij-uk’-mɛ*
 hurriedly come-M-2SG:IMP
 ‘Come hurriedly.’

eskar ‘alone’ as in (110) *eskar taken baŋ boge-a*
 alone stay NEG good-FIN
 ‘Staying alone is not good.’

nahak’ ‘presently’ as in (111) *nahak’-e hij-uk’-a*
 presently-3SG:SUBJ come-M-FIN
 ‘He will come presently.’

tis ‘when’ as in (112) *tis-em hij-uk’-a*
 when-2SG:SUBJ come-M-FIN
 ‘When will you come?’

Some other words are *nit* ‘now’, *cɔʔ* ‘immediately’, *laha* ‘first’, *un* ‘then’, *netar* ‘at present’, *seday* ‘in old days’, *acka* ‘suddenly’, *pəhil* ‘at first’, *dinəm* ‘daily’, *enaŋ* ‘then’, *dhinəŋ* ‘after a while’.

The suffixes *-ka*, *-ɖɛ*, *-te* are added to the demonstrative and interrogative roots to derive adverbs denoting location, motion, and manner: *noŋka* ‘in this manner’; *nɔte* ‘in this direction’; *nɔndɛ* ‘here’; *okare* ‘where’; *okate* ‘to which side’; *ceka* ‘how’.

Reduplicated words with or without *-te* and words along with their echo counterparts are used as adverbs.

(113) *pəri pəri-te kəmi-me*
 turn-INS do-2SG:IMP
 ‘Do by turns.’

(114) *kəc’ kəc’ dak’-et’-tahēkan-a-e*
 little rain-IMP:FIN-3SG:SUBJ
 ‘It was raining a little.’

(115) *hakopako hij-uk’-mɛ*
 quickly come-M.-2SG:IMP
 ‘Come quickly.’

(116) *sen-ək’ sen-ək’-te-e mən-ked- a*
 go-M go-M-CV-3SG:SUBJ say-PST:A-FIN
 ‘While going he spoke.’

3.2 Verbal morphology

The Santali verb is defined by the fact that it obligatorily takes tense, aspect and mood markers, markers for voice, pronominal arguments, marker for the finiteness of action, and sentential modality. The words that denote action, event, or condition are not those treated as verbs, but any word is treated as such provided it takes the above markers. The verb with the finite marker *-a* is treated as finite, and without this and the gerundial suffix is treated as non-finite.

Typical of the Munda languages and Santali is the fact that the verb may be any type of lexeme, any nominal with or without a genitive suffix, interrogatives, and indefinite stems, besides an exclusive class of verb roots. In the case of nominals,

interrogatives, and indefinites, there is always the need of a copula (light verb) to be used with it. Consider the following examples:

- (117) *uni dɔ am-ak'-kan-a e*
 3SG TOP 2SG-GEN-COP-FIN.-3SG:SUBJ
 'He is yours.'
- (118) *ɔkɔe-kan-a-m*
 who-COP-FIN-2SG:SUBJ
 'Who are you?'
- (119) *oka-reak'-kan-a*
 which-GEN-COP-FIN.
 'Where does it belong to?'
- (120) *ɔjɔn-ad-e-a-ŋ*
 medicine-APPL:PST:A-3SG:OBJ-FIN.-1SG:SUBJ
 'I gave him medicine.'
- (121) *māi dɔ pargana-e kimin-ked-e-tjŋ-a*
 daughter TOP Pargana-3SG:SUBJ daughter-in-law-PST:A-3SG:OBJ-POSS-1POSS-FIN
 'A Pargana made my daughter his daughter-in-law.'

Note that in examples 117–121 a pronoun, interrogative, and noun are used in the predicate position with either a copula or a tense marker.

3.2.1 Subject

Santali is a nominative/accusative type of language in terms of verb agreement. The verb obligatorily agrees with the subject NP in terms of person and number. The subject is marked by the clitic pronominals in the verb phrase freely standing after the verb or with the word preceding it. The pronominals have the same shape as personal pronouns, except that the third person is marked by *-e*, *-kin*, and *-ko* in the singular, dual, and plural. When an animate noun stands as the subject NP, it agrees with the verb in the form of third person clitic pronouns. When the subject NP is a pronoun, it agrees with the verb by its clitic form. Table 2.14 gives an overview of subject markers on the verb:

TABLE 2.14: SANTALI AGREEMENT MARKERS

Person/Number	Singular	Dual		Plural	
		INCL	EXCL	INCL	EXCL
First	<i>-ŋ(ijŋ)</i>	<i>-laŋ</i>	<i>-liŋ</i>	<i>-bon</i>	<i>-le</i>
Second	<i>-m</i>		<i>-ben</i>		<i>-pe</i>
Third	<i>-e</i>		<i>-kin</i>		<i>-ko</i>

Note that the clitic form of the first personal pronoun is replaced by the full form *ijŋ* when it is added to a consonant-ending word. The first person plural exclusive gets an increment *n* in its clitic form. The second person singular *-m* gets an increment *e* when added to a consonant-ending word, thus *em*. The third person clitic forms

are actually number suffixes for the dual and plural. When a negative particle precedes the verb, the clitic pronominals are obligatorily marked in the particle. There is one marker, *-k'*, for an inanimate subject used exclusively with the verbs *mena* 'to be' and *hena* 'to have'.

- (122) *cet' jinis hena-k'-taben-a*
 what thing exist-M-POSS-2nd dl.POSS-FIN
 'What thing is there of you?'
 (123) *uni-ak' oʔak'-re mit'-ʔaŋ ʔhili mena-k'-a*
 3S.-GEN house-LOC one-CLSSFR pitcher exist-M-FIN
 'There is a pitcher in (his) house.'

It should be noted that the verbs *mena* and *hena* take the subject agreement in the position where the object normally comes, that is, not after a finite *a*, but before it.

Examples for illustrating positions of subject pronominal clitic:

- (124) *khan-ge paʔoari do-e dəʔ-ked-a*
 then-FOC Patoari TOP-3SG:SUBJ run-PST:A-FIN

Also

khan-ge paʔoari do dəʔ-ked-a-e
 'Then Patoari ran away'

- (125) *ba-e sen-len-a*
 NEG-3SG:SUBJ go-PLUP:M-FIN
 'He had not gone.'

3.2.2 Object types

The marking of pronominal object plays an important role in Santali. The animate objects are marked in the verb in the form of infix clitic pronominals. There are two types of objects in Santali: one direct, marked by the clitic pronominals in infix form; and the other indirect, marked by an Applicative *a-*. In the simple present/future the Applicative is prefixed to the clitic pronominals, and with TAM it is prefixed to the TAM markers.

- (126) *dal-a-ŋ-a-e*
 strike-APPL-1SG:OBJ-FIN-3SG:SUBJ
 'He strikes/will strike for me.'
 (127) *dal-a-ŋ-kan-a-e* 'He is striking for me.'
 -COP-
 (128) *dal-ad-ŋ-a-e* 'He struck for me.'
 -APPL.PST:A
 (129) *dal-akawad-ŋ-a-e* 'He has struck for me.'
 -APPL.PRF:A-

The clitic pronominals for the object stand after the TAM markers and before copula *kan* and *tahēkan*. The clitic pronominals for the object are the same as the subject

forms, except for second person singular, which is <me>. Examples illustrating the position of direct object in the verb are shown in Table 2.15:

TABLE 2.15: DIRECT OBJECT PRONOMINAL INFIXES

	Singular		Dual		Plural
First	<i>dal-ij-a-e</i>	<i>dal-lay-a-e</i>	<i>dal-lij-a-e</i>	<i>dal-bon-a-e</i>	<i>dal-le-a-e</i> ‘He strikes me/us’
Second	<i>dal-me-a-e</i>		<i>dal-ben-a-e</i>		<i>dal-pe-a-e</i> ‘He strikes thou/you’
Third	<i>dal-e-a-e</i>		<i>dal-kin-a-e</i>		<i>dal-ko-a-e</i> ‘He strikes him/them’

Further examples:

- (130) *dal-kid-**ij**-a-e*
strike-PST:A-1SG:OBJ-FIN-3SG:SUBJ
‘He struck me.’
- (131) *dal-led-e-a-ko*
strike-PLUP:A-3SG:OBJ-FIN-3PL:SUBJ
‘They had struck him.’
- (132) *dal-akat’-**le**-a-e*
strike-PRF:A-1PL:OBJ-FIN-3SG:SUBJ
‘He has struck us.’
- (133) *dal-k-**ij**-a-e*
strike-OPT-1SG:OBJ-FIN -3SG:SUBJ
‘He should strike me.’
- (134) *dal-le-**m**-a-e*
strike-IRR-2SG:OBJ-FIN-3SG:SUBJ
‘He would strike you.’

3.2.3 Tense

In Santali two copulas, *kan* and *tahēkan*, express time dimension for the present and past, respectively. There is the concept of time for present, past, and future, and the aspect suffixes along with their own functions denote time dimension as well.

The copula *kan* denotes present time. It is quite different from other aspect-cum-tense suffixes. While aspect-cum-tense suffixes take an object after them, *kan* takes its object before it. It has, again, nothing to do with voice markers for active or middle; it can be used with both active and middle. It can also be independently conjugated, thus:

- (135) *kan-a-ŋ* ‘I am’
be-FIN-1SG:SUBJ
- (136) *kan-a-m* ‘Thou art’
- (137) *kan-a-e* ‘He is’

It is secondarily shifted to the non-completive determinative progressive in the middle and with first and third person singular object in the active.

(138) Active : *uni dal-ij-kan-a-e*
 3sS beat-1SG:OBJ-COP-FIN-3SG:SUBJ
 ‘He is beating me.’

Middle : *uni cal-ak'-kan-a-e*
 3sS go-M-COP-FIN-3SG:SUBJ
 ‘He is going.’

It is also used with the Iterative/Intensive stem to denote non-completive determinative progressive.

(139) *dadal-kan-a-e*
 beat intensively-COP-FIN-3SG:SUBJ
 ‘He is beating intensively.’

Corresponding to the *kan* copula for the past is *tahēkan*. Like *kan*, it is independently conjugated.

(140) *tahēkan-a-ŋ* *tahēkan-a-m* *tahēkan-a-e*
 COP:PST-FIN-1SG:SUBJ COP:PST-FIN-2SG:SUBJ COP:PST-FIN-3SG:SUBJ
 ‘I was’ ‘Thou were’ ‘He was’

Unlike *kan*, it never functions as an aspect marker. In combination with different aspect suffixes, it always denotes past action or state. It can combine with the copula *kan* as well to convey the sense of past progressive. The tense nuances expressed by this copula will be shown in the aspect-cum-tense network. Zero or the absence of any marker denotes the simple present/future.

3.2.4 Aspect

Santali employs a number of suffixes to denote different aspects. It is interesting to note that the suffixes employed to denote different states of action are also employed to denote time dimension, thus serving a dual role. Aspects in Santali may be divided broadly into non-completive and completive, each having two broad divisions – indeterminate and determinative. Non-completive indeterminate is indefinite, while the same determinative is progressive. Completive indeterminate is aorist while the same determinative is either resultative, stating an accomplished action with a result which is still present; or non-resultative, stating an action which happened long ago with a result which is no longer present or is an action to be taken up prior to some other action. In a tabular form the aspect framework of Santali may be represented as follows:

<i>Non-completive</i>		
Indeterminative:	Indefinite	∅
Imperfective:	Progressive	
	Active	-e
	Active & Middle	<i>kan</i>
<i>Completive</i>		
Indeterminative:	Aorist	
	Active	-ke
	Middle	-e

Determinative	Resultative	-aka
		(accomplished with the result still present)
	Non-resultative	-le
		(accomplished with the result not present)

The non-resultative *-le* is also used to denote priorative action. All aspect suffixes except \emptyset and *kan* are added with active and middle voice markers when they are suffixed to the verb stems. Aspect suffixes with active and middle markers are shown in Table 2.16:

TABLE 2.16: ASPECT SUFFIXES WITH ACTIVE AND MIDDLE MARKERS

<i>Non-completive</i>			
Indeterminative	Active	\emptyset	
	Middle	-ok' \emptyset	
Imperfective progressive	Active	-et'	
	Middle	-ok' kan	
<i>Completive</i>			
Indeterminative aorist	Active	-ket'	
	Middle	-en	
Determinative resultative	Active	-akat'	
	Middle	-akan	
Determinative non-resultative	Active	-let'	
	Middle	-len	
Determinative priorative	Active	-le	
	Middle	-len	

Note that determinative priorative active does not take any marker in the active. The aspect-cum-tense network can be enumerated as follows:

Non-past

Non-completive indeterminative as *simple present/future*:

Active	\emptyset
Middle	-ok' \emptyset

Non-completive determinative progressive as *present progressive*:

Active	-et'
Middle	-ok' kan

Past

Completive indeterminative aorist as *simple past*:

Active	-ket'
Middle	-en

With Applicative²¹ Active -at'

Middle -an

With Benefactive²² Active -kat'

Completive determinative resultative as *perfect*:

Active	-akat'
Middle	-akan

With Applicative	Active	-akawat'
	Middle	-akawan
Completive determinative resultative with <i>tahēkan</i> as <i>past perfect</i> :		
	Active	-akat' <i>tahēkan</i>
	Middle	-akan <i>tahēkan</i>
Completive determinative non-resultative as <i>pluperfect</i> :		
	Active	-let'
	Middle	-len
With Applicative	Active	-at'
	Middle	-an
With Benefactive	Active	-kat'
Completive determinative non-resultative with <i>-tahēkan</i> as <i>anterior pluperfect</i> :		
	Active	-let' <i>tahēkan</i>
	Middle	-len <i>tahēkan</i>
Non-completive determinative progressive as <i>past progressive</i> :		
	Active	-et' <i>tahēkan</i>
	Middle	-ok' <i>kan tahēkan</i>

Examples illustrating use of TAM suffixes with two groups of stems – active and middle:

Simple present:

Active:

Transitive root

dal-e-a-e

beat-3SG:OBJ-FIN-3SG:SUBJ

'He beats him.'

Transitive-Intransitive root

gitic'-e-a-e

lay-3SG:OBJ-FIN-3SG:SUBJ

'He lays him.'

Causative²³

ne1-oco-ɲ-a-e

see-CAUS-1SG:OBJ-FIN-3SG:SUBJ

'He causes/allows me to see.'

Benefactive stem

tɔl-ka-e-a-e

bind-BEN-3SG:OBJ-FIN-3SG:SUBJ

'He binds (the cow) for someone.'

Compound stem(TR.+TR.)

sendra-ɲam-e-a-e

seek and find-3SG:OBJ-FIN-3SG:SUBJ

'He searches and finds him.'

Middle:

Intransitive root

senɔk'-a-e

go-M-FIN-3SG:SUBJ

'He goes.'

Transitive-Intransitive root

gitic'-a-e

lie-FIN-3SG:SUBJ

'He lies.'

Reciprocal stem

dapal-a-kin

beat each other-FIN-3DL:SUBJ

'They (two) fight.'

Passive/Reflexive stem

dal-ok'-a-e

beat-M-FIN-3SG:SUBJ

'He is beaten/beats himself.'

Compound stem(INTR.+INTR.)

rakap'-hij-uk'-a-e

ascend and come-M-FIN-3SG:SUBJ

'He ascends and comes.'

Present progressive:

Transitive root

dal-ed-e-a-e

beat-PROG:A-3SG:OBJ-FIN-3SG:SUBJ

'He is beating him.'

dal-e-kan-a-e

beat-3SG:OBJ-COP-FIN-3SG:SUBJ

'He is beating him.'

Causative stem

nel-oco-et'-me-a-e

see-CAUS-IMPRF:A-2SG:OBJ-FIN-3SG:SUBJ

'He is causing/allowing me to see.'

dal-oco-e-kan-a-e

beat-CAUS-3SG:OBJ-COP-FIN-3SG:SUBJ

'He is causing/allowing him to beat.'

Benefactive stem

tɔl-ka-e-kan-a-e

bind-BEN-3SG:OBJ-COP-FIN-3SG:SUBJ

'He is binding it(cow) for someone.'

Compound stem(TR.+TR.)

mak'əgu-ed-a-e

cut bring-IMPRF:A-FIN-3SG:SUBJ

'He is cutting and bringing.'

Simple past:

Transitive root

dal-ked-e-a-n

beat-PST:A-3SG:OBJ-FIN-1SG:SUBJ

'I beat him.'

With Applicative

ləi-at'-ko-a-e

tell-APPL.PST:A-3PL:OBJ-FIN-3SG:SUBJ

'He told them.'

Causative stem

dal-oco-kid-ijn-a-e

beat-CAUS-PST:A-1SG:OBJ-FIN-3SG:SUBJ

'He caused/let me to beat.'

Benefactive stem

uyuk'-kad-e-a-n

cover-BEN.PST:A-3SG:OBJ-FIN-1SG:SUBJ

'I covered it (to help someone).'

Compound stem(TR.+TR.)

nel-ɲam-ked-a-e

see find-PST:A-FIN-3SG:SUBJ

'He saw and found.'

Intransitive root

sen-ɔk'-kan-a-e

go-M-COP-FIN-3SG:SUBJ

'He is going.'

Reciprocal stem

dapal-kan-a-kin

beat each other-COP-FIN-3DL:SUBJ

'They are fighting.'

Passive/Reflexive stem

nel-ɔk'-kan-a-e

see-M-COP-FIN-3SG:SUBJ

'He is being seen/seeing himself.'

Compound stem (INTR.+INTR.)

ruəŋhij-uk'-kan-a-e

return- come-M-COP-FIN-3SG:SUBJ

'He is returning.'

Intransitive root

sen-en-a-e

go-PST:M-FIN-3SG:SUBJ

'He went.'

With Applicative

ləi-an-a-e

tell-APPL.PST:M-FIN-3SG:SUBJ

'He told for his satisfaction.'

Causative stem

dal-oco-en-a-e

beat-CAUS-PST:M-FIN-3SG:SUBJ

'He was caused to/let beat.'

Passive/Reflexive stem

nel-en-a-e

see-PST:M-FIN-3SG:SUBJ

'He was seen/saw himself.'

Compound stem (INTR.+INTR.)

ruəŋ-hec'-en-a-e

return- come-PST:M-FIN-3SG:SUBJ

'He came back.'

Perfect:

Transitive root

dal-akad-e-a-m

beat-PRF:A-3SG:OBJ-FIN-2SG:SUBJ

‘You have beaten him.’

With Applicative ²⁴*l̥i-akawad-e-a-m*

tell-PRF.Ap:A-3SG:OBJ-FIN-2SG:SUBJ

‘You have told him.’

Causative

dal-oco-akad-e-a-e

beat-CAUS-PRF:A-3SG:OBJ-FIN-3SG:SUBJ

‘He has been caused to beat.’

Compound stem (TR.+TR.)

mak’-əgu-akad-a-e

cut bring-PRF:A-FIN-3SG:SUBJ

‘He has cut and brought.’

Pluperfect:

Transitive root

dal-led-a-e

beat-PLUP:A-FIN-3SG:SUBJ

‘He had beaten.’

With Applicative

l̥i-ad-e-a-m

tell-APPL.PLUP:A-3SG:OBJ-FIN-2SG:SUBJ

‘You had told him.’

Causative

dal-oco-led-e-a-m

beat-CAUS-PLUP:A-3SG:OBJ-FIN-2SG:SUBJ

‘You had caused/permitted him to beat.’

Benefactive stem

pɔɕm-kad-a-e

cover-BEN.PLUP:A-FIN-3SG:SUBJ

‘He had covered (for someone).’

Intransitive root

sen-akan-a-e

go-PRF:M-FIN-3SG:SUBJ

‘He has gone.’

With Applicative

dal-akawan-a-e

beat-PRF.Ap:M-FIN-3SG:SUBJ

‘He has beaten (paddy) for himself.’

Causative

gɔɕ’-oco-akan-a-e

die-CAUS-PRF:M-FIN-3SG:SUBJ

‘He has caused/permitted to die.’

Reciprocal

dapal-akan-a-kin

beat each other-PRF:M-FIN-3DL:SUBJ

‘They have fought.’

Passive/Reflexive

dal-akan-a-e

beat-PRF:M-FIN-3SG:SUBJ

‘He has been beaten/has beaten himself.’

Compound stem (INTR.+INTR.)

ruəɾhɛc’-akan-a-e

return come-PRF:A-FIN-3SG:SUBJ

‘He has come back.’

Intransitive root

sen-len-a-e

go-PLUP:M-FIN-3SG:SUBJ

‘He had gone.’

With Applicative

ɲɛl-an-a-e

see-APPL.PLUP:M-FIN-3SG:SUBJ

‘He had seen for himself.’

Causative

dal-oco-len-a-m

beat-CAUS-PLUP:M-FIN-2SG:SUBJ

‘You had been caused/permitted to be beaten.’

Passive/Reflexive stem

ɲɛl-len-a-e

see-PLUP:M-FIN-3SG:SUBJ

‘He had been seen/had seen himself.’

Compound stem (TR.+TR.)

mak'-əgu-led-a-e

cut-bring-PLUP:A-FIN-3SG:SUBJ

'He had cut and brought.'

Past Progressive:

Transitive root

dal-et'-tahēkan-a-e

beat-IMPRF:A-COP:PST-FIN-3SG:SUBJ

'He was beating.'

Permissive with Applicative

men-ad-e-tahēkan-a-m

say-APPL.PST:A-3SG:OBJ-COP:PST-

FIN-2SG:SUBJ

'You were telling him.'

Causative stem

dal-oco-ed-e-tahēkan-a-m

beat-CAUS-PROG:A-3SG:OBJ-COP:PST-

FIN-3SG:SUBJ

'You were causing him to beat.'

Benefactive stem

əl-ka-e-kan-tahēkan-a-e

bind-BEN-3SG:OBJ-COP-COP:PST-

FIN-3SG:SUBJ

'He was binding it (cow) for somebody.'

Compound stem (TR.+TR.)

mak'əgu-et'-tahēkan-a-e

cut bring-IMPRF:A-COP:PST-FIN-3SG:SUBJ

'He was cutting and bringing.'

Anterior pluperfect:

Transitive root

idi-let'-ko-tahēkan-a-e

take away-PLUP:A-3PL:SUBJ-COP:PST-

FIN-3SG:SUBJ

'He had taken them them away.'

With Applicative

ləi-ad-ij-ṅ-tahēkan-a-e

tell-APPL.PST:A-1SG:OBJ-COP:PST-

FIN-3SG:SUBJ

'He had told me.'

Reciprocal stem

dapal-len-a-kin

beat each other-PLUP:M-FIN-3DL:SUBJ

'They had fought each other.'

Compound stem (INTR.+INTR.)

ārgo-hec'-len-a-e

descend-come-PLUP:M-FIN-3SG:SUBJ

'He had descended and come.'

Intransitive root

sen-ok'-kan-tahēkan-a-e

go-M-COP-COP:PST-FIN-3SG:SUBJ

'He was going.'

Permissive with Applicative

ṅel-an-tahēkan-a-e

see-APPL.PST:M-COP:PST-FIN-3SG:SUBJ

'He was seeing for himself.'

Causative stem

ṅel-oco-k'-kan-tahēkan-a-e

see-CAUS-M-COP-COP:PST-FIN-3SG:SUBJ

'He was being caused to be seen.'

Passive/Reflexive stem

ṅel-ək'-kan-tahēkan-a-e

see-M-COP-COP:PST-FIN-3SG:SUBJ

'He was being seen/was seeing himself.'

Compound stem (INTR.+INTR.)

ruəṅhij-uk'-kan-tahēkan-a-e

return come-M-COP-COP: :PST-FIN-

3SG:SUBJ

'He was returning.'

Intransitive root

hec'-len-tahēkan-a-ṅ

come-PLUP:M-COP:PST-FIN-1SG:SUBJ

'I had come.'

With Applicative

dal-an-tahēkan-a-e

beat-APPL.PST:M-COP:PST-FIN-

3SG:SUBJ

'He had beaten for himself.'

Causative stem

dal-oco-let'-kin-tahēkan-a-e
 beat-CAUS-PLUP:A-3DL:OBJ-COP:PST-
 FIN-3SG:SUBJ
 'He had caused them to beat.'

Permissive stem

nel-oco-ad-e-tahēkan-a-ko
 see-CAUS-APPL.PST:A-3SG:OBJ-COP:PST-
 FIN-3PL:SUBJ
 'They had permitted him to see.'

Benefactive stem

dapal-kad-e-tahēkan-a-e
 cover-BEN:PLUP:A-3SG:OBJ-COP-
 FIN-3SG:SUBJ
 'He had covered it (cow) for someone.'

Compound stem (TR.+TR.)

neljam-led-e-tahēkan-a-e
 see find-PLUP:A-3SG:OBJ-COP-FIN-
 3SG:SUBJ
 'He had seen and found him.'

Causative stem

dal-oco-len-tahēkan-a-e
 beat-CAUS-PLUP:M-COP:PST-FIN-
 3SG:SUBJ
 'He had been caused to beat.'

Reciprocal stem

dapal-len-tahēkan-a-kin
 beat<RECIP>-PLUP:M-COP:PST-FIN-3DL
 'They two had fought.'

Passive/Reflexive stem

nel-len-tahēkan-a-e
 see-PLUP:M-COP-FIN-3SG:SUBJ
 'He had been seen/seen himself.'

Compound stem (INTR.+INTR.)

āṛgohec'-len-tahēkan-a-e
 descend come-PLUP:M-COP-FIN-
 3SG:SUBJ
 'He had descended and come.'

3.2.5 Mood

The action or event in Santali can be divided as real (Indicative), desired (Optative), simply possible (Subjunctive), advisable (Imperative), and welcoming (Irrealis). The optative and subjunctive are subsumed under Optative, since the marker is the same for both and the meanings expressed by them are often fluid, and in some respects almost identical. There is a special marker for the Indicative – any finite clause with a finite *-a* is treated as Indicative. A verbal form with a suffixed second person *-me* functions as Imperative, and with particle *alo* as negative Imperative, the clitic pronoun for second person added to the particle. Irrealis is expressed by the suffix *-le*.

There is no special marker for conditional; the conditional particle *khan*, with or without a priorative suffix *-le* (active) and *-len* (middle), denotes conditionality.

(141) *janwar-ijn jam-le-ko-khan əgu-ko-a-ṅ*
 animal-1SG:SUBJ find-ANT-3PL:OBJ-COND bring-3PL:OBJ-FIN-1SG:SUBJ
 'If I find animals I shall bring them.'

(142) *ba-ṅ sen-len-khan uni ba-e hij-uk'-a*
 NEG-1SG:SUBJ go-ANT:M-PART 3SG NEG-3SG:SUBJ come-M-FIN
 'If I do not go he will not come.'

3.2.5.1 Optative

Markers for the Optative are *-ke* in the active and *-kok'* in the middle. The modal suffix for the active takes object after it.

3.2.6 Orientation/directionality

The concept of orientation or directionality does not play any morphological role in Santali. It is, however, achieved through compound verb constructions.

3.2.7 Voicelversion

Santali has two voices, active and middle, corresponding to semantical transitivity and intransitivity, respectively. Verbs appearing in the active are transitive and in the middle, intransitive. Some can appear in both the active and middle, and are transitive in the active and intransitive in the middle. Below are some of the verbs which appear in both the active and the middle:

<i>Verbs</i>	<i>Intransitive</i>	<i>Transitive</i>
<i>ap'</i>	alight	make alight
<i>rakap'</i>	rise, ascend	pull up
<i>paṛhao</i>	read	teach
<i>perēc'</i>	be full	fill
<i>caba</i>	end	finish
<i>rəput'</i>	be broken	break
<i>gitic'</i>	lie down	lay down
<i>busək'</i>	be born	give birth
<i>burum</i>	lie down	lay down(of animal)

Two sets of endings are added to the TAM markers to express active and middle. One set consists of one member, *-t'*, for the active, and the other consists of two members, *-ok'* and *-n*, for the middle. The ending *-ok'* is used in the simple present/future, present and past progressive, imperative, and optative while *-n* appears with the other TAM suffixes. There are three verbs – *topo* 'bathe', *teṅgo* 'stand', and *āṛgo*²⁵ 'descend' – which take the suffix *-n* directly with the stem in the simple present/future, present and past progressive, imperative, and optative in place of *-ok'*, as is normally followed in the middle.

- | | | |
|-------|--|--|
| (155) | <i>topo-n-a-e</i>
bathe-M-FIN-3SG:SUBJ
'He bathes/will bathe.' | <i>topo-n-kan-a-e</i>
bathe-M-COP-FIN-3SG:SUBJ
'He is bathing.' |
| (156) | <i>teṅgo-n-a-e</i>
stand-M-FIN-3SG:SUBJ
'He stands.' | <i>teṅgo-n-kan-a-e</i>
stand-M-COP-FIN-3SG:SUBJ
'He is standing.' |
| (157) | <i>āṛgo-n-a-e</i>
descend-M-FIN-3SG:SUBJ
'He descends.' | <i>āṛgo-n-kan-a-e</i>
descend-M-COP-FIN-3SG:SUBJ
'He is descending.' |

It should be noted that semantically transitive verbs take middle endings when conjugated in the reflexive, passive, and mediopassive. Reciprocal stems are also conjugated in the middle. Causative stems when conjugated in the passive also

take middle endings. Examples given in Section 3.2.4 illustrate voice endings in the Passive, Reflexive, Mediopassive, Reciprocal, and Passive Conjugation of the Causative.

3.2.8 Finiteness

A finite verb in Santali gets the finite ending *-a*. It gives the sentence a demonstrated reality. This is attached to all kinds of finite verbs except the Imperative.²⁶ In the subordinate clause it is never attached.

- (158) *uni ato-re cal-ak'-kan-a-e* (159) *mit' ghəric' taken-me*
 3SG village-LOC go-M-COP-FIN-3SG:SUBJ one time wait-2SG:IMP
 'He is going to the village.' 'Wait for a while.'
- (160) *ba-m hɔhɔ-ɨn-khan jur-k-ok'-a-ɨ*
 NEG-2SG:SUBJ call-1SG:SUBJ-COND fall-OPT-M-FIN-1SG:SUBJ
 'If you did not call me I might have fallen in the ditch.'

Non-finite forms used in various types of subordination are discussed in the section on Syntax. Non-finite refers to converbs in *-katε* and *-tε* which never occur with any TAM suffix, but they occur with stem formatives like *-oco* (causative) and *-jɔŋ* (mediopassive).

3.2.9 Negation

Negation is expressed by three particles in Santali: *baŋ* in interrogative and declarative sentences, *ɔhɔ* as emphatic negative in declarative sentences, and *alo* as prohibitive negative in the imperative. When the sentence is negated the subject clitic is attached to the particle instead of the verb. The negative particle obligatorily precedes the verb it negates. It cannot take any of the TAM suffixes. Consider the following pairs:

- (161) *em-ad-ɨn-a-e* *ba-e* *em-ad-ɨn-a*
 give-APPL.PST:A-1SG:OBJ-FIN-3SG:SUBJ NEG-3SG:SUBJ give-APPL.PST:A-1SG:OBJ-FIN
 'He gave me.' 'He did not give me.'
- (162) *cal-ak'-a-ɨ* *ɔhɔ-ɨ* *cal-ak'-a*
 go-M-FIN-1SG:SUBJ NEG-1SG:SUBJ go-M-FIN
 'I go/shall go.' 'I do not/shall not go.'
- (163) *em-a-e-mε* *alo-m* *em-a-e-a*
 give-A-3SG:OBJ-2SG:IMP PROHIB-2SG:SUBJ give-APPL-3SG:OBJ-FIN
 'Give him.' 'Do not give him.'
- (164) *sen-ok'-mε* *alo-m* *sen-ɔk'-a*
 go-M-2SG:IMP PROHIB-2SG:SUBJ go-M-FIN
 'Go' 'Do not go.'

baŋ is realized as *ba* when pronominal clitics for the subject are added to it. Sometimes *baŋ* is used with pronominal clitics too, but it is purely optional and has nothing to do with strength or force, as noted by Neukom (2001:149).

The particle *baŋ* is unmarked and the most frequently used negative. In contrast to what has been claimed by Neukom, that 'verb forms marked for the Past (*-ket'/-en*)

and for the Optative (-*ke*) cannot be negated; in negated clauses the Pluperfect (-*let’/len*) and the Irrealis (-*lel-len*) are used instead’ (Neukom 2001:149), Past and Optative verb forms are negated also with *baŋ*.

- (165) *uni ba-e men-ked-a* (166) *ba-e ruəphec’-en-a*
 3SG NEG-3SG:SUBJ say-PST:A-FIN NEG-3SG:SUBJ return.COME-PST:M-FIN
 ‘He did not say.’ ‘He did not come back.’
- (167) *ba-e dal-ke-a* (168) *ba-m sen-kok’-a*
 NEG-3SG:SUBJ beat-OPT-FIN NEG-2SG:SUBJ go-OPT:M-FIN
 ‘He should not beat.’ ‘You should not go.’

In (169) the use of *əhɔ* rather than *baŋ* does not mean that the Negative does not take PST TAM in the verb, but it can be interpreted as Irrealis:

- (169) *adɔ uni hɔɔ koɔa dɔ-e men-ked-a əhɔ-ŋ bɔɔɔ-len-a*
 then that man boy FOC-3SG:SUBJ say-PST:A-FIN NEG-1SG:SUBJ enter-IRR:M-FIN
 ‘Then that boy told, “I would never enter”.’ [P.O.Bodding]

The particle *əhɔ* is an emphatic negative used with the Optative, Declarative verb forms, and with Irrealis in the apodosis of a conditional sentence.

- (170) *ɪŋ dɔ əhɔ-ŋ ləi-ke-a*
 1SG:SUBJ FOC NEG-1SG:SUBJ tell-OPT:A-FIN
 ‘I might not say.’
- (171) *onko-ko ləi-ed-a aɛ ləi benao-dare-a-e-a*
 3PL:SUBJ-3PL:SUBJ tell-IMPFR:A-FIN 1PL:SUBJ tell make-able.to-APPL-3SG:OBJ-FIN
tɔbe jivi dɔ əhɔ-lɛ em-ke-a
 but life FOC NEG-1PL:SUBJ give-OPT:A-FIN
 ‘They are saying, “we can make him by uttering (mystical formula) but we are unable to give him life”.’
- (172) *ba-m dal-li-ŋ-khan əhɔ-ŋ dal-le-m-a*
 NEG-2SG:SUBJ beat-IRR:A-1SG:OBJ-if NEG-1SG:SUBJ beat-IRR:A-2SG:OBJ-FIN
 ‘If you did not beat me I would certainly not beat you.’

In (172), two negative particles have been used in the same sentence – *baŋ* in the protasis and *əhɔ* in the apodosis – and both have taken Irrealis as they are conditional sentences. This does not indicate that *baŋ* cannot be used in the simple past.

The particle *alo* indicates prohibition and occurs with the simple present/future verb forms to convey a negative imperative sense.

- (173) *handɛ alo-m cal-ak’-a*
 yonder PROHIB-2SG:SUBJ go-M-FIN
 ‘Do not go there.’
- (174) *setoŋ-re alo-m dɔɔbaɔay-a*
 sun-LOC PROHIB-2SG:SUBJ run.about-FIN
 ‘Do not run in the sun.’

There is only one instance in my data where *alo* is used in the final clause without prohibitive sense:

khub sɔktɔkate tɔl-me jemɔn alo laɣ-a
 very hard-CV tie -2SG:SUBJ so.that NEG shake-FIN
 ‘Tie it very hard so that it does not shake.’

Both *baŋ* and *alo* can be used predicatively with the meaning ‘not to be’.

(175) *ɪn alo-khan huɣeʻ khəʃi-ge-e jɔm-ke-pe-ge-a*
 1SG:SUBJ PROHIB-if if surely-FOC-3SG:SUBJ eat-OPT:A-2PL:OBJ-FOC-FIN
 ‘If I had not been, he would surely have eaten you.’

(176) *hij-uk’-me alo-m baŋ-a* ‘Come, don’t say no.’

3.2.10 Derivation

There are four processes of verbal derivation in Santali: prefixation, infixation, suffixation, and reduplication.

3.2.10.1 Causative *a-/oco*

Causative stems are derived from simple verb roots by means of the causative morphemes *a-* and *-oco*. Of these, *-oco* is employed to all types of roots, namely transitive, intransitive, transitive-intransitive, and reciprocal stems. The prefix *a-* is restricted to two transitive roots only, *jɔm* ‘eat’ and *ju* ‘drink’, which can take the suffix also, but never two at a time.

Permissive is also derived by the same suffix *-oco* as of Causative but the difference lies in the status of the person acted upon. In the Causative the person acted upon is in the accusative, while in the Permissive the person acted upon is in the dative (with the Applicative marker *a-*).

The suffix for Causative/Permissive *-oco*²⁷ has a different phonological shape in the Southern dialect, while the Causative prefix is *a-* in both.

By prefixing *a-*:

jɔm ‘eat’ *ajɔ** ‘feed’; *ju* ‘drink’ *aju* ‘give to drink’

The roots also derive their Causative/Permissive by suffixing *-oco* as in *jɔmoco* ‘feed’ and *juoco* ‘make drink’. Derivation by prefixing *a-* to *jɔm* and *ju* is used in the Northern dialect and in the older generation of the Southern dialect. The younger generation of the Southern dialect prefers suffixation to prefixation.

By suffixation *-oco*:

dal ‘beat’ > *daloco* ‘cause/allow to beat’ *ɲel* ‘see’ > *ɲeloco* ‘cause/allow to see’
sen ‘go’ > *senoco* ‘cause/allow to go’ *hec* ‘come’ > *hec’oco* ‘cause/allow to come’
gitic ‘lie,lay’ > *gitic’oco* ‘cause/allow to lie/lay’
gɔc ‘die,kill’ > *gɔc’oco* ‘cause/allow to die.kill’
ɲepel ‘see each other, meet’ > *ɲepeloco* ‘cause/allow to see each other/meet’
ɲapam ‘get together’ > *ɲapamoco* ‘cause/allow to get together’

3.2.10.2 Benefactive *-ka*

The Benefactive²⁸ stem is formed by suffixing *-ka* to the transitive and transitive-intransitive roots. When added to the transitive-intransitive roots the transitive meaning is prominent. As the name implies, this category denotes that the action is carried out on behalf of someone other than the subject: *tɔl* ‘bind’ > *tɔlka* ‘bind for somebody’; *dapal*, ‘cover’ > *dapalka* ‘cover for somebody’.

The Benefactive stem suffix is *-kak’* in the Southern dialect when the object is inanimate or there is no object at all. If the object is animate, the final *k’* is replaced by the pronominal clitics. In the Northern dialect it is invariably *-ka*. Conjugation of the Benefactive stem in Northern and Southern is thus:

Northern	Southern
<i>dapra tɔl-ka-e-me</i>	<i>dapra tɔl-ka-e-me</i>
bullock bind-BEN-3SG:OBJ-2SG:IMP	bullock bind-BEN-3SG:OBJ-2SG:IMP
‘Bind the bullock.’	‘Bind the bullock.’
<i>hoɣo dal-ka-m</i>	<i>hoɣo dal-ka-k’-me</i>
paddy thrash-BEN-2SG:IMP	paddy thrash-BEN-INAN:OBJ-2SG:IMP
‘Cover the paddy.’	‘Cover the paddy.’

The Benefactive stem is conjugated in the active with all TAM suffixes.

- (177) *duɣup’-kad-e-a-ɲ*
sit-BEN.PST:A-3SG:OBJ-FIN-1SG:SUBJ
‘I made her sit (for somebody).’
- (178) *gidrə-e əbuk-kad-e-a*
child-3SG:SUBJ wash-BEN.PST:A-3SG:OBJ-FIN
‘She washed the baby (on behalf of somebody).’

3.2.10.3 Reciprocal <*pV*>

The Reciprocal²⁹ stem is formed by infixing <*pV*> after the first vowel of the root, the vowel of the infix being the same root vowel after which the insertion takes place. The insertion of the infix is restricted to the transitive and transitive-intransitive roots. In the latter case the transitive meaning is prominent. It is always conjugated in the middle and occurs with any of the TAM suffixes.

dal ‘beat’ > *dapal* ‘beat each other’ *ɔr* ‘draw,pull’ > *ɔpɔr* ‘draw/pull each other’
landa ‘laugh’ > *lapanda* ‘laugh together’ *galmarao* ‘chat’ > *gapalmarao* ‘gossip’

The infix generally has a reciprocal interpretation, but with many verbs it also denotes that the two participants of an action did something together, as *lapanda* and *gapalmarao* indicate.

- (179) *dapal-kan-a-kin* (180) *onko-ko gapalmarao-kan-a*
beat each other-COP-FIN-3DL:SUBJ 3PL:SUBJ-3PL:SUBJ gossip-COP-FIN
‘They (two) are fighting.’ ‘They are gossiping.’
- (181) *alay apa hɔn-lay tapam-en-a adɔ cedak’ mɔɣɛ hɔɣ-then*
1DL father son-1DL:SUBJ fight-PST:M-FIN then why five man-with

dɔ-m nəlɪs-ket'-a
 TOP-2SG:SUBJ complain-PST:A-FIN
 'We (two), father and son, fought (lit. seized each other), then why did you complain to the village council.'

3.2.10.4 Medio-passive *-jɔŋ*

The Medio-passive is formed by suffixing *-jɔŋ* to the transitive or a limited number of intransitive roots and transitive-intransitive roots, the transitive meaning being prominent in the latter case. The stem is conjugated in the Simple present/future, present and past progressive, and Imperative. In other tense/aspects, *-jɔŋ* is dispensed with in favour of the applicative *a-* and the middle ending *-n*. It is conjugated in the middle.

- | | | | |
|-------|---|-------|--|
| (182) | <i>dal-jɔŋ-a-e</i>
beat-MP-FIN-3SG:SUBJ
'He beats for himself.' | (183) | <i>dal-jɔŋ-kan-a-e</i>
beat-MP-COP-FIN-3SG:SUBJ
'He is beating for himself.' |
| (184) | <i>dal-jɔŋ-mɛ</i>
beat-MP-2SG:IMP
'Beat for yourself.' | (185) | <i>dal-an-a-e</i>
beat-APPL.PST:M-FIN-3SG:SUBJ
'He beat for himself.' |

3.2.10.5 Passive/reflexive *-ok'*

The suffix for the Passive/Reflexive stem is *-ok'*. It is added to the transitive roots, transitive-intransitive roots, and causative stems to derive passive stems. When added to the transitive-intransitive roots, the transitive meaning is prominent. In conjunction with the transitive roots it denotes reflexivity, that is, the subject's action affects the subject itself.

Passive from transitive and transitive-intransitive roots:

ɲam 'get' > *ɲamok'* 'be got'; *ɲel* 'see' > *ɲelok'* 'be seen'; *aŋjɔm* 'hear' > *aŋjɔmok'* 'be heard'; *gɔc'* 'kill' > *gojok'/gujuk'* 'be killed'

Passive from causative stems:

ranoco 'cause to medicate' > *ranocok'* 'be caused to medicate'
jɔmoco 'feed' > *jɔmocok'* 'be fed'

Reflexive from transitive roots:

ɔl 'write' > *ɔlɔk'* 'write oneself'; *ɛm* 'give' > *ɛmɔk'* 'give oneself'; *mak'* 'cut' > *magok'* 'cut oneself'; *ir* 'reap' > *irok'* 'reap oneself'

The stem formative appears in simple present/future, present and past progressive, and Imperative. In other tenses/aspects the simple root is conjugated in the middle to denote Passivity/Reflexivity.

- | | | | |
|-------|--|-------|--|
| (186) | <i>dal-ok'-a-e</i>
beat-PASS/RFLXV-FIN-3SG:SUBJ
'He is beaten/beats himself.' | (187) | <i>dal-ok'-kan-a-e</i>
beat-PASS/RFLXV-COP-FIN-3SG:SUBJ
'He is being beaten/is beating himself.' |
| (188) | <i>oɾak' hɔy dak'-te</i>
house storm rain-INS
<i>rəput'-en-a</i>
break-PST:M-FIN
'The house was damaged by the rainstorm.' | (189) | <i>mak'-en-a-e</i>
cut-PST:M-FIN-3SG:SUBJ
'He cut himself.' |

3.2.10.6 Iterative/intensive: reduplication and infixation

The Iterative/Intensive stem is formed by reduplication of the initial consonant along with the vowel, or reduplication of the initial vowel in the consonant and vowel-initial roots, respectively. Bodding (1929:168) and Neukom (2001:126) mentioned only one way of forming the Iterative/Intensive stem of the vowel-initial roots – by infixation of *-k'*. In my field data there are two-ways of this stem formation – one by reduplication of the initial consonant and the other by infixation of *-k'*, the latter not being very frequent. The older generation generally forms the stem from vowel-initial roots by infixation. The formation is restricted to transitive roots and two intransitive roots only. Both stems are formed in the same way, and only the context determines whether Intensive or Iterative meaning is involved.

<i>get</i> 'cut' > <i>geget</i>	<i>dəɽ</i> 'run' > <i>dədəɽ</i>
<i>ɲel</i> 'see' > <i>ɲɛɲel</i>	<i>ɟək</i> 'sweep' > <i>ɟəɟək</i>
<i>mak</i> 'cut' > <i>mamak</i>	<i>rək</i> 'pierce' > <i>rərək</i>
<i>dak</i> 'rain' > <i>dadak</i>	<i>tɟɪ</i> 'pierce with arrow' > <i>tutɟɪ</i>
<i>cas</i> 'cultivate' > <i>cacas</i>	<i>tot</i> 'peck' > <i>totot</i>
<i>rɔɽ</i> 'speak' > <i>rɔrɔɽ</i>	<i>ger</i> 'bite' > <i>geger</i>
<i>təl</i> 'bind' > <i>tətəl</i>	<i>ɲam</i> 'get' > <i>ɲaɲam</i>

The roots of the CV structure are found to reduplicate the whole roots.

ɲu 'drink' > *ɲuɲu* *si* 'plough' > *sisi*

The roots of the VC structure reduplicate the initial vowel or infix *-k'* after the vowel.

ɔr 'draw,pull' > *ɔɔr/ɔk'ɔr*; *ir* 'reap' > *iir/lik'ir*;
er 'sow' > *ɛɛr/lek'er*; *em* 'give' > *ɛem/lek'em*; *ɔl* 'write' > *ɔɔll/ɔk'ɔl*

Only two intransitive roots are found undergoing reduplication to form Iterative/Intensive stems, iterative meaning being more prominent.

sen 'go' > *sesen* *ɲir* 'run' > *ɲiɲir*

The stems formed from the dissyllabic roots tend to become trisyllabic, which is not tolerated by the language. Therefore, the trisyllabic stems become dissyllabic by elision of the second vowel. Thus,

benao 'make' > *bebnao* (<*bebenao*) *banam* 'play on violin' > *babnam* (<*babanam*)
hatao 'take' > *hahtao* (<*hahatao*) *kuli* 'ask' > *kukli* (<*kukuli*)

Roots with diphthongs reduplicate the first vowel of the diphthong along with the initial C.

ləi 'tell' > *lələi* 'reiterate'

In Neukom's data, dissyllabic roots with initial consonants form an Intensive stem by the infixation of *-k'*. This type of formation is not attested to in my data.

The Iterative/Intensive takes only *kan* and *tahēkan* and does not allow any of the TAM suffixes and pronominal object clitics.

3.2.10.7 Compound stem derivation

Compound stems are formed by the root and one more element, either a root or a modifier. According to the semantic content of the last element, two types may be distinguished:

- (i) Root + Root
- (ii) Root + Modifier

In the first category two transitives or two intransitives may be juxtaposed.

Transitive + Transitive

sendra ‘hunt’ + *jam* ‘find’ = *sendrajam* ‘search and find’

mak ‘cut’ + *agu* ‘bring’ = *mak’agu* ‘cut and bring’

- (190) *uni kul-e sendra-jam-ked-e-a*
 3SG tiger-3SG:SUBJ search-find-PST:A-3SG:OBJ-FIN
 ‘He searched for and found the tiger.’

Intransitive + Intransitive

rakap ‘ascend’ + *hec* ‘come’ = *rakap’hec* ‘ascend and come’

ãrgo ‘descend’ + *hec* ‘come’ = *ãrgohec* ‘descend and come’

- (191) *subodh rakap’hec-en-a-e*
 Subodh ascend come-PST:M-3SG:SUBJ
 ‘Subodh ascended and came.’

Transitive and Intransitive as also Intransitive and transitive may be juxtaposed, but their number is very limited.

Intransitive + Transitive

sen ‘go’ + *daɽe*³⁰ ‘be able’ = *sendaɽe* ‘able to go’

Transitive + Intransitive

jel ‘see’ + *janam* ‘be born’ = *jeljanam* ‘see born’

- (192) *jeljanam-ked-e-a-n*
 see be born-PST:A-3SG:OBJ-FIN-1SG:SUBJ
 ‘I saw him being born.’

In the second roots and modifiers are juxtaposed to derive a compound stem.

Root + Adverbial modifiers

Some modifiers occur after the root to give a twist to the basic meaning of the root.

baɽa,gət’,ofo,bɽɛc’,dorok’ are some of the most commonly used modifiers. Neukom (2001:142) mentions two more modifiers – *hataɽ* and *gela*.

3.2.11 Noun incorporation and combining forms

Santali does not allow noun incorporation in the predicate, as is found in the South Munda languages in the form of full incorporation or combining forms.

3.2.12 Auxiliary verb constructions

Two of the commonly used auxiliaries are *daʒe* ‘can’ and *lega* ‘try to’, described as V2 of the compound stem by Neukom (2001:137). The auxiliaries cited specify the modality of the first root. *daʒe* takes the Applicative *a* and is conjugated in the active, even if there is no object in the surface level. *lega* is conjugated in the middle.

- (193) *uniændĩḁ dṵ kanti sudhə rəhəʒ-en-ta-e-tə ba-e*
 that male TOP throat also dry-PST:M-POSS-3SG:POSS-CV NEG-3SG:SUBJ
rəʒ-daʒe-at'-a
 speak-can-APPL.PST:A-FIN
 ‘Since even the throat of the jackal had become dry, he could not speak.’
- (194) *sereŋ-legak'-me jut-ok'-re hṵ baŋ-re hṵ*
 sing-try-M-2SG:IMP succeed-M-LOC too NEG-LOC also
 ‘Try to sing whether you will succeed or not.’

3.3 Expressives

Santali has a construction popularly termed in South Asian languages as echo-word formation. It is constructed in more than one way – by repeating the element in an identical form, by augmenting a consonant in the repeated element, and by vowel mutation. Sometimes repetition and vowel mutation occur simultaneously. The repeated form does not have any independent meaning but modifies the meaning of the first element. Different types of expressive formations are as follows:

3.3.1 Identical reduplication

The expressives under this category are formed by repeating the first element.

<i>ahal ahal</i>	‘distressed’	<i>ajak'ajak'</i>	‘clamour for’
<i>atrəm atrəm</i>	‘incompletely’	<i>baḍgak' baḍgak'</i>	‘sharp painful sensation’
<i>baḍgət' baḍgət'</i>	‘rough’	<i>cəḍuk' cəḍuk'</i>	‘noise of pumping into water’
<i>dacaŋ dacaŋ</i>	‘ubiquitous’	<i>gab gab</i>	‘sink deeply’
<i>halat' halat'</i>	‘slightly’	<i>jelep' jelep'</i>	‘flashing’
<i>kāc' kāc'</i>	‘whine as a dog’	<i>məkur məkur</i>	‘sound of crunching’

3.3.2 Partial reduplication

This type of reduplication is formed by augmenting a consonant initially in the repeated element. The augmented consonant may be any of *j, t, p, b, g, d, k, c, s, m, ph, r* and *dh*.

- (i) ØVX CVX
- | | | | |
|--------------------|------------------|-------------------|-----------------------|
| <i>əbuk'cəbuk'</i> | ‘here and there’ | <i>abe tabe</i> | ‘just at the time of’ |
| <i>acel pacel</i> | ‘abundance’ | <i>əḍəi bəḍəi</i> | ‘arrogant’ |
| <i>adha padha</i> | ‘unfinished’ | <i>ahal kahal</i> | ‘distressed’ |

alam galam ‘indistinctly’ *albaṭ salbaṭ* ‘contradictory’
ana gona ‘coming and going’ *anak phanak* ‘sundries’
əɽuk’ dhəɽuk’ ‘here and there’ *aṭap’ jaṭap’* ‘ parched with thirst’
akjak rakjak ‘lightly’

(ii) ØVIX CV2X (with vowel mutation)

adha sudhə ‘half’; *agaɽbigəɽ* ‘topsy turvy’; *agar dɨgəɽ* ‘infringe’; *əhir kuhir* ‘fix the eyes upon’; *əiṭhəṭiṭhə* ‘leavings of food’; *ajak’ bujak’* ‘irregular’; *əril kuril* ‘stare as smoke nips the eyes’; *aral koral* ‘perplexed’; *asaha dusəha* ‘evil omen’; *əṭis kuṭis* ‘innumerable’

(iii) CVIX CV2X (with vowel mutation)

bacak’ bocok’ ‘nonsensical’; *badha bidhi* ‘occult adverse influence’; *baḍak’ buḍuk’* ‘move the lips as if speaking’; *bakat’ bəkət’* ‘chatter’; *bhaɽbhuɽ* ‘crashing noise’; *caṭa cuṭu* ‘crackle’; *gasac’ gusuc’* ‘solitary’

(iv) C1VX C2VX

bajek sajek ‘sometimes’; *cas bas* ‘cultivation’; *cedro bedro* ‘rough’; *cehəɽ behəɽ* ‘warmly’; *celkə melkə* ‘wanton’; *ceɽə beɽə* ‘twitter’; *cuṅgur muṅgur* ‘be restless’

(v) C1V1X C2V2X

bacha kuchə, ‘refuse’ *kāc’kūc’* ‘stingy’

(vi) VICV V2CV

aḍe oḍe, ‘secretly’

(vii) VICV1C V2CV2C (vowel mutation)

adac’ uduc’ ‘unwieldly through corpulence’ *adaɽodoɽ* ‘fat and naked’
agar ogor ‘dumpy’ *amaɽ omoɽ* ‘diligently’
aɽae oɽoe ‘burning’ *arak’ orok’* ‘stare vacantly’
aral orol ‘perplexed’ *asam usum*, ‘leisurely’

(viii) VICV1 VICV2 (V1 is invariably *a* and V2 is *i*)

āṭa āṭi ‘dispute’ *adha ədhi* ‘half’
adra ədri ‘be ill-humoured’ *aḍra əḍri* ‘bellow’
əḡlə əḡli ‘be forward’ *agra əgri* ‘show temper’
ahka ahki ‘panting’ *ala əli* ‘tired’
andka əndki ‘a strong smell’ *ankha ənkhi* ‘disgusting’
apna əpni ‘automatically’ *aɽsa əɽsi* ‘plead an excuse’

Sometimes the initial or the medial consonant of the first form tends to change in the repeated form.

kadar kapar ‘rubbish’ *hadrak’ gasrak’* ‘stumblingly’

4 SYNTAX

4.1 Syntax of the simple sentence

Santali has the unmarked word order SOV, although the order can be twisted depending on the topic of discussion. Topical elements tend to appear first in the

sentence followed by less or non-topical elements. The unmarked word order may be shown by the following sentence (195):

- (195) *am dajra-m əgu-e-a*
 2SG bullock -2SG:SUBJ bring-3SG:OBJ-FIN
 ‘You will bring the bullock.’

Here *am*, being the unmarked subject, appears first followed by the object *dajra*, and the verb comes at the end. As the following sentence shows, the object can begin a sentence if it is the topic of discussion:

- (196) *kimin dɔ ba-m əgu-ko-a?*
 daughter-in-law TOP NEG-2SG:SUBJ bring-3PL:OBJ-FIN
 ‘Will you not bring daughter-in-law?’

As *kimin* is the topic of discussion, it appears in the initial slot without any apparent mention of the subject which is, however, marked in the word preceding the verb. There is one apparent contradiction in the object marking – the object is marked in the verb by the third personal plural clitic *ko* although *kimin* is in the singular form. It indicates that the number of objects talked about are indefinite, and the number suffix may not be overtly marked in the object. It may be compared with Bangla *bou tou* ‘wife and the like’ when a singular relation is mentioned.

The subject and object both may not be overtly marked in the sentence; their presence in the sentence in the clitic pronominal form is obligatory.

- (197) *ona khunʃi-re-ge-ŋ tɔl-led-e-a, tin-re cɔ*
 that pole-LOC-EMPH-1SG:SUBJ bind-PLUP-3SG:OBJ-FIN how-LOC ever
topak'-kate dəʃ-gɔt'-ked-a-e
 break-CV run-away-PST:A-FIN-3SG:SUBJ
 ‘I tied it (goat) to that very pole, however it ran away by breaking it.’

In this sentence neither subject nor object is overtly marked; they are incorporated in the verb or the word preceding it, which is obligatory. The locative phrase appears first in the sentence with focus marker *ge*, which is the topic of concern.

In a sentence having an object and a dative (indirect object), the unmarked order is subject–dative–object–verb.

- (198) *in uni noa katha-ŋ met-ad-e-a*
 1SG:SUBJ 3SG this word-1SG:SUBJ tell-A.PST:A-3SG:OBJ-FIN
 ‘I told him this word.’

However, the dative *uni* or the dative and the object may be dispensed with if not considered topical.

- (199) *met-ad-e-a-ŋ*
 tell- APPL.PST:A-3SG:OBJ-FIN-1SG:SUBJ
 ‘I told him.’

Sentences like (199) preclude a context in reference to which the object or the dative is dropped. The following couplet illustrates the point:

- (200) Q. *pas-akad-a-m to, cəkri-ko dɔ ba-m ɲel-baʃa-ed-a?*
 pass-PRF:A-FIN-2SG:SUBJ DISC job-PL TOP NEG-2SG:SUBJ see-INDEF-IMPRF:A-FIN
 ‘You have passed, are you not searching for job?’

A. *paŋja-baŋa-ed-a-ŋ*
 search-INDEF-IMPRF:A-FIN-1SG:SUBJ
 ‘I am searching.’

The position after the verb can be used for positing the converb or as a determining afterthought.

- (201) *ceŋ' hoy-uk'-a kuŋi gidrə ɔl-ək' sēŋa-kate*
 what be-M-FIN girl child write-M educate-CV
 ‘What will happen by educating a girl?’
- (202) *cal-ak' cal-ak'-te mit'-ŋaŋ gaŋa-ge hɔŋ-rɛ-ko*
 go-M go-M-CV one-CLSSFR river-FOC way-LOC-3PL:SUBJ
ŋam-ked-a pɛrɛc'-akan
 find-PST:A-FIN be.full-PRF:M
 ‘While going along the way they found a river full (of water).’

In the phrase level, the constituent order adjective, numeral, demonstrative, and quantifier precede the head noun. The structure of the noun phrase is as follows:

(DEM) (QUANT) (ADJ) (ADJ) NOUN

<i>nui hɔŋ</i> ‘this man’	<i>noa serma</i> ‘this year’
<i>nui daŋan hɔŋ</i> ‘this strong man’	<i>noa usul dən</i> ‘this high jump’
<i>nui ədʒi napay kuŋi</i> ‘this very beautiful girl’	<i>noa ədʒi maraŋ bir</i> ‘this very big forest’
<i>mit' hɔŋ</i> ‘one man’	<i>mit' serma</i> ‘one year’
	<i>bele jɔ</i> ‘ripe fruit’
	<i>bele</i> ³¹ <i>bele jɔ</i> ripe fruits’

Identificational sentences marked by *kan* and *tahēkan* are formed nowadays without the copulas mentioned. This happens in most of the South Asian languages. A story begins with the following sentence:

ato ŋutum baŋɛdʒi ‘Name of the village Baredi.’ Instead of *ato ŋutum baŋɛdʒi-kan-a*.

In the first paragraph of the story ‘sikhnət’³² there are six sentences, of which only two sentences are marked by a finite *-a*; the other four are without a finite marker or copula.

rasi ato ‘Big village.’
 big village

- (203) *baŋ baŋ-te mit' sae-khən hɔ bəŋti gan oŋak' dɔ mena-k'-a*
 NEG NEG-CV one hundred-ABL even.more about house FOC have -M-FIN
 ‘Never-the-less more than hundred houses are there in the village.’
- (204) *rɛŋgec' kisəŋ sanam lekan hɔŋ-ge mena-k'-ko-a*
 poor rich all like man-FOC have-M-3PL:SUBJ-FIN
 ‘All types of people poor and rich are there.’

- (205) *tɔbe kisãŋ-mente dɔ bar-ea kaɖa nahel- dhəbic'-ren*
 yet rich-said to be TOP two-CLSSFR buffalo plough-up.to-GEN
cəsi hɔŋ
 farmer man
 'Yet he is said to be rich who has two buffalo-ploughs.'
- (206) *hɔŋ ato-ko-re dɔ onko-ge andel kisãŋ*
 Santal village-pl-LOC FOC 3PL:SUBJ-FOC very rich
 'They are the rich in Santal villages.'

4.1.1 Agreement

Agreement reflects natural number. Nouns like 'crowd', 'board', or 'government' agree with a verb with 3rd-PL clitic, although these nouns may not be overtly marked for number.

Consider the following sentence:

- (207) *nui boɖ dɔ ɔkɔe hɔ cəkri ba-e em-at'-ko-a*
 this board FOC who even job NEG-3SG:SUBJ give-APPL.PST:A-3PL:OBJ-FIN
 'This board has given job to none.'

Here *boɖ* is unmarked for number, but its singularity as a body is marked in the predicate, which is incorporated in the negative particle preceding the verb in the form of the third person singular clitic *-e*. The recipient being indefinite is marked by plural *-ko*.

- (208) *nui sɔrkar dɔ ba-e bhage-a*
 this government TOP NEG-3SG:SUBJ good-FIN
 'This government is not good.'

'Government' being considered as a single entity agrees with the verb in the form of the third person singular clitic *-e*. 'Cattle herd' is considered plural and is marked by the plural clitic *-ko* in the verb.

- (209) *uni-ren dɔ pal pal gəi mena-k'-ko-a*
 3S-GEN TOP herd REDPL COW have-M-3PL:OBJ-FIN
 'They have herd of cattles.'

Indefinite reference also does not overtly mark the number in the noun, rather it is marked in the verb. Here I refer to the sentence in (192) in which the object *kimin* is not marked for number but is marked in the verb.

4.1.2 Verb serialization

Root serialization is quite common in Santali as we have seen in compound verb constructions.³³ Along with this verb, roots with TAM may be serialized. According to Neukom (2001:176) 'the two verbs either refer to distinct sub-events of the same situation, or they are (quasi)-synonyms and denote the same situation.' Consider the sentence quoted in Neukom (2001:176):

- (210) *bhəgtɛ-ko raɣa-led-e ɲam-led-e*
 quickly-3PL:SUBJ release-PLUP:A-3SG:OBJ find-PLUP:A-3SG:OBJ

uni tərup' dɔ-e rɔŋ-gɔt'-ked-a
 that tiger TOP-3SG:SUBJ speak-V2-PST:A-FIN

'No sooner had they let him out and found him than the leopard said.'

The pronominal object can appear twice, as in the above example, but the finite *-a* occurs only once with the final verb. This type of verb serialization is very rare, and in modern-day language it is not found, at least in my data.

4.2 Complex sentence structure

Santali has a number of conjunctions and disjunctions that coordinate clauses. Although they do not have any effect on word order, they play an important role in complex sentences. Some of the most common conjunctives and disjunctives are:

ar 'and' *menkhan* 'but' *se* 'or'
adɔ 'then, thereupon' *bickom* 'rather' (L) *khan/khac* 'then'
baŋkhan 'otherwise' *bɔrɔŋ* 'rather' *baŋma* 'that is to say, namely'

- (211) (i) *ac' baba subodh-e hɔhɔ-ad-e-a*
 3SG father Subodh-3SG:SUBJ call-APPL.PST:A-3SG:OBJ-FIN

ar-e met-ad-e-a
 and-3SG:SUBJ tell-APPL.PST:A-3SG:OBJ-FIN
 'His father called Subodh and told him.'

- (ii) *ba-bon jɛl-led-e-a-e baŋma uni*
 NEG -1PL:SUBJ see-PLUP:A-3SG:OBJ-FIN-3SG:SUBJ that is to say 3SG

hɔpɔn era ədʒi-e cɔrɔk-a
 child girl very-3SG:SUBJ nice-FIN
 'We did not see her, people say, the girl is very beautiful.'

4.2.1 Coordination

Coordinate clauses are formed by particles denoting conjunction, disjunction, adversative, and conclusive.

4.2.1.1 Conjunctive

ar, *adɔ* and *khan* are the coordinate conjunctives. While *ar* operates within the sentence, *adɔ* and *khan* operate across sentences.

- (212) (i) *ɪn jodi mantar-ɪn ləi-a ar dak'-ɪn*
 1SG:SUBJ if magic.chant-1SG:SUBJ utter-FIN and water-1SG:SUBJ

chiŋkəu-a noa jaŋ-ko joŋon-gɔd-ɔk'-a
 sprinkle-FIN this bone-PL join-V2-M-FIN

'If I utter mystical formula and sprinkle water these bones will be joined.'

- (ii) *adɔ topon somɔy uni buɖhi dɔ ona*
 thereupon bath time 3SG old lady TOP that

alo-e sap'-ked-e-a
 light-3SG:SUBJ catch-PST:A-3SG:OBJ-FIN

'Thereupon at the bathing time that old lady caught hold of that light.'

- (213) *khan dɔ koʒa gidrə rɔʔ-ruəʔ-ked-a* 'henda baba
 then TOP boy child speak-return-PST:A-FIN look father
ona gundli ma bele-jut-akan-a
 that millet MOD ripe-properly-PRF:M-FIN
 'Then the boy replied, "Look father that millet has properly ripened."
adɔ and *khan* occur in continuing discourse, that is, they preclude something that is spoken before, and the present sentences are uttered in reference to that.

4.2.1.2 Disjunctive

se 'or' and *baŋkhan* 'otherwise' are used to form coordinate disjunctive clauses.

- (214) *ɔŋjɔm-ked-a-m se-m baŋ-a*
 hear-PST:A-FIN-2SG:SUBJ OR-2SG:SUBJ NEG-FIN
 'Did you hear or not?'
- (215) *onka dɔ alo-m rɔʔ-a baŋkhan-laŋ dal-me-a*
 like that TOP PROHIB-2SG:SUBJ speak-FIN otherwise-1DL:SUBJ beat-2SG:OBJ-FIN
 'Do not speak thus otherwise we will beat thee.'

4.2.1.3 Adversative

menkhan, *bickom*, *bɔrɔŋ* and *hutkə* are used as coordinating conjuncts denoting the adversative.

- (216) *ɲel-led-e-a-ŋ menkhan bə-ŋ galmarao-ad-e-a*
 see-PLUP:A-3SG:OBJ-FIN-1SG:SUBJ but NEG-1SG:SUBJ discuss-APPL.PST:A-3SG:OBJ-FIN
 'I had seen him but did not discuss with him.'

menkhan is also used as a switch reference marker.

- (217) *unkin dɔ din-ge ədʒi kurumuʔu-kin kəmi-a menkhan*
 3DL TOP day-FOC very diligently-3DL:SUBJ work-FIN but
ceka-kate-e mit' din uni hɔʔ-rən oʔak' boŋga dɔ bɔhək'
 how-CV-3SG:SUBJ one day 3SG man-GEN house goddess TOP head
lac' haso ɲam-ked-e-a
 stomach pain get-PST:A-3SG:OBJ-FIN
 'They two work very hard, but one day for unknown reason the man's wife was affected by pain in stomach and head.'
- (218) *nitok' dɔ bickom bə-ŋ cal-ak'-a*
 now TOP rather NEG-1SG:SUBJ go-M-FIN
 'Rather I shall not go now.'
- (219) *hɔʔ dɔ dʒher dɔ ba-ko sen-len-a bɔrɔŋ deko*
 Santal TOP many TOP NEG-3PL:SUBJ go-PLUP:M-FIN rather Hindu
bəʔti-ge-ko tahēkan-a
 more-FOC-3PL:SUBJ COP:PST-FIN
 'Many Santals had not gone there rather the Hindus were more.'

- (220) *am-em* *kuli-n-khan* *hutkə-n* *em-ke-m-a*
 2SG:SUBJ-2SG:SUBJ ask-1SG:OBJ-if then-1SG:SUBJ give-OPT-2SG:OBJ-FIN
 ‘If thou hadst asked me I would have given it to you’

hutkə is used in conditional sentences to introduce the apodosis, in which the protasis is supposed not to have been realized, and therefore, the apodosis would not have occurred.

4.2.2 Complement clauses

Complement clauses perform a nominal function which may be subsumed as subject or object.

4.2.2.1 In subject function

Verbs in complement clauses in subject function may contain the middle voice marker:

- (221) *dinəm on-rɛ cal-ak’ dɔ baŋ boge-a*
 daily that-LOC go-M TOP NEG good-FIN
 ‘It is not good to go over there daily.’
- (222) *hamal jinis tul-baɾa dɔ baŋ ʈhik-ge-a*
 heavy thing lift-INDEF TOP NEG proper-FOC-FIN
 ‘It is not proper to lift heavy weight.’

4.2.2.2 In object function

Complement clauses in an object function are quite common, especially with predicates of speech. The verb form is without TAM suffixes and pronominal clitics. Intensive verb stems are sometimes used:

- (223) *uni iŋ-ʈhen kəkɔy-e hec’-en-a*
 3SG 1SG-place ask for -3SG:SUBJ come-PST:M-FIN
 ‘He came to me to ask for.’
- (224) *uni əgu baron-ko-m*
 3SG bring forbid-3PL:OBJ-2SG:IMP
 ‘Forbid them to bring him.’
- (225) *iŋ dɔ uni kombɔo-n ɲel-akad-e-a*
 1SG:SUBJ TOP 3SG steal-1SG:SUBJ see-PRF:A-3SG:OBJ-FIN
 ‘I saw him stealing.’
- (226) *ac’ baba subodh hɔhɔ-kate-e met-ad-e-a,*
 3SG father Subodh call-CV-3SG:SUBJ say-APPL.PST:A-3SG:OBJ-FIN
 ‘*teheŋ dɔ iŋ bir-te ɔhɔ-n calao-kok’-a.*
 today TOP 1SG:SUBJ forest-LOC NEG-1SG:SUBJ go-OPT-FIN
 ‘Calling Subodh his father told him, “Today I wont be able to go to the forest.”’

There are complement clauses marked for TAM, personal clitics as well as for the finite *-a*.

- (227) *gendak'* *oʃok'-kate-e* *ɲel-ked-e-a* *rugi*
 cloth remove-CV-3SG:SUBJ see-PST:A-3SG:OBJ-FIN patient
dɔ-e *gɔj-akan-a*
 TOP-3SG:SUBJ die-PRF:M-FIN
 'Having removed the cloth he saw the patient is dead.'
- (228) *kuɾi-ko dak' lo hij-uk'-e* *ɲel-ket'-ko-a*
 girl-PL water fetch come-M -3SG:SUBJ see-PST:A-3PL:OBJ-FIN
 'He saw the girls coming to fetch water.'

Complement clauses may be marked by the quotative *mente* (lit. by saying).³⁴

- (229) *onko uni ako apat mente bay-ko* *ɲel-ʃhəukə-e-kan-a*
 3PL:SUBJ 3SG 3PL father QUOT NEG-3PL:SUBJ see-fully-3SG:OBJ-COP-FIN
 'They do not recognize him as their father.'

Direct speech may be introduced by the quotative marker *bayma*.

- (230) *seday dɔ men-a bayma kayra sakam-leka jɔjɔ sakam*
 long ago TOP say-FIN QUOT banana leaf-like tamarind leaf
 'Long ago it was said tamarind leaf (is) like banana leaf.'

4.2.3 Subordinate clauses

Subordinate clauses in an oblique function are marked by converb *kate* and *te*, ablative *khən*, place marker *ʃhen*,³⁵ temporal *khan*, and purposive *jemon*.

The converb *kate* refers to an adverbial subordinate clause in the form of non-finite verb forms. It occurs with bare roots or with the middle voice marker and the reflexive marker *-jɔŋ*, but never with tense-aspect suffixes and the marker for finite.

- (231) *ac' baba subodh hɔhɔ-kate-e* *met-ad-e-a-e ...*
 3s father Subodh call-CV-3SG:SUBJ tell-APPL.PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'Calling Subodh his father told him ...'

Here I also refer to sentence (227) in which the subordinate clause with *kate* begins the sentence. In narrative, the converb in *kate* can be used as sentence linker, repeating the verb of the preceding clause.

- (232) *uni buɖhi dɔ ona sara-re-ko* *dɔhɔ-ked-e-a* *dɔhɔ-kate*
 that old lady TOP that pyre-LOC-3PL:SUBJ put-PST:A-3SG:OBJ-FIN put-CV
seŋgel-ko lagao-ad-e-a
 fire-3PL:SUBJ apply-APPL.PST:A-3SG:OBJ-FIN
 'They put the old lady on the funeral pyre; having put her on it they set her on fire.'
- (233) *ɲel-jɔŋ-kate-ɲ* *calao-kok'-a*
 see-MP-CV-1SG:SUBJ go-OPT-FIN
 'Seeing it I would go.'

The instrumental *-te* with a verb functions as a converb with simultaneous or sequential meaning. *-te* as a converb formative generally occurs with reduplicated

verb forms, Neukom's data (2001:188) explicate a non-reduplicated verb with *-te*. The verb-forms indicate time and manner of action.

- (234) *cal-ak' cal-ak'-te mit'-ṭaṅ toyo jɛl-ked-e-a-e*
 go-M go-M-CV one-CLSSFR jackel see-PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'While going he saw a jackal.'

- (235) *laha laha-te calao-en-a-e*
 precede REDPL-CV go-PST:M-FIN-3SG:SUBJ
 'He went in front.'

Subordinate clauses marked by the place marker *ṭhen* function as locative adverbial clauses.

- (236) *gapa dɔ am-ge si-ok'-ṭhen ḍaṅra dɔ laga-əgu-kin-mɛ*
 tomorrow TOP 2SG-FOC plough-M-PL bullock TOP drive-bring-3DL:OBJ-2SG:IMP
 'Tomorrow you will drive the bullocks to the place of ploughing.'

Verbs with *khɔn* in the subordinate clause indicate the time from which the action of the main clause holds. Verbs occurring with *khɔn* can take TAM and pronominal clitics.

- (237) *dare-khɔn-e jur-en-khɔn dɔ kəhil ge mena-e-a*
 tree-ABL-3SG:SUBJ fall-PST:M-ABL TOP ill FOC exist-3SG:OBJ-FIN
 'Since he fell down from the tree he became ill.'

Temporal clauses are formed by the suffix *-khan*. The temporal sense is manifested when the verb of the subordinate clause is marked for past. Temporal sense is also manifested when the verb is marked for non-past and irrealis.

- (238) *ona men-kate ti-e jɔṭet'-led-e-khan-ge uni*
 that say-CV hand-3SG:SUBJ touch-PLUP:A-3SG:OBJ-TEMP-FOC that
maejju dɔ-e men-ked-a, baṅma 'dohai ṭhakur.'
 woman TOP-3SG:SUBJ say-PST:A-FIN, QUOT help God
 'Saying that when he had touched her hand the woman said, "Oh, God, help."'

- (239) *jāhānak'-ko kuli-ben-khan cet' hɔ alo-m*
 anything-3PL:SUBJ ask-2DL:OBJ-TEMP what DISC PROHIB-2SG:IMP
ləi-a-ko-a
 tell-APPL:3PL:OBJ-FIN
 'When they ask you do not tell them anything.'

Temporal clauses may also be marked by *tarpɔre* postposed to the main clause.

- (240) *laha-te jɔm-le-m tarpɔre-lay cal-ak'-a*
 before-LOC eat-IRR-2SG:IMP then-1DL:SUBJ go-M-FIN
 'Eat first then we will go.'

Conditional clauses are formed by *-khan* with or without the irrealis *-le* and *-len*. Sometimes the particle *jodi* 'if' is used when introducing the protasis.

- (241) *am-em hukum-en-khan noa bə-ɲ kəmi-ke-a*
 2SG-2SG:SUBJ order-PST:M-COND this NEG-1SG:SUBJ work-OPT-FIN
 'If you order I would not do it.'

- (242) *jodi thik dam-ijn nam-le-khan pasec'-ijn əkrɪj-ki-kin-a*
 if actual price-1SG:SUBJ get-IRR-COND probably-1SG:SUBJ sell-OPT-3DL:OBJ-FIN
 'If I get actual price I would probably sell the two.'

Purpose clauses are introduced by the postposition *ləgit'* 'for' with verb forms as well as with *jəmən* in between the main and purpose clauses. Verb forms taking *ləgit'* are also marked for applicative and pronominal clitics.

- (243) *ijn am noa katha met-a-m-ləgit'-ijn həc'-len-a*
 1SG:SUBJ 2SG this word tell-APPL-2SG:OBJ-for-1SG:SUBJ come-PLUP:M-FIN
 'I had come to tell you.'

A purposive clause with *jəmən* 'so that' is postposed to the main clause, and the verb of the purpose clause is marked by TAM suffixes and pronominal clitics.

- (244) *khub səkɔ-kate tɔl-me jəmən alo laɣ-a*
 very hard-CV tie-2SG:IMP so that PROHIB shake-FIN
 'Tie it hard so that it does not shake.'

Causal clauses are expressed by *cedak' se* occurring before the causal clause. The causal clause comes after the main clause, and the clause is marked for TAM, suffixes, and pronominal clitics.

- (245) *uni dɔ chuɟi-i hatao-ked-a cedak' se uni dɔ-e*
 3SG TOP leave-3SG:SUBJ take-PST:A-FIN because 3SG TOP-3SG:SUBJ
ruə-k'-kan-a
 fever-M-COP-FIN
 'He took leave because he is ill.'

4.2.4 Relative-type clause

Santali shows two types of relative clauses: one has the indefinite pronouns *jāhāe* and *jāhā*; while the first one is for the animate, the second one is for the inanimate. The other has the borrowed relative pronoun *ja* 'which'.

- (246) *jāhā kɔlɔm-tɛ-m ɔl-akad-a ona dɔ oka-re*
 any pen-INS-2SG:SUBJ write-PRF:A-FIN that TOP which-LOC
 'Where is the pen which you have written with?'
- (247) *jāhā dare-rɛ-m dɛc'-len-a on-re*
 Any tree-LOC-2SG:SUBJ climb-PLUP:M-FIN that-LOC
mit'-taŋ tɛrɔm cak mena-k'-a
 one-CLSSFR honey-comb exist-M-FIN
 'There is a honey-comb in the tree which you climbed.'
- (248) *jāhāe ato oɣak'-rɛ-ko taken-a onko dɔ ədɪ-ko*
 any one village house-LOC-3PL:SUBJ stay-FIN 3PL:SUBJ TOP very-3PL:SUBJ
pɛɣa-a
 friend-FIN
 'Those who stay in the village are very friendly.'

- (249) *am ja-m men-ked-a thik-ge-a*
 2SG that-2SG:SUBJ say-PST:A-FIN right-FOC-FIN
 ‘The word you said is right.’

4.2.5 Correlatives

Correlative constructions in Santali are formed by using pronouns or correlative particles in both the main and attributive clauses.

- (250) *je hilok’ uni-ŋ nel-led-e-a un hilok’*
 which day 3SG-1SG:SUBJ see-PLUP:A-3SG:OBJ-FIN that day
do sombar tahēkan-a
 TOP Monday COP:PST-FIN
 ‘The day I saw him was Monday.’
- (251) *oka disom-re onko gaḍel hḥ-ko*
 which country-LOC 3PL:SUBJ crow man-3PL:SUBJ
jarwa-akan-tahēkan-a ona disom-ren raj do gḥj-akan-a
 gather-PRF:M-COP:PST-FIN that country-GEN king TOP die-PRF:M-FIN
 ‘The king of the country where the crowd of people had gathered has died.’
- (252) *jḥkhṇ uni-ŋ met-a-e-kan tahēkan-a un-jokhan uni*
 when 3SG-1SG:SUBJ tell-APPL-3SG:OBJ-COP COP:PST-FIN that-time 3SG
bhagi-ḥkḥ-ṭe ḥnjḥm-et’-tahēkan-a
 good-attentive-INS hear-IMPRF-COP:PST-FIN
 ‘When I was telling him he was listening to it attentively.’
- (253) *jodi uni laha-te ba-e onḍo-kok’-a table hanḍe*
 if 3s gearly-LOC NEG-3SG:SUBJ come out-OPT-FIN then there
uni ba-e tiyok’-ke-a
 3SG NEG-3SG:SUBJ reach-OPT-FIN
 ‘If he would not come out earlier then he would not reach there’
- (254) *jēmṇ-ŋ men-led-a tēmṇ-ge cando ḥn-ak’ sana-e*
 as-1SG:SUBJ say-PLUP:A-FIN so-FOC Cando 1s-GEN wish-3SG:SUBJ
purḥu-keṭ’-t-ŋ-a
 fulfill:PST:A-POSS-1SG-FIN
 ‘Chando fulfilled my wish as I had asked.’
- (255) *jodi-pe uddhar-ŋ-a tḥbe ape-ren gate-pe*
 if -2PL:SUBJ rescue-1SG:OBJ-FIN then 2PL-GEN friend-2PL-2PL:SUBJ
ḥnam-e-a
 find-3SG:OBJ-FIN
 ‘If you rescue me you will find your friend.’

jāhā: ona, jāhāe: unilonko, and *jāhā: on-re* are types of constructions given in 246–248 and may be considered as correlative constructions.

5 SEMANTICS/DISOURSE

5.1 Semantics

Here I will briefly address a phenomenon, namely semantic agreement typical of Santali. Consider example (256):

- (256) *puṭkə-ko* *halaŋ-ket'-ko-a*
 mushroom-3PL:SUBJ collect-PST:A-3PL:OBJ-FIN
 'They have collected mushrooms.'

Consider the apparent contradiction between the object NP *puṭkə* 'mushroom' and its agreement in the verb. Apparently the object is inanimate and as a rule the inanimate is not marked in the verb. Here, in the case of Santali, the Santal's perception of animacy is reflected in the verb agreement, as the so-called inanimate objects like puff-balls, ear-wax, or thorns being pricked are considered animate in Santali. The correlation between the linguistic phenomenon of verb agreement and the concept of animacy is very close in Santali. The subjects and objects considered animates by the Santals are invariably marked in the verb in the form of third personal clitic pronouns. Consider the following example where an idol is considered animate:

- (257) *uni mit'-taŋ* *kəli boŋga-e* *benao-akad-e-a*
 3SG one-CLSSFR Kali goddess -3SG:SUBJ make-PRF:A-3SG:OBJ-FIN
 'He has made a Kali idol.'

Similarly sun, moon, and stars are also considered animate.

- (258) *sijcando* *rakap'-kan-a-e*
 sun rise-COP-FIN-3SG:SUBJ
 'The sun is rising.'

Another area I would like to address is the distinction between direct and indirect causation, and the problem of the degree of control retained in the causative macro-situation by the causee. The distinction between true causation and permission deserves special mention. Consider the following examples:

- (259) *dal-oco-ked-e-a-e*
 beat-CAUS-PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'He caused him to beat.'
- (260) *dal-oco-ad-e-a-e*
 beat-CAUS-APPL:PST:A-3SG:OBJ-FIN-3SG:SUBJ
 'He permitted him to beat.'

There is only one difference between the two constructions, that is, in (258) the object is marked by the applicative *a-*, here attached to the TAM, which is absent in (257). Thus while the cause in (257) is direct, making it a direct causation thereby involving the direct object, the cause in (258) is rather indirect, involving the indirect object with the applicative marker *a-*. It is easy to see the relationship between true causative and permissive. In both the constructions the marker for causation is the same, but while in the true causation (257) the causee is forced to

beat, in the indirection (258) the causee is persuaded and/or permitted to beat. In other words,

with the true causative, the anterior event/agent has the power to bring the effect about; in the permissive, the anterior event/agent has the power to prevent the effect from coming about. In both types, the realization of the effect is, at least partially, within the control of the causer/permitter. (Comrie 1989:171)

5.2 Discourse

adɔ is the marker of continuing discourse in Santali. *tea* is also used for the same purpose.

- (261) ... *təkhən bela duʔor-moto hoi-akan-a. adɔ*
 then part of the day two-about be-PRF:M-FIN thereupon
sahan-ko mit'-tʰen-ko dɔhɔ- jaora- ket'-tɛ jhɔtɔ
 firewood-PL one-place-3PL:SUBJ put down-together-PST:A-CV all
hɔʔ mit'-tʰen-ko duʔup'-en-a
 man one-place-3PL:SUBJ sit-PST:M-FIN
 'Then it was about 2 o'clock. Thereupon having put down all the firewoods together all men sat together in one place.'

- (262) *ona jhurkə-rɛ dɔ tʰik-em ɲam-a. tea dɔ uni kuʔi gidrə*
 that door-LOC TOP actual-2SG:SUBJ get-FIN then TOP DEM girl child
pəpəp sat-tʰi kapaʔ-e pəɾəm-idi-ked-a.
 consecutive seven-DEF door-3SG:SUBJ cross-away-PST:A-FIN
 'You will get the actual door in that. Thereupon that girl crossed seven doors one after another.'

The anaphoric *ac'* 'he/she/him/her' is used more often in continuing discourse than in ordinary speech. The particles like *du* 'look' are used as quotative in the discourse. The borrowed particle *je* 'that'³⁶ is used as a subordinate clause marker.

- (263) *adɔ hɔʔ-e kuli-ket'-ko-a*
 thereupon man-3SG:SUBJ ask-PST:A-3PL:OBJ-FIN
je maŋjhi haʔam oʔak' dɔ oka
 that village head oldman house TOP where
 'Thereupon the man asked them that where the house of the village headman was.'

Another area that needs to be mentioned is the use of pronouns among certain kin-relations in discourse. Consider the following examples:

- (264) Q. *ceka-en-a-ben, bəhu* A. *ruə-k'-kan-a-lɪŋ*
 how-PST:M-FIN-2DL, daughter-in-law Fever-M-COP-FIN-1DL:EX
 'How are you, daughter-in-law?' 'I am getting fever.'

When the father-in-law addresses the daughter-in-law he uses 2DL for a singular addressee, and in reply the daughter-in-law uses 1DL to talk about herself. This involves power and solidarity among in-law relations; the more powerful figure,

here the parent-in-law, uses 2DL to address the less powerful figure, here the daughter-in-law. The less powerful figure in turn uses 1DL while reporting about him- or herself to the more powerful figure to express solidarity with his- or herself, that is, with husband and wife. The concept of more and less powerful figures could better be exemplified by the speech pattern among the symmetric relations:

- (265) *henda ho sumdhi, cet'-leka mena-k'-bon-a*
 o (co)parent-in-law how be-M-1PL:INC-FIN
A. əđi muskil-rə-bon paṛao-akan-a
 very trouble-LOC-1PL:INC fall-PRF:M-FIN
 'Oh, co-parent-in-law, how are we (you)?' 'I (we) am in great trouble.'

This is a part of a discourse between two parents-in-law. In both the question and the answer 1PL inclusive is used to refer to and address, which denotes that the two in-laws belong to a symmetric power relation, and that is why they use the same kind of pronoun. When a plural form is used, the choice gives rise to certain speculation about the social status of the participants in the discourse. When a parent-in-law inquires about his counterpart, the latter's dependents are not left out. In Santal society the use of a singular pronoun in this case shows disrespect. Similarly, in addressing and referring to son/daughter-in-law the parent-in-law will always use 2DL, as socially they are supposed to address and refer to both spouses. The son/daughter-in-law, in reply, will use the first person dual exclusive to denote that they (both spouses) are mentally included.

The same kind of social relation and pronominal selection hold in the case of *bahṅpharea* ('a man and his younger brother's wife') and *ajhnarea* ('a man and the wife of his wife's younger brother') relations. But because of paucity of space it cannot be detailed.

6 LEXICON

As a Munda language, the status of Santali is much more faithfully preserved than any other language of the group. The core areas of the lexicon – the pronouns, demonstratives, grammatical morphemes, and numerals – are preserved despite prevailing bilingualism of the speakers of the language. The terms for kinship and for body parts are preserved in the vocabulary of the elders, with some changes in the vocabulary of the younger generation. The terms for daily needs are affected much by the influence of the neighbouring languages. The influence of the neighbouring languages cannot be avoided, as in everyday activity they either have to interact with the local non-Munda people or depend on them. In the states they are distributed in, they are the minorities. They have to have education in the medium of the neighbouring language, as education in their own language is still not available. In this kind of situation it is not always possible for them to maintain a clear-cut dichotomy between L1 and L2 – one being at home and among the in-group, and the other outside their home area in the out-group. So it is quite natural that their own language is affected, and more and more foreign components will find a place in their language. Still it is very deserving that the language is maintaining its distinct identity, as compared with many Munda languages like Kharia and Juang, which have almost lost their Munda identity. The reason may be in the high number of speakers Santali claims as opposed to the low number of speakers of Kharia and

Juang. Love for their own language is another factor helping the language to preserve its 'purity'. It is this love and persistent struggle for the language of its speakers which has earned the language a constitutional status. As the language has gained constitutional status and the script has been recognized as the medium of reading, writing, and printing of text, it can be assumed that the speakers of the language will be much more conscious than before in the maintenance of the language.

6.1 Austroasiatic/Munda components

Personal pronouns are more resistant to loans. Not a single pronoun is taken from the neighbouring languages. Moreover, the inclusive–exclusive distinction in the first person, and their use within a special set of relations, is maintained as faithfully as ever before. Most of the grammatical markers such as TAM suffixes, genitive suffixes, and applicative are of Munda origin. Terms relating to beliefs, rituals such as marriage, and funeral and ancestry are retained in the original form. The cardinal numerals are preserved fully. Though the younger generation is more accustomed to Indo-Aryan numerals, they can recognize the original numerals, and use them on special occasions. The only cardinal borrowed is for hundred, *sae*, used by both the older and younger generations. The explanation behind this may be that because counting by a hundred is not a Munda system, the term connected with the system has to be borrowed by the speakers of the language.

6.2 Loan strata

The majority of loans come from the neighbouring Indo-Aryan languages, like Bengali, Hindi, Oriya, and Assamese, with which the language has been in contact for centuries. The most notable of the loans from these languages are the ordinal numerals like *pəhil* 'first', *dəsar* 'second', and *tesar* 'third'; many postpositions like *ləgit* 'for', *sāo* 'with', *səŋge* 'along with', *upər* 'above', *bhītər/bhītri* 'inside', *then* 'with', *sathe* 'along with', etc. are borrowed either from Bengali or Hindi. The particles like *jodiljudi* 'if', *jemən* 'so that', *tahle* 'then', *to* emphatic, *ar* 'and', *təkhən* 'then', and *jəkhən* 'when' are from the same source. The masculine–feminine distinction ending in *-a* or *-i* as in *kala* vs. *kəli* 'deaf' and *koŋka* vs. *kuŋki* 'mad', is of Indo-Aryan origin. Although there is no exact figure, almost 20% of the lexemes of daily needs are borrowed from Indo-Aryan. As there are no statistics of the regional loans, it is very difficult to discuss the loans from Oriya or Assamese, which need further survey. Loans from other Munda languages like Mundari and Ho, with which it is also in contact, need in-depth investigation and comparative study. Otherwise it is difficult to specify the figure. It is also very risky in the sense that if the languages under consideration are of the same stock, the common vocabulary could be considered as cognates, rather than loans from one another.

7 BRIEF ANALYSED TEXT

The following text was collected from Gobinda Hansda, aged 40, of the village Jitpur, Jamtara, Santal Parganas, Jharkhand. The story describes the stepmother's attitude towards the stepson. The area from where the story was collected falls in the Northern dialect area of Santali.

Kəki gə
Stepmother

- (i) *mit'ʔec' casa həŋ-e tahēkan-a.*
one-CLSSFR farmer-man-3SG:SUBJ COP-FIN
- (ii) *uni-ren era dɔ ədi khaŋoa həŋ-e tahēkan-a.*
3S-GEN wife FOC very diligent (wo)man-3SG:SUBJ COP-FIN.
- (iii) *unkin-ren mit'ʔec' koŋa gidrə tahēkan-ta-kin-a.*
3DL-GEN one-CLSSFR boy child COP-POSS-DL-FIN.
- (iv) *unkin dɔ din ge ədi kurumuŋu-kin kəmi-i.*
3DL:SUBJ FOC day EMPH very diligent-3DL:SUBJ work-FIN.
- (v) *mənkan cikə-kate cɔ mit' din uni həŋ-ren oŋak'*
but how-CV ever one day 3SG man-GEN house
boŋga dɔ bəhək' lac' haso ŋam-ked-e-a.
god(wife) FOC head belly pain get-PST:A-3SG:OBJ-FIN
- (vi) *ar acka ge unkin apa hən*
and suddenly EMPH 3DL father son
bəgi-at'-kin-a-e.
leave-APPL.PST:A-3DL:OBJ-FIN-3SG:SUBJ.
- (vii) *khan ədi duk-rə-kin paŋao-en-a.*
then very trouble-LOC-3DL:SUBJ fall-PST:M-FIN.
- (viii) *kəmi hɔ nit dɔ baŋ-kin kəmi-daŋe-ak'-kan-a.*
work even now FOC NEG-3DL:SUBJ work-can-M-COP-FIN
- (ix) *cedak'je mayjiu kəmi-ko dɔ ar unkin oŋak'-re*
because that woman worker-PL FOC and 3DL house-LOC
ekkal bənuk'-ko-a.
absolutely NEG-3PL:OBJ-FIN
- (x) *adɔ ayma hudis-baŋa-kate mit'ʔec'*
thereupon much thought-about-CV one-CLSSFR
kəki-gɔ-kin saŋgha³⁷ əgu-ked-e-a.
aunt-mother-3DL:SUBJ marry a widow bring-PST:A-3SG:SUBJ-FIN.
- (xi) *uni əgu-kate thora din dɔ ʔhik-ge din-ko khema-ked-a.*
3S bring-CV some day FOC right-EMPH day-3PL:SUBJ pass-PST:A-FIN.
- (xii) *inə tayɔm dɔ uni gidrə kəki gɔ-ak' met'*
just that after FOC that boy aunt mother-GEN eye
samaŋ-re dɔ ədi sikɪŋ-ge ŋel-e-a.
front-LOC FOC very hate-EMPH see-3SG:OBJ-FIN.
- (xiii) *adɔ ona-te mit' din uni ayo dɔ ac'-ren*
thereupon that-INS one day that woman FOC 3S-GEN
herel-tet'-e met-a-e-kan-a, 'nui gidrə dɔ
husband-DEF-3SG:SUBJ say-APPL-3SG:OBJ-COP-FIN this boy FOC

- jāhā-sen idi-ofo-ka-e-me ar baŋkhan*
any-to drive-away-BEN-3SG:OBJ-2SG:IMP and otherwise
gɔc'-gidj-ka-e-me.'
kill-off-BEN-3SG:OBJ-3SG:IMP.
- (xiv) *khan herel-tet' ona katha aŋjom-toray*
then husband-INAL that word listen away
tiŋgit'-gɔt'-en-a.
deafen-instantly-PST:M-FIN.
- (xv) *ar mɔne mɔne-te men-jɔŋ-an-a je nui ayo dɔ*
and mind mind-LOC say-MP-APPL.PST:M-FIN that this woman FOC
cit katha-e met-ad-ɪn-a.
what word-3SG:SUBJ say- APPL.PST:A-1SG:OBJ-FIN.
- (xvi) *adɔ ədʒi hudis-rɛ-e paɣao-en-a.*
then very think-LOC-3SG:SUBJ fall-PST:M-FIN.
- (xvii) *adɔ-e kuli-ruəɣ-ked-e-a, 'cedak'-em*
then-3SG:SUBJ ask-return-PST:A-3SG:OBJ-FIN why-2SG:SUBJ
əɣis-a-e-kan-a'?
worry-APPL-3SG:OBJ-COP-FIN?
- (xviii) *adɔ uni ayo-e rɔɣ-ruəɣ-ked-a 'hɛ, ɪn dɔ*
then that woman-3SG:SUBJ tell-return-PST:A-FIN yes 1SG:SUBJ FOC
əɣis-gi-ŋ jɛl-e-kan-a'.
worry-EMPH-1SG:SUBJ see-3SG:OBJ-COP-FIN.
- (xix) *khan herel-tet'-e men-ked-a, 'am-ge ləi-me*
then husband-INAL-3SG:SUBJ say-PST:A-FIN 2S-EMPH tell-2SG:IMP
tɔbe cikə-kate-ŋ gɔj-e-a'.
then how-CV-1SG:SUBJ kill-3SG:OBJ-FIN.
- (xx) *adɔ uni ayo-e ləi-a-e-kan-a, 'am-ak' isi*
then that woman-3SG:SUBJ tell-APPL-3SG:OBJ-COP-FIN 2SG-GEN plough
dɔ jɔkhɔn si-ok'-ben jɔɣao-idi-a un jɔkhen
FOC when plough-M-2DL:SUBJ link-continue-FIN that time
gidrə dɔ laha-ka-e-me ar am dɔ tayɔm-re
boy FOC front-BEN-3SG:OBJ-2SG:IMP and 2SG FOC behind-LOC
si-me ar am-ren ɖaŋra khub laga laga-kin-me.
plough-2SG:IMP and 2S-GEN bullock very drive drive 3DL-2SG:IMP.
- (xxi) *un jɔkhen-ge uni gidrə dɔ ona isi-te-e*
that time-EMPH that boy FOC that plough-INS-3SG:SUBJ
gutu gɔjɔk'-a.
insert kill-M-FIN.
- (xxii) *khan ona aŋjom-kate goɬa bəd bəyhar-kin*
then that listen-CV whole upland low.land-3DL:SUBJ
si-caba-ked-a.
plough-finish-PST:A-FIN.

- (xxiii) *mɛnkhan uni gidrə gɔj-e-ləgit' ɔkte-ge bay hɛc'-len-a.*
but that boy kill-3SG:OBJ-for time-EMPH NEG come-PLUP:M-FIN
- (xxiv) *khan ayo-e mɛn-ked-a, 'sanam khet-ben si-*
then woman-3SG:SUBJ say-PST:A-FIN whole land-2DL:SUBJ plough
caba-ked-a. adɔ ɛnrɛ hɔ̃ nui gidrə dɔ ba-m
finish-PST:A-FIN yet still even this boy FOC NEG-2SG:SUBJ
gɔc'-daɣe-ad-e-a'.
kill-can-APPL.PST:A-3SG:OBJ-FIN
- (xxv) *khan ona ɔkte uni ayo gidrə gɔj-e-ləgit' mit'-tɛc' kuɽpaɽ-ked-a.*
then that time that woman boy kill-3SG:OBJ-for one-INAL effort-PST:A-FIN.
- (xxvi) *met-ad-e-a-e, 'hana tɛndʒi-rɛ gundli-bon*
say-APPL.PST:A-3SG:OBJ-FIN-3SG:SUBJ that.yonder plains-LOC millet-1PL:SUBJ
cas-akat' ona-ge si-ben'.
cultivate-PRF:A that-EMPH plough-2DL:IMP.
- (xxvii) *un jɔkhɛc'-ge uni gidrə dɔ ona isi-tɛ sɔb-ɔk'*
that time-EMPH that boy FOC that plough-INS pierce-M
gɔc'-oco-y-e-m.
die-CAUS-EUPH-3SG:OBJ-2SG:IMP
- (xxviii) *ar ona jayga-rɛ dɔ ɛtak' cas-laj lagao-a.*
and that place-LOC FOC any cultivate-1DL:SUBJ apply-FIN.
- (xxix) *adɔ uni hɛrɛl-tɛt' dɔ ona-ge hɛ-ad-a.*
thereupon that husband-INAL FOC that-EMPH yes-APPL.PST:A-FIN
- (xxx) *khan dɔsar hilok'-ge setak'-re gidrə-e*
then second day-EMPH morning-LOC boy-3SG:SUBJ
met-a-e-kan-a 'dɛlay-ta si-ok'-lay idi-a.
say-APPL-3SG:OBJ-COP-FIN look-DEF plough-M-1DL:SUBJ drive-FIN.
- (xxxi) *ona gundli-laj si-oco-g-a ar ɛtak'*
that millet-1DL:SUBJ plough-CAUS-M-FIN and another
cas-bon lagao-a'.
cultivate-1PL:SUBJ apply-FIN.
- (xxxii) *khan koɽa gidrə rɔɽ-ruɽ-ked-a, 'henda baba, ona gundli*
then boy child tell-return-PST:A-FIN oh father that millet
ma³⁸ bili-jut-akan-a.
OPT ripe-properly-PRF:M-FIN.
- (xxxiii) *gapa meaj khan-ge jɔm-jut-uk'-a.*
tomorrow after tomorrow then-EMPH eat-suitable-M-FIN.
- (xxxiv) *ona dɔ cedak'-laj si-bərij-a.*
that FOC why -1DL:SUBJ plough-waste-FIN.
- (xxxv) *khan uni gidrə-rɛn apa-t'-tɛt'-e hudis-ked-a, 'səri-ge*
then that boy-GEN father-INAL-DEF-3SG:SUBJ think-PST:A-FIN right-EMPH
nui gidrə dɔ bhage solha-ge-i ləi-a-ɲ-kan-a'.
this boy FOC good advice-EMPH-3SG:SUBJ tell-APPL-1SG:OBJ-COP-FIN

- (xxxvi) *adɔ mənɛ mənɛ-tɛ hudis-jəŋ-kan-a-e.*
 thereupon mind mind-LOC think-MP-COP-FIN-3SG:SUBJ
- (xxxvii) *hudis-katɛ gidrə-rɛn apa-t-tɛt' dɔ ac'-rɛn ayo-e*
 think-CV boy-GEN father-INAL-DEF FOC 3S-GEN wife-3SG:SUBJ
met-ad-e-a, 'jɪ dɔ nui gidrə dɔ ʒɔ-jɪ
 say-APPL.PST:A-3SG:OBJ-FIN 1SG:SUBJ FOC this boy FOC NEG-1SG:SUBJ
gɔc'-daŋe-ke-a'.
 kill-can-OPT-FIN
- (xxxviii) *ona aŋjəm-sāotɛ uni ayo dɔ ədɪ raŋgao-gət'-en- a-e*
 that listen-with that wife FOC very be angry-instantly-PST:M FIN-3SG:SUBJ
ar boge-tɛ-kin jhogɾa-en-a.
 and good-INS-3DL:SUBJ quarrel-PST:M-FIN
- (xxxix) *adɔ uni ayo dɔ-e laga-gidɪ-kad-e-a.*
 then that woman FOC -3SG:SUBJ drive-away-BEN.PST:A-3SG:OBJ-FIN
- (xxxx) *mucɛt'-en-a.*
 finish-PST:M-FIN

TRANSLATION

1. There was a farmer. 2. He had a diligent wife. 3. They had a son. 4. They always worked very hard. 5. The farmer's wife had a sudden attack of headache and stomach-ache. 6. Soon she passed away leaving behind her husband and son. 7. Thereafter these two faced lots of problems. 8. They could not even go to work. 9. The absence of a woman to take care of household chores led to this situation. 10. Thereupon after a lot of contemplation they brought a stepmother for the boy by *sangha* marriage. 11. After she arrived things worked well for sometime. 12. After sometime the boy was looked upon with hatred by the stepmother. 13. One day the woman told her husband, 'Drive away the child anywhere or otherwise kill him.' 14. Having heard that the husband felt as if he had been deafened. 15. He reflected over what the woman told him. 16. He then fell into deep thought. 17. He asked her, 'What bothers you about the boy?' 18. Then she replied, 'I am looking at him with great fear.' 19. He told her, 'Suggest to me how to kill him.' 20. She suggested, 'When you two go to work in the field, you should link your plough then keep the child in front of you, and you plough from behind driving your bullocks very hard.' 21. 'Then the child will die being pierced by the yoke.' She continued. 22. After that the father and son ploughed up the whole high land and low land for many days. 23. But he never got around to killing his son. 24. The woman told the farmer, 'You have ploughed the whole field and still you could not kill the boy.' 25. Then the woman made another suggestion on how to kill the boy. 26. She told the farmer, 'In that far off plain (where) you plough for our millet cultivation.' 27. 'At that time when you are ploughing that field, let the boy die by being pierced by the yoke', she said. 28. 'And in that piece of land we will cultivate other crops.' 29. The farmer told her he would do as she told him. 30. The next morning he told his son, 'Look child, we will take the plough.' 31. 'We will plough up the millet and cultivate another crop.' 32. Then the boy replied, 'Oh father, that millet has ripened.' 33. 'Tomorrow or the day after it will be edible.' 34. 'Why shall we waste the

crop by ploughing?' 35. The farmer thought, 'The child is giving me advice in good spirit.' 36. He then made up his mind. 37. He told his wife, 'I can never kill this boy.' 38. On hearing that the woman became very angry and quarreled with him. 39. Then he drove the woman away. 40. The story ends here.

NOTES

- 1 S.K. Chatterji, 'Foreword' to *HOR BAPLA PUTHI (Pahil hatin)* by Stephen H. Murmu, Benagaria: Mission Press, pp. v–xi, 1961.
- 2 L.O. Skrefsrud, 'Introduction' to *A Grammar of the Santhal Language*, Benares: Medical Hall Press, 1873.
- 3 S.K. Chatterji, 'Foreword' to *HOR BAPLA PUTHI (Pahil hatin)*.
- 4 Pinnow in his later papers took up the position of the Munda languages and comparing the external relationship of the Munda languages with other Austroasiatic ones he revised his earlier thesis slightly, but the position of Munda remains the same, that is, Munda-Nahali is separated off as the western branch; cf. H.J. Pinnow's (1960, 1963) 'Über den Ursprung der voneinander abweichenden Strukturen der Munda Und Khmer-Nikobar Sprachen', *Indo-Iranian Journal* 4(1): 81–103. 'The position of the Munda languages within the Austroasiatic family', in H.L. Shorto (ed.), *Linguistic Comparison in South-East Asia and Pacific*, pp. 140–152.
- 5 Here the name conforms to that of Robert Shaffer (1940), F.B.J. Kuiper (1962), and S. Bhattacharya (1957) rather than Pinnow's Nihali.
- 6 Norman H. Zide, 'Munda and non-Munda Austroasiatic languages', in T.A. Sebeok (ed.), *Current Trends in Linguistics*, 5: 411–433. The Hague: Mouton.
- 7 S. Bhattacharya, 'A New classification of Munda', *Indo-Iranian Journal*, XVII:1, (1975): 97–101.
- 8 EDITOR'S NOTE: Both Remo and Gta? have dual somewhere active in their grammars.
- 9 *Ethnologue* 14th edition 2000.
- 10 Census of India, 1981 Series 1, INDIA-Part iv-B (i), Population by Language/ Mother Tongue (table C-7), New Delhi: Registrar General and Census Commission, India.
- 11 The sadar sub-division of Bankura covers the western and southern parts of the district.
- 12 The districts having Santal population in the northern part of West Bengal are Murshidabad, Malda, North and South Dinajpur, Jalpaiguri and Cochbehar.
- 13 Ghosh (2003:11) argues that 'If all the ... problems are resolved we may come up with different type of classification of the Munda where South Munda may or may not find a place. We may even come up with a different theory of migration where North Munda including Kharia-Juang may be found to enter India through the Himalayan range, leaving traces of their migration in the pronominalized Himalayan languages.'
- 14 Census of India, 1981, Series 1, India-Part iv-B(ii), Population by Bilingualism (table C8), 4.
- 15 R.N. Cust, *A Sketch of the Modern Languages of the East Indies*, London: Trubner, 1878.
- 16 *The Ethnologue* (Grimmes 1996) mentions six dialects in India: Karmali (Khole), Kamali-Santali, Lohari-Santali, Mahali (Mahle), Manjhi, Paharia.
- 17 A. Campbell (1988: preface) notes that 'Northern Santali or that spoken in Bhagalpur, Munghyr, the Santal Parganas, Birbhum, Bankura, Hazaribagh and Manbhum, is the language of the overwhelming majority of the tribe, and

- is more polished than Southern Santali. The former is, therefore, regarded as the Standard, and Southern Santali, or that spoken in the remaining districts, as a dialect, or, possibly, a group of dialects of it.
- 18 Ghosh (1994:8) notes 'Covering only West Bengal and Santal Parganas of Bihar dialectal situation in Santali has been chalked out and in that case also Campbell's demarcation as Northern and Southern Santali has been confirmed.'
 - 19 /p>[b]/-V is optional in case verb roots, obligatory in case of TAM.
 - 20 Occurrence of /h/ in final position is found only in interjectives, otherwise /h/ does not occur finally.
 - 21 With Applicative *a*- Past and Pluperfect are realized as *at'* < *a+ket'* and *a+let'*.
 - 22 With Benefactive *-ka* Past Pluperfect are realized as *-kat'* < *ka+ ket'* and *ka+ let'*.
 - 23 Causative and Permissive are expressed by the same suffix *-oco*, which in Southern Santali is realized as *-hɔcɔ*.
 - 24 Applicative in the perfect is infixted between the perfect suffix and the voice marker *-t'* in the active, i.e. *-aka+a+t'*, euphonic *w* comes between the final *-a* of *aka* and Applicative *a*.
 - 25 In the morphophonology it is shown that this *n* is a phonological increment. Considering the verbal paradigms it is better to consider it as a middle voice marker.
 - 26 An exception is Negative Imperative which takes the finite marker *-a*.
 - 27 An exception is Negative Imperative which takes the finite marker *-a*.
 - 28 Neukom (2001:123) tries to describe this formation as Completive. The stem is conjugated throughout with TAM suffixes. For a similar formation in Kharia see Peterson (this volume).
 - 29 The infixation is the only process of forming the reciprocal stem in the Southern dialect while in the Northern the stem is sometimes formed by suffixing *-ok'* as in *dal-ok'-kan-a-kin* 'They (two) are fighting each other.' The suffixed reciprocal is used by male teenagers in the Northern dialect. It is then, unlike infixted stem, conjugated in the Simple present/Future, Present and Past progressive and in the Imperative. In other tenses and moods simple roots are conjugated in the middle.
 - 30 The root also functions as an auxiliary, vide section 3.2.12.
 - 31 Reduplicated adjective indicate plurality of head nouns, compare Bangla *paka paka am* 'ripe mangoes', *choṭo choṭo gach* 'small trees'.
 - 32 The story 'sikhṇət' written by Upen Kisku is taken from a collection of short stories 'Mit' sae mit kəhni' (one hundred and one stories) published by West Bengal Tribal Development Corporation Ltd with introduction and annotation by S.K.Bhowmik.
 - 33 Examples of verb serialization given by Neukom (2001:176) are drawn from Santali Bakher which is connected with the ritual performed after death. These constructions are like Hindu *mantra*, not of colloquial language.
 - 34 Compare *bole* in Bangla as in *ami jabo bole janiechi* 'I have informed I would go.'
 - 35 Compare Bangla *kache*, Old Bangla *thiṇe* as in *amar kachelṭh iṇe ṭaka nei* 'I have no money with me.'
 - 36 *je* is a relative pronoun in Bangla, used also as relative clause marker.
 - 37 *saṅgha* is a kind of marriage prevalent among the poor Santals. In this type of marriage no social function is organized. A widow or divorcee is taken as a mate and no vermilion is applied in this type of marriage.
 - 38 *ma* is a particle used regularly with the Optative of the verb; with the Imperative it has an optative sense. In connection with clauses it implies statement of reason or cause.

BIBLIOGRAPHY

- Ahmed, E. (1969) 'Santal parganas', *Encyclopaedia Britannica 19*, Chicago: William Benton.
- Bagchi, P.C. (1929) *Pre-Aryan and Pre-Dravidian in India*, Calcutta: University of Calcutta.
- Banerjee, G.C. (1894) *Introduction to the Kharia Language*, Calcutta: Bengal Secretariate Press.
- Bhaduri, M.B. (1931) *A Mundari-English Dictionary*, Calcutta: Calcutta University.
- Bhat, D.N.S. (1997) 'Noun-verb distinction in Munda languages', in A. Abbi (ed.) *Languages of Tribal and Indigenous Peoples of India*, Delhi: Motilal Banarasidas, 227–251.
- Bhattacharya, S. (1954) 'Studies in the Parengi language', *Indian Linguistics*, 14(III): 45–63.
- (1957) 'Field-notes on Nahali', *Indian Linguistics*, 17: 245–258.
- (1965) 'Glottal stop and checked consonants in Bonda', *Indo-Iranian Journal*, IX(1): 69–71.
- (1966) 'Some Munda etymologies', in N.H. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton, 28–40.
- (1968) *A Bonda Dictionary*, Poona: Deccan College.
- (1969) *Some More Munda Etymologies*, Poona: Pratidanam.
- (1970) 'Kinship terms in the Munda languages', *Anthropos*, 65: 444–465.
- (1975) 'Munda studies: A new classification', *Indo-Iranian Journal*, XVII:1, 97–101.
- Biligiri, H.S. (1965) *Kharia: Phonology; Grammar and Vocabulary*, Poona: Deccan College.
- (1965) 'The Sora verb, a restricted study', in G.B. Milner and E.A. Henderson (eds) *Indo-Pacific Linguistic Studies II*, Amsterdam: North-Holland Publishing Company, 231–250.
- Bodding, P.O. (1922) *Materials for a Santali Grammar I (Mostly Phonetic)*, Dumka: Santal Mission of the Northern Churches.
- (1925) *Santali Folk-Tales 1-3*, Oslo: H. Ascheboug & Co.
- (1929) *Materials for a Santali Grammar II (Mostly Morphological)*, Dumka: Santal Mission of the Northern Churches.
- (1929) *A Santali Grammar for Beginners*, Benagaria: Santal Mission Press.
- (1929–1936) *A Santal Dictionary, 1-5*, Oslo: A.W. Broggers Boktrykkeri.
- Burrows, L. (1915) *Ho Grammar*, Calcutta: Catholic Orphan Press.
- Campbell, A. (1953) *Santali-English Dictionary*, 3rd edn, R.M. Macphail (ed.), Benagaria: Santal Christian Council.
- (1954) *English-Santali Dictionary*, 3rd edn, R.M. Macphail (ed.) Benagaria: Santal Christian Council.
- Chatterji, S.K. (1923) 'The study of Kol', *The Calcutta Review*, VIII: 451–473.
- (1926) *The Origin and Development of the Bengali Language*, Calcutta: Calcutta University.
- (1953) 'The languages of the Adivasis', *March in India*, 6(2): 14–16.
- (1961) 'Foreword' in S.H. Murmu *HOR BAPLA PUTHI(pahil hatin)*, Benagaria: Mission Press, V–XI.
- Cole, Rev. F.T. (1875) 'Santali riddles', *Indian Antiquary*, IV: 164ff.
- (1879) 'List of words and phrases with their Santali equivalents', *Indian Antiquary*, VIII: 194–196.
- (1906) *Santali Primer*, Pokhuria: Santal Mission Press
- Comrie, B. (1985) 'Causative verb formation and other verb-deriving morphology', in T. Shopen (ed.) *Language Typology and Syntactic Description*, Cambridge: CUP, 309–348.

- Comrie, B. (1989) *Language Universals and Linguistic Typology*, Chicago: The University of Chicago Press.
- Cook, W.A. (1965) *A Descriptive Analysis of Mundari: A Study of Structure*, unpublished PhD dissertation, Georgetown University.
- Cust, R.N. (1878) *A Sketch of the Languages of the East Indies*, London: Trubner.
- Friesen, V.I. (1983) 'Finite and non-finite verbal forms in Santali', *Indian Journal of Linguistics*, 10: 169–184.
- Ghosh, A. (1980) 'Santali kinship terminology: A sociolinguistic perspective', *Calcutta Papers in Linguistics*, I(1): 7–23.
- (1980) 'Some notable findings in Santali as compared with that of P.O. Boddling (1929)', Paper presented at the All India Conference of Dravidian Linguists, Osmania University, Hyderabad.
- (1988) *BIBLIOTHECA AUSTRASIATICA*, Calcutta: Firma KLM Pvt. Ltd.
- (1992) 'Tagmemic analysis of some aspects of the Santali transitive verb', in J.C. Sharma (ed.) *From Sound to Discourse: A Tagmemic Approach to Indian languages*, Mysore: Central Institute of Indian Languages, 45–78.
- (1994) *Santali – A Look into Santali Morphology*, New Delhi: Gyan publishing house.
- (2003) *An Ethnolinguistic Profile of Eastern India – A Case of South Orissa*, Burdwan: University of Burdwan.
- Greenberg, J.H. (1978) 'Generalizations about numeral systems', in J.H. Greenberg (ed.) *Universals of Human Language*, Stanford: Stanford University Press, 249–295.
- Grierson, G.A. (1906) *Linguistic Survey of India IV*, Calcutta: Superintendent of Government Printing.
- Grimes, B.F. (ed.) (1996) *Ethnologue. Languages of the World*, Dallas: SIL.
- Gumperz, J. and Biligiri H.S. (1957) 'Notes on the phonology of Mundari', *Indian Linguistics*, XVII: 6–15.
- Haspelmath, M. (1995) 'The converb as a linguistically valid category', in M. Haspelmath and E. König (eds) *Converbs in Cross-Linguistic Perspective*, Berlin: Mouton de Gruyter, 1–55.
- Hoffmann, Rev. J. (1903) *Mundari Grammar*, Calcutta: Bengal Secretariate Press.
- (1950) *Encyclopaedia Mundarica*, Patna: Superintendent Govt. printing.
- Hunter, W.W. (1868) *A Comparative Dictionary of India and High Asia*, London: Trubner.
- Konow, S. (1904) 'Mundas and Dravidas', *The Indian Antiquary*, 33: 121–125.
- Kuiper, F.B.J. (1962) 'Nahali – a comparative study', *Mededelingen der Koninklijke Nederlandse Akademie Van Wetenschappen*, 25: 239–352.
- (1965) 'Consonant variation in Munda', in G.B. Milner and E.J.A. Henderson (eds) *Indo-Pacific Linguistic Studies I*, Amsterdam: North Holland Publishing Co, 54–87.
- Macphail, R.M. (1935) 'Notes on Santals', in *Census of India 1931, 1: III*, Simla: Government of India Press.
- (1983) *An Introduction to Santali*, (Reprint) Calcutta: Firma K.L. Mukhopadhyay.
- Mahapatra, K. and Zide N.H. (1972) 'GTA? Nominal combining form', *Indian Linguistics*, 33(3): 179–202.
- Man, E.G. (1867) *Sonthalia and the Sonthals*, Calcutta: Geo Wyman & Co.
- Martin, A. (1898) *English-Santali Vocabulary*, Benares: Medical Hall Press.
- Mitchell, Rev. M. (1875) 'Santali songs with translations', *Indian Antiquary*, IV: 34f.
- Munda, R. (1969) 'Aspects of Mundari verb', *Indian Linguistics*, 30(2): 27–49.
- Muscat, G. (1989) *Santali. A New Approach*, Sahibganj: Santali Book Depot.
- Neukom, L. (2000) 'Argument marking in Santali', *Mon-Khmer Studies*, 30: 95–113.
- (2001) *Santali*, Muenchen: Lincom Europa.
- Osada, T. (1996) 'Notes on the Proto-Kherwarian vowel system', 39: 245–258.
- Phillips, J. (1845) *A Santali Primer*, Calcutta: School Book Society.
- (1850) *Sequel to a Santali Primer*, Calcutta: School Book Society.

- (1852) *An Introduction to the Santal Language*, Calcutta: School Book Society.
- Pinnow, H.J. (1959) *Versuch einer historischen Lautlehre der Kharia-Sprache*, Wiesbaden: Otto-Harrassowitz.
- (1965) 'Personal pronouns in the Austroasiatic languages: A historical study', in Milner and Henderson (eds) *Indo-Pacific Linguistic Studies I*, Amsterdam: North Holland publishing Co., 3–42.
- (1963) 'The position of the Munda languages within the Austroasiatic language family' in H.L. Shorto (ed.) *Linguistic Comparison in South East Asia and Pacific*, 140–152.
- (1966) 'A comparative study of the verb in the Munda languages', in N.H. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton, 96–103.
- Puxley, Rev. E.L. (1868) *A Vocabulary of the Santhali Language*, London: W.M.Watts.
- Sebeok, T.A. (1942) 'An examination of the Austroasiatic language family', *Language*, XVIII(III): 206–217.
- (1943) 'The phonemic system of Santali', *Journal of the American Oriental Society*, 63: 66–68.
- Skrefsrud, L.O. (1873) *A Grammar of the Santhal Language*, Benares: Medical Hall Press.
- Stampe, D. (1965–1966) 'Recent work on Munda Linguistics I–IV', *IJAL*, 31: 332–341, 32: 74–80, 164–168, 390–397.
- Starosta, S. (1967) *Sora Syntax*, unpublished PhD dissertation, University of Wisconsin.
- (1971) 'Derivation and case in Sora verbs', *Indian Linguistics*, 32:3, 194–206.
- Suryakumari (1991) *A Descriptive Study of Santali*, PhD dissertation, Central Institute of Indian Languages, Mysore.
- Zide, A.K. (1978) 'A Note on glottalization and release in Munda', *Indian Linguistics*, 39(1–4): 70–75.
- Zide, N.H. (1958) 'Final stops in Korku and Santali', *Indian Linguistics*, I: 44–48.
- (1966) 'Low tone and proto-Korku-Kherwarian vowel system', in N.H. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton, 214–229.
- (1972) 'A Munda demonstrative system: Santali', in J.M.C. Thomas and L. Bernot (eds) *Langues et techniques. Nature et Societe. Tom I*, Paris: Klincksieck, 267–274.
- (1978) *Studies in Munda Numerals*, Mysore: Central Institute of Indian Languages.

ANNOTATED BIBLIOGRAPHY

The Santals are the most numerous of the Munda groups and the study of their language and culture dates back to the early nineteenth century. The language was first brought to light by foreign missionaries. Rev. J. Phillips (1845) was a pioneer among the European scholars who studied the Santali language. His *A Santali Primer* (1845), *Sequel to a Santali Primer* (1850), and *An Introduction to the Santal Language* (1852) were sensations in the scholarly world. In his introduction he submits with 'unfeigned diffidence': '...a first effort... necessarily exposed to many errors, which time and careful observation alone can correct. In the absence of all established authority, much difficulty has been experienced in fixing the orthography and ascertaining the exact meaning of many words, as they are differently used and differently pronounced by different persons.' As the language had no character of its own, he used Bengali characters 'to supply the deficiency'. The verb, being the most complicated, has been given much attention. Some thirty-five pages have been devoted to verbs (pp. 13–51). Still, vocabulary has been given the most attention (pp. 83–186).

L.O. Skrefsrud's *A Grammar of the Santhal Language* (1873) was the first authoritative and comprehensive study of the Santali language, although Skrefsrud had adopted the old name Turanian as the ultimate source of the language. The grammar

is divided into two parts – the first dealing with phonology and morphology, although not named as such, and the second with syntax. The conditions under which the four checked consonants are changed to their corresponding voiced counterparts are discussed in the section ‘of the permutation of letters’. Santali verb morphology is of the greatest complexity, and Skrefsrud’s realization of it can be seen from the range of discussion; while the entire phonology and morphology have been discussed in 40 pages, the verb is discussed in 250 pages.

F.T. Cole attempted in his *Santali Primer* (1906) ‘to clear up some of the difficulties experienced by beginners in their first efforts to acquire their difficult and idiomatic language’. The verb morphology is discussed in detail with the remark, ‘The wonderful intricacy of the verbal system is a marvel to the student and considering the mental caliber of the Santals of the present day, one is all the more surprised to find such a complex, and yet perfectly regular, verbal system.’ To some extent it is a simplified version of Skrefsrud.

G.A. Grierson (1906) in his *Linguistic Survey of India*, Vol. IV, discussed the sounds and grammatical features. A skeleton grammar of Santali was also given, in which nouns, pronouns, ‘conjugational bases’, ‘inflexional bases’, and negative particles are entered. He recognized locality-based differences in the language, and those have in recent years given rise to differences in the languages spoken in the east where most loanwords come from Bengali of the west which chiefly borrows from Bihari, and the south where the influence of Oriya is felt.

P.O. Bodding’s *Materials for a Santali Grammar* Parts I and II (1922, 1929) is the first monumental and the most authentic of the grammars written in the first half of the twentieth century, and still remains unsurpassed. The sounds, stress, syllabification, and euphonic harmony are described with minute observations. In the part of morphology (named Mostly Morphological) ‘much syntactic matter has come in, partly because omission would make it more difficult for the student to understand the peculiar working of this agglutinating language’. The verb made unnecessarily complicated by Skrefsrud is made simple and accessible. Had Bodding not died early, he could have completed the proposed part of syntax. Yet the work done by him in the field of Santali linguistics is still unsurpassable, and future scholars will remember him with reverence. His *A Santali Grammar for Beginners* (1929) is a concise and simplified version of his *Materials for a Santali Grammar*. His *A Santal Dictionary* Vols 1–5 (1929–1936) remains a thesaurus of the Santali language. His *Santali Folk Tales* Vols 1–3 (1925) is still the best Santali text collection.

In the very preface of his *Introduction to Santali* (1983), R.M. Macphail submits, ‘As its name implies, this book does not pretend to be a full grammar. Such a work already exists, in Bodding’s “Materials for a Santali Grammar,” parts I and II; but that is not an easy book for the beginner, and the need has been felt for something simpler. This is an attempt to meet that need.’ It is actually meant to help the new learners of the language to find easily how the different forms and functions are formalized. A. Campbell’s *Santali-English Dictionary* (1953) and *English-Santali Dictionary* (1954) deal with common Santali words and their use.

Three works have been done recently, two of them in the last quarter of the last century, Ghosh (1994) and Suryakumari (1991), and one in the beginning of this century, Neukom (2001). Ghosh has dealt with the morphology of the language. Suryakumari and Neukom have dealt with phonology, morphology, and syntax. While Ghosh and Suryakumari have analysed data collected through a field survey, Neukom’s data are mostly drawn from Bodding. Muscat (1989) is basically meant for pedagogical purpose.

MUNDARI*

Toshiki Osada

1 INTRODUCTION

Mundari is mainly spoken in the state of Jharkhand, which was recently set up by the Government of India on 15 November 2000, and in the adjoining states of Orissa and West Bengal in India. *Munḍa* means ‘village-headman’ in Mundari. But the language name *munḍārī* is given by the neighbouring peoples, the indigenous name is *hoḥo jagar* ‘human language’ or *munḍa jagar* ‘Munda language’.

As Gregory Anderson shows us in the Introduction of this book, Mundari belongs to the Kherwarian group of the North Munda branch. According to the Census of India 1991, the number of speakers of Mundari is 861,378. The same Census reports the number of speakers of Munda as 413,894. The names Munda and Mundari seem to confuse. It is likely that the census officer did not have perfect criteria for naming the languages in India. There is actually no difference between the Munda language and the Mundari language linguistically. Thus, the total number of the speakers of Mundari is likely to be more than one million. From a linguistic point of view, the designation Munda is used for the language family. Mundari, on the other hand, refers to an individual language, namely the language of Munda people.

As Hoffmann reported in the *Encyclopaedia Mundarica*, Vol.1, page (6), Mundari has four dialects; that is, Hasada from *hasa-da?* ‘(literally) land water (place name)’ in Mundari, Naguri from *naguri* (place name), Tamaria from *tamar-ia* ‘language of Tamar (place name)’, and Kera from *kera?* (perfect ending, instead of *keda* in another dialect). Munda (1980:kha) has proposed the name Latar dialect (*latar* means ‘low’) instead of Tamaria. I do not adopt this term here because I have never heard *latar jagar* in Mundari.

The Hasada dialect is considered as the standard variety among Munda peoples. Hasada speakers are located on the eastern side of Ranchi–Chaibasa Road while Naguri speakers are situated on the western side. The Tamaria dialect is distributed in the Panchpargana area (Bundu, Tamar, Silli, Baranda and Rahe). Further, Kera is mainly spoken by the inhabitants of Ranchi city and the adjacent area, who ethnically belong to the Oraon tribe. According to Pinnow (1959:2), Ho should be considered as a dialect of Mundari from a linguistic point of view.¹ We, however, regard the Ho language as a separate language on the basis of the ethnic identity of its speakers (see the chapter on Ho and the other Kherwarian languages in this book).

The study of Mundari started in the nineteenth century; for example, Haldar (1871), Whitley (1873), Nottrott (1882). These works are neither comprehensive, nor reliable from a linguistic point of view. For example, the glottal stops were not described in these works. Linguistically oriented grammars have been written by Hoffmann (1903), Cook (1965) (his data are collected not by him but by Hoffmann), Sinha (1975) (his descriptions contain a lot of self-contradiction and some data are not

reliable), Munda (1980) (this is written in Hindi and contains reliable data but is not comprehensive), and Osada (1992) (the section on syntax is very poor); phonology by Gumperz with Biligiri (1957) and Sinha (1974) (the data are not reliable; the same as Sinha 1975); verbal morphology by Langendoen (1966, 1967) (his data are based on the Naguri dialect; he applied Mundari data to the standard theory by Chomsky but unsuccessfully as I show in section 3.2), Munda (1971) (this paper is focussed on aspect but incomplete); morpho-syntax by Osada (1999, 2007). A dictionary of Mundari has been compiled by Hoffmann (1930–1978), Bhaduri (1931), Prasad (1973, 1976) (in these dictionaries she missed a description of the glottal stops), and Mundu (1995). The most influential work is Hoffmann’s *Mundari Grammar* (=MG) and *Encyclopaedia Mundarica* (=EM). The descriptions in MG and EM differ in dialect. MG is mainly based on Naguri while EM mainly on Hasada. The data in EM are more comprehensive and reliable than those in MG. In addition to EM, Munda as a native speaker has given us reliable data. Thus I utilize the data from EM and Munda (1971, 1980).

2 PHONOLOGY

2.1 Phonemic inventory

Mundari has a five-vowel system as shown in Table 3.1. Vowel length and nasalization are not phonemic. It is, however, very important to make the distinction phonetically. As regards vowel length, an open and monosyllabic /CV/ is realized as two morae; for example, *lrul* ‘to beat a drum’ [ru:].²

Vowel nasalizations are found in the following circumstances:

- (i) / $(C)V\eta V(C)$ / for example, *lceŋel* ‘bird’ [tʃɛŋɛ̃], *laŋeʔl* ‘to pour out a liquid’ [ãŋɛ̃ʔ], etc.
- (ii) / $C^N V$ / (C^N means a nasal consonant)
for example, *lmul* ‘nose’ [mũ:], *lnul* ‘to drink’ [nũ:]
- (iii) / jV / (optionally)
for example, *ljil* ‘smell’ [dʒĩ:], *ljal* ‘any’ [dʒã:] or [dʒa:], but *ljol* ‘fruit’ [dʒɔ:].
- (iv) *loel*, *loal*, *lual* (optionally) for example, *lkoel* ‘beggar’ [kõẽ] *lkoasil* ‘fog’ [kõãsi], *lcual* ‘to extract a liquid by fire’ [tʃũã] or [tʃua].

If expressives are considered, nasalization becomes (very marginally) contrastive. We note the following minimal pair in the expressives *soe soe* ‘sound of boiling water’ and *sõẽ sõẽ* ‘to sit in a slovenly fashion’.

TABLE 3.1: MUNDARI VOWELS

	Vowel inventory		
	Front	Central	Back
High	i		u
Mid	e		o
Low		a	

Mundari has 23 consonants (Table 3.2) including 10 stops: *p, b, t, d, ʈ, ɖ, c, j, k, g*; one sibilant: *s*; three liquids: *r, ʎ, l*; five nasals: *m, n, ŋ, ñ, ŋ*; two glides: *w, y*; and two glottals: *h, ʔ* given in the consonant inventory chart. All stop consonants except the two glottals *h, ʔ* appear in word-initial and word-medial position. In word-final position, the distinction between voiced stops *p, t, k* and voiceless stops *b, d, g* is neutralized and realized as checked consonants *b, d* and a glottal stop *ʔ* except in recent loanwords. The retroflex stops *ʈ, ɖ* in word-final position only occur in loan words from adjoining Indo-Aryan languages; for example, *haaʈ* ‘market’ from Hindi *hāṭ*. The two stops *c, j* are phonetically realized as affricates [tʃ, dʒ] and occur in word-final position in recent loanwords; for example, *ãc* ‘flame’ from Hindi *āc*, *kagoj* ‘paper’ from *kāgaz/kāgoz* in Persian through adjoining Indo-Aryan. The sibilant *s* appears in all positions but occurs in word-final position only for loanwords; for example, *bes* ‘good’ from *bez* in Persian (through adjacent Indo-Aryan). The two liquids *r* and *l* can occur in all positions while another liquid *ʎ* can occur only in word-medial position. Two nasals *m* and *n* can appear in all positions. But among nasal consonants, *ŋ* occurs only in inter-vocalic position and *ŋ* occurs only in word-final position. The palatal nasal *ñ* appears only in one word; that is, *añ* ‘I (1st person singlar)’. *añ* is realized as [ajŋ] or [aiŋ] when used independently but as [aŋ] before genitive suffix *-aʔ, añaʔ* ‘my’. The frequency of *añ* is very high, so I recognize /ñ/ as a distinct phoneme. The two glides *w, y* never occur in initial position.

As far as the differences among dialects are concerned, the Hasada, on which my description is based, and Tamaria dialects have no aspirated stops while the Naguri and Kera dialects have them. The same goes for intervocalic *h*. Another difference among the dialects is that *ʎ* in Hasada corresponds (Table 3.3) to *ɖ* in others.

TABLE 3.2: MUNDARI CONSONANTS

	Consonant inventory					
	labial	dental	retrolex	palatal	velar	glottal
stop voiceless	<i>p</i>	<i>t</i>	<i>ʈ</i>	<i>c</i>	<i>k</i>	<i>ʔ</i>
stop voiced	<i>b</i>	<i>d</i>	<i>ɖ</i>	<i>j</i>	<i>g</i>	
fricative		<i>s</i>				<i>h</i>
nasal	<i>m</i>	<i>n</i>	<i>ŋ</i>	<i>ñ</i>	<i>ŋ</i>	
flap		<i>r</i>	<i>ʎ</i>			
lateral		<i>l</i>				
semivowel	<i>w</i>			<i>y</i>		

TABLE 3.3: MUNDARI DIALECT COMPARISON

Differences among dialects				
gloss	Hasada	Naguri	Tamaria	Kera
‘flower’	<i>baa</i>	<i>baha</i>	<i>baa</i>	<i>baha</i>
‘pole’	<i>kunja</i>	<i>khunja</i>	<i>kunja</i>	<i>khunja</i>
‘river’	<i>gaɽa</i>	<i>gaɖa</i>	<i>gaɖa</i>	<i>gaɖa</i>

2.2 Checked consonants

The most peculiar feature of consonants is the so-called checked consonant series. The stop phonemes /b/ and /d/ are realized as checked consonants in morpheme-final position.

The phonetic description of these checked consonants is as follows: first, the glottis is closed and the tongue or the lips simultaneously form an oral closure. The tongue or lip position is the same as that of the corresponding normal stops. Then the glottal closure is released, which is optionally followed by nasal release and voicing. Thus, [ʔb^m], [ʔd^m].

It is mentioned that nasal release after the glottal release is optional. In my observation, whether nasal release occurs or not is determined by the syllable structure of the morpheme. There is no nasal release in polysyllabic words but only in monosyllabic ones. For example,

lubl ‘hair’ [uʔb^m] but *ludubl* ‘to tell’ [uduʔb];

lridl ‘to grind’ [riʔd^m] but *lbiridl* ‘to stand up’ [biriʔd].

Gumperz (1957) considers checked stops in word-final position as allophones of the voiceless stops /p/ and /t/. But I treat these checked consonants as allophones of voiced stops /b/ and /d/ as Hoffmann did. The following morphophonological change is very clear: *ldub-al* (/a/: IND) ‘will sit’ [*duba*] not [*dupa*] and *lbirid-al* ‘will stand up’ [*birida*] not [*birita*].

Glottal stops are also regarded as checked consonants because these are allophones of /g/ and /y/. Glottal stops are followed by echo-vowel release in monosyllabic morpheme but never in polysyllabic morphemes. For instance, *lragl* ‘to call’ [raʔa], but *lracag* ‘to pull’ [racaʔ]; *lpoyl* ‘to rinse’ [pɔʔɛ], but *ltukuyl* ‘to saw’ [tukuiʔ].

Recent loanwords in Mundari allow morpheme-final *g*. Thus the following minimal pair can be found:

[nɛg] ‘religious feast’

[nɛʔɛ] ‘here take it’ (interjection)

I phonemicize the first word as *neg* and the second one as *neʔ*. In addition to this final /g/, I describe final /y/ as *iʔ* or *eʔ*.

2.3 Syllable structure and phonotactics

A phonological word in Mundari can be syllabified by a simple rule due to the simplicity of consonant clusters. There are only three types of syllable boundary, that is, (a) between two successive vowels, (b) between a vowel and a following consonant, and (c) between two consonants which form a consonant cluster.

The middle vowels in trisyllabic words can optionally be deleted. It seems to me that a phonological word in Mundari has a tendency to keep two morae. We have already seen the examples of monosyllabic words in the form of CV, CVb, CVd, CVʔ, and CVy in sections 2.1. and 2.2. Hence we discuss here only monosyllabic words of the remaining forms which may be counted as having two morae. In most C_1VC_2 words, C_2 is either a liquid or a nasal if it is not /b, d/ or /ʔ/. Denoting a mora boundary by %, we have *sim* [si%ɪm] ‘chicken’, *dul* [du%l] ‘to pour’. We have exceptions in the following loanwords, for example, *bes* [be%s] ‘good’, *soj* [sɔ%ɔɟ] ‘straight’. We consider that the consonants which occur as C_2 in all these cases constitute one mora.

We now list all the possible combination of C and V in the syllable structure of a phonological word, as seen in Table 3.4.

TABLE 3.4: MUNDARI WORD CV STRUCTURE

Monosyllabic	
VC	<i>ub</i> 'hair'
CV	<i>jo</i> 'fruit'
CVC	<i>jo?</i> 'to sweep'
Disyllabic	
V.V	<i>au</i> 'bring'
V.CV	<i>uku</i> 'hide'
V.CVC	<i>udub</i> 'tell'
CV.V	<i>bai</i> 'make'
CV.CV	<i>bulu</i> 'thigh'
CV.VC	<i>tain</i> 'live'
CV.CVC	<i>buluŋ</i> 'salt'
VC.CV	<i>eyga</i> 'mother'
VC.CVC	<i>umbul</i> 'shade'
CVC.CV	<i>doŋdo</i> 'fool'
CVC.CVC	<i>seŋgel</i> 'fire'
Trisyllabic	
V.V.CV	<i>auri</i> 'not yet'
V.V.CVC	<i>aosan</i> 'bring about an improvement' (EM)
V.CV.V	<i>apia</i> 'three'
V.CV.CV	<i>asadi</i> 'feel ennui'
V.CV.VC	<i>ale-a?</i> 'our' (plural and exclusive)
V.CV.CVC	<i>eperaŋ</i> 'quarrel'
V.CVC.CV	<i>aꞤandi</i> 'marriage'
V.CVC.CVC	<i>okonꞤo</i> 'lift the head whilst lying down' (EM)
CV.V.CV	<i>baila</i> 'deaf'
CV.V.CVC	<i>saitan</i> 'evil'
CV.CV.V	<i>balae</i> 'difficulty'
CV.CV.CV	<i>rasika</i> 'rejoice'
CV.CV.VC	<i>balae-n</i> 'worry'
CV.CV.CVC	<i>tutukun</i> 'cold'
CV.CVC.CV	<i>salangi</i> 'tall'
CV.CVC.CVC	<i>hasaŋgar</i> 'live coal'
VC.CV.V	<i>eyga-o</i> 'mother also'
VC.CV.CV	<i>oŋꞤoka</i> 'human sacrifice'
VC.CV.VC	<i>inku-a?</i> 'their'
VC.CV.CVC	<i>eyga-te?</i> 'his/her mother'
VC.CVC.CV	<i>enbanꞤa</i> 'rather' (EM)
CVC.CV.CV	<i>kumbuꞤu</i> 'thief'
CVC.CV.CVC	<i>nimnaŋge</i> 'enough (for food)'
CVC.CVC.CV	<i>pampalad</i> 'butterfly'

2.4 Intonation/stress

Mundari is not a tone language, unlike some Mon-Khmer languages in the other branch of Austroasiatic. Among Munda languages, Korku, which forms the North Munda language group along with Kherwarian languages (including Mundari, Santali, and Ho), has a tonal contrast (Zide 1960, 1966).³ Mundari does not have stress but pitch accent. Previous studies have mentioned only stress (Cook 1965: 100, Langendoen 1963: 14–15, N.K. Sinha 1975: 39).

Word accent in Mundari can be described in the following manner:

- (i) Word accent is not phonemic because it is predictable. In each phonological word an accent is assigned to only one syllable which is marked by a high pitch.
- (ii) The accent patterns are as follows:
 - (a) A monosyllabic phonological word is always accentuated;⁴ for example, *lbal* ‘flower’ [*baā*] or [*bāa*], *da*² ‘water’ [*dā²a*] or [*da²ā*].
 - (b) Accent is normally assigned to the second syllable in disyllabic words; for example, *bulū* ‘thigh’, *bulūŋ* ‘salt’, *seŋgēl* ‘fire’.
 - (c) Exceptions to this rule are the result of syllable weight. When the first syllable in a disyllabic word is heavier than the second syllable, the accent normally falls on the first syllable. When a syllable boundary is located between the nasal and homorganic stop sequences, the accent is assigned not to the first syllable but to the second syllable; for example, *sīrma* ‘sky, year’, *gōmke* ‘lord’ but *dondō* ‘to lift’, *campā* ‘a kind of flower’.
 - (d) Further, in a trisyllabic word, accent is never assigned to the first syllable even if the first syllable is the heaviest. The second syllable in a trisyllabic word cannot be accentuated unless the element in the last syllable is a suffix. An unaccented vowel in the second syllable can optionally be deleted; for example, *pampalād* ‘butterfly’, *arandī* ‘marriage’, *apī-a* ‘three’, *kumb(u)ṛū* ‘thief’.
 - (e) A quadrisyllabic word is divided into two bisyllabic phonological words. Accent is allocated to each phonological word; for example, *akādandā* ‘to feel astonished’.

As far as sentence intonation is concerned, the major role of intonation is to provide contrast between several sentence types which may be marked by the distinctive use of patterns of pitch. Furthermore, intonation functions as a signal of grammatical structure such as the marking of sentence boundaries. Moreover, intonation conveys paralinguistic features, that is, information about the speaker’s emotion, attitude, social background, etc. However we do not discuss these features in detail here. We mainly discuss the terminal contour.

Word accent also keeps its high level pitch at the sentence level. Besides high-level pitches, falling /↘/, rising /↗/ and falling-rising /↘↗/ contours play a major role in sentence intonation. As for level-pitches there are three; high¹, middle² and low³.

- (i) In an affirmative (declarative) sentence a falling intonation is allocated to the final syllable.

jom¹-ke³-d-a²-ko ↘
 eat-COMPL-TR-IND-3PL
 ‘They ate something.’

- (ii) In an interrogative sentence a falling–rising intonation is assigned to the final syllable of the sentence.

jom¹-ke³-d-a²-ko ↘ ↗
 eat-COMPL-TR-IND-3PL
 ‘Did they eat something?’

- (iii) When sentence particles occur in sentence-final position, the sentence intonations are different from the pattern of (ii).

- (a) In sentences with the question marker *ci*, *ci* is always high level pitch.

jom¹-ke³-d-a²-ko³ ↘ *ci¹*
 COMPL-TR-IND-3PL Q
 ‘Did they eat something?’

- (b) In sentences with question marker *ci* + negation marker *ka*, *ci* is not marked, but *ka* has a marked rising intonation.

jom¹-ke³-d-a²-ko³ *ci²* *ka* ↗
 ‘I wonder whether they ate something or not.’

- (c) A sentence with the negation marker *ka* is characterized by a falling contour.

jom¹-ke³-d-a²-ko² *ka* ↘
 ‘Did they eat something?’

- (iv) Negative and declarative sentences have the same pitch patterns as (i). Further, negative and interrogative sentences have the same pitch patterns as (ii).

ka¹=ko¹ *jom¹-ke³-d-a²* ↘
 ‘They didn’t eat something.’

ka¹=ko¹ *jom¹-ke³-d-a²* ↘ ↗
 ‘Didn’t they eat something?’

2.5 Morphophonology

Mundari has a kind of phonological restriction, which divides vowels into two distinct subsets, that is, high vowels and mid-vowels which do not co-occur within a morpheme. This phenomenon is well-known as vowel harmony.

These distinct subsets can be described by their distinctive features (Table 3.5). The subsets 1 and 2 cannot co-occur within a morpheme while 3 can co-occur with 1 and 2. This rule can be extended beyond a morpheme boundary within a

**TABLE 3.5: MUNDARI HARMONY
 FEATURE SETS**

1.	[+high]
2.	[-high, -low]
3.	[+low]

phonological word. For instance, *in* ‘that’ + *ko* (plural marker) = *inku* ‘those ones (animate)’; *ni* ‘this’ + *ko* (plural marker) = *niku* ‘these ones (animate)’.

Interestingly, not all phonological words can be generated by the rule. Hence it seems that only personal pronominal suffixes, including dual and plural suffixes, undergo the vowel harmony rule beyond a morpheme. Moreover, this rule can be adopted in a newly borrowed word. Thus *suri* < English *sorry*. In this case regressive assimilation has occurred, whereas progressive assimilation has occurred in the case of personal pronominal suffixes.

In the verbal morphology morphophonological changes frequently occur. The transitive marker *d* becomes *?* with the first and third person singular object. At the same time the completion aspect marker *ke* becomes *ki*. Thus,

- (1) *biꞥ coke=? jom-ja-?-i-a.*
 snake frog=3SG:SUBJ eat-INGR-TR-3SG:OBJ-IND
 ‘The snake is eating the frog.’
- (2) *biꞥ coke=? jom-ki-?-i-a.*
 snake frog=3SG:SUBJ eat-COMPL-TR-3SG:OBJ-IND
 ‘The snake ate the frog.’

3 MORPHOLOGY

3.1 Word class

For Mundari there has been a lengthy discussion of the difficulties in categorizing words into classes in terms of the traditional definitions of the parts of speech since Hoffmann (1903:xxi) declared the following:

Thus the same unchanged form is at the same time a Conjunction, an Adjective, a Pronoun, an Adverb, a Verb, and a Noun, or, to speak more precisely, it may become a Conjunction, an Adjective, and so on, but by itself alone it is none of them. It is simply a vague elastic word, capable of signifying, in a vague manner, several distinct concepts, that is of assuming a variety of functions.

This means that a prototypical lexical verb like *jom* ‘eat’ can be used as a noun without any morphological change, while a prototypical noun like *buru* ‘mountain’ can only be verbalized by attaching verbal endings. For example,

- (3) *buru=ko bai-ke-d-a.*
 mountain=3PL:SUBJ make-COMPL-TR-IND
 ‘They made the mountain.’
- (4) *saan=ko buru-ke-d-a.*
 firewood=3PL:SUBJ mountain-COMPL-TR-IND
 ‘They heaped up the firewood.’
- (5) *maꞥdi=ko jom-ke-d-a.*
 food=3PL:SUBJ eat-COMPL-TR-IND
 ‘They ate the food.’
- (6) *jom=ko nam-ke-d-a.*
 food=3PL:SUBJ get-COMPL-TR-IND
 ‘They got the food.’

In (3) *buru* is used as an argument, with the meaning ‘mountain’, while in (4) it is used as a two-place predicate with the meaning ‘heap up’. To illustrate the other direction of deployment, in (5) the word *jom* is used as a two-place predicate with the meaning ‘eat’, while in (6) it is used as an argument with the meaning ‘food’.

Nicholas Evans of Melbourne University and I published a paper titled ‘Mundari: the myth of language without word classes’ in *Linguistic Typology* in 2005. We introduced three criteria for establishing lack of word class distinctions, that is, equivalent combinatorics (members of both classes should have equivalent combinatorics), compositionality (the semantic results of using a member of one putative class in a constructional slot prototypically associated with the other putative class should be derivable through strict compositional principles) and bidirectionality (members of X should be deployable in the environments associated with Y, and members of Y should be deployable in the environments associated with X). Further, these three criteria should be exhaustive across the lexicon, that is, the same test should yield the same results for all lexemes in the putative class, not just for a few well-chosen ones. In our paper, we have seen that applying these three criteria decisively demonstrates that Mundari is not a monocategorical language.⁵

Thus I describe nouns and verbs as follows: nouns can be morphologically marked for certain grammatical categories such as noun class (animate/inanimate) and number (singular/dual/plural). Verbs can be marked for grammatical features such as aspect and mood. Second, they can take affixes for voice and transitivity which are related to grammatical functions such as subject and object. The verb agrees with subject and object in person and number which are marked by a personal suffix.

In addition to noun and verb we set up the following word classes:

Pronoun, adjective, postposition, adverb, numeral, conjunction, particle, interjection, and expressive.

3.2 Nominal morphology

3.2.1 Noun class and number

Nouns are divided into animate and inanimate in terms of a system of concord between subject, object, and verb. Animate nouns refer to human beings and animals. In fact most grammatically animate nouns denote human beings and animals. Besides them the following nouns are considered animate:

- (i) Heavenly bodies: *caŋdu?* ‘moon’, *siŋgi* ‘sun’, *ipil* ‘star’.

In relation to heavenly bodies the following verbs can be coded by animate marking: *gama* ‘to rain’, *hoyo* ‘to blow (the wind)’.

- (7) *gama-ja-d-a-e?*
rain-INGR-TR-IND-3SG:SUBJ
‘It is raining.’
- (8) *hoyo-le-d-a-e?*
wind-ANT-TR-IND-3SG:SUBJ
‘It had blown.’

- (ii) Supernatural beings: *boŋga* ‘spirit’, *siŋ boŋga* ‘supreme God’.

As for gender distinction, some animate nouns can be divided into female and male nouns, marked morphologically by the endings *i* and *a*, respectively under the influence of adjoining Indo-Aryan varieties. For instance,

kuṛi ‘woman’ *koṛa* ‘man’
kaki ‘aunt’ *kaka* ‘uncle’

In order to express a distinction of sex in Mundari, a following modifier is preposed to the noun: *eṅga* (originally means ‘mother’) is used for female while *saṅḍi* (originally means ‘cock’) is used for male. Thus,

eṅga seta ‘bitch’ *saṅḍi seta* ‘dog’
eṅga sim ‘hen’ *saṅḍi sim* ‘rooster’

As far as kinship terminology is concerned, *koṛa* may be used for male and *kuṛi* for female. For instance,

<i>hon-te</i> child-his/her	<i>koṛa</i> man	‘his/her son’
<i>hon-te</i> child-his/her	<i>kuṛi</i> woman	‘his/her daughter’
<i>boko-ñ koṛa</i> younger sister/brother-my	<i>koṛa</i> man	‘my younger brother’
<i>boko-ñ</i> younger sister/brother-my	<i>kuṛi</i> woman	‘my younger sister’

The number marking system for nouns in Mundari has three tiers, that is, singular-dual- plural. Singular is unmarked, and the dual and plural markers are *kin* and *ko*, respectively. Count nouns are marked for number irrespective of their animacy.

<i>hon</i> ‘a child’	<i>hon-kin</i> ‘two children’	<i>hon-ko</i> ‘children’
<i>ipil</i> ‘a star’	<i>ipil-kin</i> ‘two stars’	<i>ipil-ko</i> ‘stars’
<i>kitab</i> ‘a book’	<i>kitab-kin</i> ‘two books’	<i>kitab-ko</i> ‘books’
<i>lija?</i> ‘a piece of cloth’	<i>lija?-kin</i> ‘two pieces of cloth’	<i>lija?-ko</i> ‘pieces of cloth’

3.2.2 Case

Mundari NPs do not inflect for case: both the subject and object of a sentence are morphologically unmarked. The subject and object of a sentence are determined by word order. The unmarked word order is as follows: **S + O + Verb**.

Examples are given as in (9) and (10).

- (9) *pusi-kin seta-ko=kin hua-ke-d-ko-a.*
 cat-DL dog-PL=3DL:SUBJ bite-COMPL-TR-3PL:OBJ-IND
 ‘The two cats bit the dogs.’

- (10) *seta-ko pusi-kin=ko hua-ke-d-kin-a.*
 dog-PL cat-DL=3PL:SUBJ bite- COMPL-TR-3DL:OBJ-IND
 ‘The dogs bit the two cats.’

In addition to two arguments a postpositional phrase or adverb denoting location or time can be inserted in any position before verb. We can illustrate this in (11).

- (11) *seta?re seta-ko maŋdi=ko jom-ke-d-a.*
 morning-LOC dog-PL food=3PL:SUBJ eat-COMPL-TR-IND
 'In the morning the dogs ate the food.'

Case relations in Mundari are mainly marked by postpositions. Thus, instrumental is expressed by the postposition *te* following a noun or pronoun. Comitative is expressed by postposing *lo?* after a noun or pronoun. Benefactive is expressed by the postposition *naŋgen* following a noun or pronoun. There are several dialectal variants; *nagen/natin/naten*. Source is expressed by postposing *ate* in Hasada and Tamaría dialects or *ete* in Naguri and Kera dialects after a noun or pronoun.

The possessive is expressed by the suffixes *-a?*, *-rea?*, *-ra?*, and *-ren*. The possessive suffix *-a?* denotes alienable possession by an animate noun, while *-rea?/-ra?*, and *-ren* indicate alienable possession by an inanimate noun. The distinction between *-rea?/-ra?*, and *-ren* is made on account of the animacy of the head noun. We demonstrate it in Table 3.6 as follows:

TABLE 3.6: MUNDARI GENITIVE

Possessor	Possessed	
	Animate	Inanimate
animate	<i>-a?</i>	<i>-a?</i>
inanimate	<i>-ren</i>	<i>-rea?/-ra?</i>

3.2.3 Pronouns

Personal pronouns exhibit a 3 (First, Second, and Third) × 3 (Singular, Dual, and Plural) system (see Table 3.7).

We have found *abin* as a variant of second person dual and *akiŋ* as a variant of third person dual. I used this variant *akiŋ* in my previous works as Munda (1971) did. The form *akin*, however, is more common. Thus I, henceforth, use *akin* for the third person dual.

The possessive pronoun is formed by adding the genitive suffix *-a?* to a pronoun. Further, the Mundari equivalent of the independent possessive in English such as 'mine, yours', etc. is expressed by postposing the genitive suffix *-a?* to the possessive pronoun. The independent possessive is found only in the singular system.

TABLE 3.7: MUNDARI PRONOUNS

	Full form			Short form		
	Singular	Dual	Plural	Singular	Dual	Plural
1 (inclusive)	<i>añ</i>	<i>alaj</i>	<i>abu</i>	<i>-ñ</i>	<i>-laj</i>	<i>-bu</i>
1 (exclusive)		<i>aliŋ</i>	<i>ale</i>		<i>-liŋ</i>	<i>-le</i>
2	<i>am</i>	<i>aben</i>	<i>ape</i>	<i>-m</i>	<i>-ben</i>	<i>-pe</i>
3	<i>ae?</i>	<i>akin</i>	<i>ako</i>	<i>-e</i>	<i>-kin</i>	<i>-ko</i>

Thus, the possessive pronoun and independent pronoun can be described as follows (Table 3.8):

TABLE 3.8: MUNDARI POSSESSIVE PRONOUNS – I

	Possessive pronoun			Independent pronoun
	Singular	Dual	Plural	Singular
1 (inclusive)	<i>añ-aʔ</i>	<i>alaŋ-aʔ</i>	<i>abu-aʔ</i>	<i>añ-ag-aʔ</i>
1 (exclusive)		<i>aliŋ-aʔ</i>	<i>ale-aʔ</i>	
2	<i>am-aʔ</i>	<i>aben-aʔ</i>	<i>ape-aʔ</i>	<i>am-ag-aʔ</i>
3	<i>ay-aʔ</i>	<i>akin-aʔ</i>	<i>ako-aʔ</i>	<i>ay-ag-aʔ</i>

Furthermore, we have an old system of possessive pronouns (Table 3.9). This is performed by the reduced pronominal suffixes which follow the genitive marker *-ta*.

In colloquial Mundari, this system has been almost completely replaced by the construction, pronoun + genitive suffix *-aʔ*, though the language of poetry in Mundari still retains this system. For instance, we have *disum-tabu* (*disum* ‘country’) ‘our country’ in poetry, but *abu-aʔ disum* ‘our country’ in colloquial speech.

TABLE 3.9: MUNDARI POSSESSIVE PRONOUNS – II

	Possessive pronouns		
	Singular	Dual	Plural
1 (inclusive)	<i>ta-ñ</i>	<i>ta-laŋ</i>	<i>ta-bu</i>
1 (exclusive)		<i>ta-liŋ</i>	<i>ta-le</i>
2	<i>ta-m</i>	<i>ta-ben</i>	<i>ta-pe</i>
3	<i>ta-eʔ</i>	<i>ta-kin</i>	<i>ta-ko</i>

3.2.4 Demonstratives

Demonstratives in Mundari make a 3×2 contrast set (Proximate: Intermediate: Remote x marked: unmarked) as follows (Table 3.10):

The variants *ne-/ni-*, *e-/i-*, and *he-/hi-* are defined by the vowel harmony rule. The demonstratives have a rich derivational system. I illustrate them in Table 3.11.

Apart from the demonstratives, we have interrogative and indefinite pronouns as word classes. These three have a similar word process.

TABLE 3.10: MUNDARI DEMONSTRATIVES – I

	Proximate	Intermediate	Remote
Demonstrative bases			
Unmarked	<i>ne-/ni-</i>	<i>i-/e-</i>	<i>hi-/he-</i>
Marked	<i>na-</i>	<i>a-</i>	<i>ha-</i>

TABLE 3.11: MUNDARI DEMONSTRATIVES – II

	Proximate	Intermediate	Remote
Demonstrative adjectives			
Unmarked	<i>ne-/ni-</i>	<i>in-/len-</i>	<i>hin-/hen-</i>
Marked	<i>na-</i>	<i>an-</i>	<i>han-</i>
Demonstrative pronouns (animate)			
Singular			
Unmarked	<i>niⁱ?</i>	<i>iniⁱ?</i>	<i>hiniⁱ?</i>
Marked	<i>naiⁱ?</i>	<i>aniⁱ?</i>	<i>haniⁱ?</i>
Dual			
Unmarked	<i>nikin</i>	<i>inkin</i>	<i>hinkin</i>
Marked	<i>nakin</i>	<i>akin</i>	<i>hankin</i>
Plural			
Unmarked	<i>niku</i>	<i>inku</i>	<i>hinku</i>
Marked	<i>nako</i>	<i>ako</i>	<i>hanko</i>
Demonstrative pronouns (inanimate)			
Unmarked	<i>nea</i>	<i>ena</i>	<i>hena</i>
Marked	<i>naya</i>	<i>ana</i>	<i>hana</i>
Definite demonstratives			
Adjectivals			
Unmarked	<i>nimin/nimun</i> 'this much'	<i>imin/immun</i> 'that much'	
Marked	<i>namin/namun</i> 'this much' (more than one's expectation)	<i>amin/amun</i> 'that much'	
Nominals			
Unmarked	<i>niminay/nimunay/ niminuy/nimunuy</i> 'this much'	<i>iminay/imunay/ iminuy/imunuy</i> 'that much'	
Marked	<i>naminay/namunay/ naminuy/namunuy</i> 'this much' (more than one's expectation)	<i>aminay/amunay/ aminuy/amunuy</i> 'that much'	
Emphatic			
Unmarked	<i>nimpuruy/nimpiruy/ nimpiray/nimpinay</i> 'this so much'	<i>impuruy/impiruy/ impiray/impinay</i> 'that so much'	
Marked	<i>nampuruy/nampiruy/ nampiray/nampinay</i> 'this so much' (more than one's expectation)	<i>ampuruy/ampiruy/ ampiray/ampinay</i> 'that so much'	
Demonstrative adverbials			
Adverbs of place			
Unmarked	<i>nere, nete, neate</i> 'here, there, from here'	<i>enre, ente, enate</i> 'there, thither, from there'	<i>henre, hente, henate</i> 'yonder, to yonder, from yonder'
Marked	<i>nare, nate, naate</i> 'here, hither, from here'	<i>enre, ente, enate</i> 'there, thither, from there'	<i>hanre, hante, hanate</i> 'yonder, to yonder, from yonder'

(Table 3.11 continued)

TABLE 3.11: CONTINUED

	Proximate	Intermediate	Remote
Adverbs of time			
Unmarked	<i>nimtay/nimtuŋ</i> 'this time'	<i>imtay/imtuŋ</i> 'that time'	
Marked	<i>namtay/namtuŋ</i> 'this time' (against one's expectation)	<i>amtay/amtuŋ</i> 'that time' (against one's expectation)	
Adverbs of manner			
Unmarked	<i>neka</i> 'like this'	<i>enka</i> 'like that'	<i>henka</i> 'like yonder'
Marked	<i>naka</i> 'like this' (against one's expectation)	<i>anka</i> 'like that'	<i>hanka</i> 'like yonder'

There are four interrogative bases (Table 3.12):

TABLE 3.12: MUNDARI INTERROGATIVES

<i>oko</i>	'which'
<i>ca</i> and <i>ci</i>	'what'
<i>ci-lika</i>	'how'

The interrogative *ci-lika* can be derived from the other interrogative base *ci* plus *leka* 'like'. The interrogatives *ca* and *ci* may be related to the demonstrative bases *a* and *i*. These four interrogatives are derivational bases. These derivational formations are identical to the demonstrative ones. The interrogative bases *oko* and *cilika* can be used independently while *ca* and *ci* can act only as a derivational base. The interrogative *oko* functions as a modifier whereas *cilika* as an adverb of manner.

Mundari has three indefinite bases (Table 3.13):

TABLE 3.13: MUNDARI INDEFINITES

<i>oko</i>	'some'
<i>ja</i>	'any'
<i>jeta</i>	'any'

The distinction between interrogative *oko* and indefinite *oko* is somewhat dependent on the context. They, however, have at least one distinguishing syntactic criterion. The indefinite *oko* can be followed by the topic marker *do* but the interrogative *oko* cannot. For instance,

- (12) *oko-e hiju?-aka-n-a.*
 INTER-3SG come-CONT-ITR-IND
 'Who has come?'

- (13) *oko-e do hiju?-aka-n-a.*
 INDEF-3SG TOP COME-CONT-ITR-IND
 ‘Someone has come (but not all).’

The indefinites *ja* and *jeta*, which are nearly synonyms, have the same syntactic function but the indefinite *jete* implying emphatic is more frequently used with the negative. We should pay attention to the semantic distinction between *a* (more than one’s expectation) and *ile* (less than one’s expectation) here. I think that the semantic feature ‘negative’ may be related to ‘less than one’s expectation’.

Demonstrative, Interrogative and Indefinite pronouns have a similar word formation process. We summarize it here in the following way:

- (i) Adjectivals
 DB(=Demonstrative bases), INTB(=Interrogative bases), INDB(=Indefinite bases) + *-n*. For example, *ca-n* ‘what a kind of’, *ja-n* ‘any kind of’.
- (ii) Pronoun (Animate)
 DB, INTB, INDB + *(-n-)* + *-el-i?* (for singular), *-kin* (for dual), *-kol-ku* (for plural). For example, *oko-kin* ‘who (dual)’, *ja-n-ku* ‘any persons’.
- (iii) Inanimate
 DB, INTB, INDB + *(-n-)* + *-a*. For example, *ca-n-a* ‘which things’, *ja-n-a* ‘any things’.
- (iv) Possessive pronoun (Animate)
 DB, INTB, INDB + *(-n-)* + *-el-i?* (for singular), *-kin* (for dual), *-kol-ku* (for plural) + *a?* (Genitive). For example, *ca-n-kin-a?* ‘whose (dual)’, *jeta-n-ku-a?* ‘of any persons’.
- (v) Possessive pronoun (Inanimate)
 DB, INTB, INDB + *(-n-)* + *-a* + *-rea?l-ra?* (Genitive). For example, *oko-rea?* ‘of which thing’, *ja-n-a-ra?* ‘of any things.’
- (vi) Definites
 DB, INTB, INDB + *(-i-)* + *-min/mun-* + *(aŋ/luŋ)*. For example, *ci-min-aŋ* ‘how much’.
- (vii) Emphatic definites
 DB, INTB, INDB + *(-i-)* + *mpuruŋ/mpiruŋ/mpiraŋ/mpinaŋ/mpinuŋ/mpunuŋ*. For example, *ci-mpuruŋ* ‘how much exactly’.
- (viii) Adverbs of time
 DB, INTB, INDB + *(-i-)* + *-mtaŋ/mtuŋ*. For example, *ci-mtaŋ* ‘when.’
- (ix) Adverbs of place
 DB, INTB, INDB + *(-n-)* + *(sa?, ta?)* + *re, te, ate*. For example, *oko-sa?-te* ‘to which side’.
- (x) Adverbs of manner
 DB, INTB, INDB + *(-i-)* + *-lekalka*. For example, *ja-leka* ‘any ways’.

- (a) DB = Demonstrative bases
- | | Proximate | Intermediate | Remote |
|----------|-------------|--------------|-------------|
| unmarked | <i>neli</i> | <i>eli</i> | <i>heli</i> |
| marked | <i>na</i> | <i>a</i> | <i>ha</i> |
- (b) INTB = Interrogative bases
- ca-lci-* ‘what’
oko- ‘which’
- (c) INDB = Indefinite bases
- oko-* ‘some’
ja- ‘any’
jeta- ‘any’
- (d) Semantic features
- ile* unmarked or less than expected
a more than expected

3.2.5 Numerals

Table 3.14 presents Cardinal numerals.

As we have seen below, Mundari has a vigesimal counting system. According to Norman Zide (1978:1), ‘presumably Proto-Austroasiatic as well as old Indo-Aryan and Dravidian (old and modern) lacked vigesimal counting systems, but both Munda and modern Indo-Aryan use them. Whether the Indo-Aryan vigesimal systems “come from Munda” – as has been claimed – is questionable’.

The following short forms are used for the modifier of a head noun:

miad/mod ‘one’, *bar* ‘two’, *api* ‘three’, *upun* ‘four’, *moŋe* ‘five’, *turui* ‘six’, *ee* ‘seven’, *iral* ‘eight’, *are* ‘nine’, *gel* ‘ten’.

The counting forms consist of the addition of *-ialea* in postconsonantal position or *a* in post-vocalic position to the short forms, as is shown below.

TABLE 3.14: MUNDARI NUMERALS

<i>miad/moyod</i>	‘one’
<i>bar-ia</i>	‘two’
<i>api-a</i>	‘three’
<i>upun-ia</i>	‘four’
<i>moŋe-a</i>	‘five’
<i>turui-alturi-a</i>	‘six’
<i>ee-a</i>	‘seven’
<i>iral-ia</i>	‘eight’
<i>are-a</i>	‘nine’
<i>gel-ea</i>	‘ten’
<i>gel miad/moyod</i>	‘10+1=11’
<i>mod/mid hisi</i>	‘1×20=20’
<i>mod/mid hisi miad/moyod</i>	‘1×20+1=21’
<i>bar hisi</i>	‘2×20=40’
<i>api hisi</i>	‘3×20=60’
<i>moŋe hisi or mod/mid sau</i>	‘5×20=100 or 1×100=100’

Distributive numerals are expressed by reduplication of the cardinal numerals. Distributive forms are a partial reduplication of cardinal forms for the numeral forms for 'one' to 'six' and 'ten', while complete reduplication is required for the numeral forms 'seven', 'eight', and 'nine'. These coincide with the distributive form of Santali numerals.

<i>mi-miyad</i> 'one each'	<i>tu-turi-a</i> 'six each'
<i>ba-bar-ia</i> 'two each'	<i>ee-a ee-a</i> 'seven each'
<i>ap-api-a</i> 'three each'	<i>iral-ia iral-ia</i> 'eight each'
<i>up-upun-ia</i> 'four each'	<i>are-a are-a</i> 'nine each'
<i>mo-moŋe-a</i> 'five each'	<i>ge-gel-ea</i> 'ten each'

Ordinal numeral forms are as follows:

<i>sida</i> 'first'
<i>eŋa?</i> 'second'

The following variant forms are notable:

- (i) *milmo* 'one' in *mi-salmo-sa* 'once' (c.f. *bar-sa* 'twice', *api-sa* 'three times', etc.)
- (ii) *mu* 'one' in *mu-sij* 'one day' (c.f. *bar-sij* 'two days', *api-ma* 'three days', *upun-ma* 'four days', etc.)

As Emeneau (1956/1980: 115) has pointed out, numeral classifiers are an Indian areal feature. Mundari uses *hoŋo* 'person', *oŋa?* 'house', *boo?* 'head' as classifiers. Thus,

<i>api</i>	<i>hoŋo</i>	<i>hon-ko</i>	'three children'
three	person	child-PL	

The word *janljon* (from Indo-Aryan) is also currently used in Mundari. However, *janljon* always co-occurs with Indo-Aryan numerals. For example,

<i>tin</i>	<i>janljon</i>	<i>hon-ko</i>
three	Numeral Classifier	child-PL
'three children'		

3.2.6 Postpositions

Postpositions can be placed in a postnominal position and can form a postpositional phrase which may be used as a complement standing in a functional relationship with the verb.

The main postpositions can be divided in the following way:

- (i) *re* 'in' *te* 'to, by' *atelete* 'from'
- (ii) *sa?* 'on the side' *ta?* 'vicinity' *lo?* 'with'
- (iii) *ko* 'approximate'

The postposition *ko* does not appear independently but with (i) or (ii) and following (i) and (ii). Compound postpositions can be formed in the following order:

- (ii) + (iii) + (i)

The process is detailed as follows:

	<i>saʔ</i> ‘on the side’	<i>taʔ</i> ‘vicinity’	<i>loʔ</i> ‘with’
<i>re</i> ‘in’	<i>saʔ-re</i> ‘in the side’	<i>taʔ-re</i> ‘in the place’	<i>loʔ-re</i> ‘with’
<i>te</i> ‘to’	<i>saʔ-te</i> ‘to the side’	<i>taʔ-te</i> ‘to the place’	
<i>te</i> ‘by’	<i>saʔ-te</i> ‘by the side’	<i>taʔ-te</i> ‘by the place’	<i>loʔ-te</i> ‘along’
<i>ate</i> ‘from’	<i>sag-ate</i> ‘from the side’	<i>tag-ate</i> ‘from the place’	
<i>ko-re</i> ‘near in’	<i>saʔ-ko-re</i>	<i>taʔ-ko-re</i>	<i>loʔ-ko-re</i> ‘with’
	‘near the side’	‘near the place’	‘shortly before’
<i>ko-te</i> ‘near to’	<i>saʔ-ko-te</i>	<i>taʔ-ko-te</i>	<i>loʔ-ko-te</i> ‘along’
	‘near to the side’	‘near to the side’	‘shortly after’

Other postpositions will be illustrated below.

(i) *nanɣen* ‘for’.

The semantic functions of this postposition are benefactive and purpose.

(ii) *jaked, habiʔ/lhamiʔ* ‘until, up to’.

This may refer to location as in the following example:

- (14) *Ranci-jaked (habiʔ/lhamiʔ)=ko sen-ke-n-a.*
 Ranci-up to=3SG:SUBJ go-COMPL-ITR-IND
 ‘They went up to Ranchi.’

It may also refer to time as in the following instance:

- (15) *sombar-jaked (habiʔ/lhamiʔ) Ranci-re=ko tai-n-a.*
 Monday-until Ranci-LOC=3PL:SUBJ stay-ITR-IND
 ‘They will stay at Ranchi until Monday.’

Moreover, it may refer to quantity as follows:

- (16) *ne keʔa api-hisi-jaked (habiʔ/lhamiʔ)=e gonon-ɔʔ-a.*
 this buffalo three-twenty-up to=3SG:SUBJ COST-PASS-IND
 ‘This buffalo will cost up to 60 rupees.’

3.2.7 Derivation

We have already discussed the lexical semantic ambiguity involving the distinction between nouns and verbs in section 3.0. This distinction should be maintained because of the fact that a noun can be derived from a verb by a morphological process, namely, infixation:

(C)VC(VC)	→	(C)V<nV>C(VC)
<i>dub</i> ‘to sit’	→	<i>du<nu>b</i> ‘a meeting’
<i>rakab</i> ‘to rise’	→	<i>ra<na>kab</i> ‘a slope’
<i>ol</i> ‘to write’	→	<i>o<no>l</i> ‘the writing’
<i>eʔeʔ</i> ‘to begin’	→	<i>e<ne>ʔeʔ</i> ‘an origin’
<i>tebaʔ</i> ‘to arrive’	→	<i>te<ne>baʔ</i> ‘arrival’
<i>tukuiʔ</i> ‘to sew’	→	<i>tu<nu>kuiʔ</i> ‘the sewing’
<i>tagoeʔ</i> ‘to chew’	→	<i>ta<na>goeʔ</i> ‘the molar teeth’

The other nominalizing affixation is the possessive suffix *-aʔ* following a postpositional phrase which consists of a verb and the instrumental postposition *-te*. Thus,

- ol* ‘to write’ → *ol-te-aʔ* ‘an instrument for writing: pen, pencil, etc.’
jom ‘to eat’ → *jom-te-aʔ* ‘an instrument for eating: spoon, chopstick, etc.’
dub ‘to sit’ → *dub-te-aʔ* ‘an instrument for sitting: chair, stool, etc.’

3.2.8 Adjective

The distinction between verbs and adjectives is problematic. It seems to me that Mundari is a typical adjectival-verb language; that is, ‘the usual verbal equivalent of a predicate adjective is a predicate verb in a non-relative construction while the usual verbal equivalent of a modifying adjective is a verb in a relative construction’ (Schachter 1985: 18–19). Thus *marəŋ* ‘big, great’ when used predicatively can be marked for aspect, mood, voice and (in)transitivity like a predicate verb. It might be said that one word class covers two semantically different classes, that is, adjectives and verbs. For example,

- (17) *en marəŋ hoʔo*
 that great person
 that great person’ or
 ‘that person who will be great’

en jom hoʔo
 that eat person
 ‘that person who will eat’

- (18) *en hoʔo marəŋ-a.*
 that person great-IND
 ‘That person will be great.’

en hoʔo jom-a.
 that person eat-IND
 ‘That person will eat.’

In my book I wrote ‘the definition of the adjective is rather notional’ (Osada 1992: 123). I have, however, introduced one morphological criterion: infixation possibilities distinguish verbs and adjectives. There is an infix $\langle pV \rangle$ as a reciprocal marker which is illustrated in section 3.2.7. A number of adjectives take a formally identical marker but they do not acquire the reciprocal meaning ‘each other’, which is but natural as these adjectives are one-place words. Instead, they acquire the intensive meaning ‘very’. This is the sole test justifying the setting up of adjectives as a distinct word class. It is interesting to note that a head noun modified by an adjective with the infix $\langle pV \rangle$ takes the plural marker though it may be either singular or plural in meaning:

- (19) *en marəŋ hoʔo*
 that great person
 ‘that great person’

- (20) *en ma<pa>raŋ hoʔo-ko*
 that great<INTENS> person-PL
 (a) ‘that very great person’
 (b) ‘those very great persons’

At least seven adjectives denoting size, shape, and the like take the intensifying infix, because of semantic limitations on intensification. We contain seven adjectives in Mundari in terms of this derivation:

- (21) *maraj* ‘big, great’ → *ma<pa>raj* ‘very big, great’
huɽij ‘small’ → *hu<pu>ɽij* ‘very small’
jilij ‘long’ → *ji<pi>lij* ‘very long’
salangi ‘tall’ → *sa<pa>langi* ‘very tall’
dijgae ‘short’ → *dj<pi>ijgae* ‘very short’
cakar ‘wide’ → *ca<pa>kar* ‘very wide’
moɽo ‘fat’ → *mo<po>ɽo* ‘very fat’

It is noteworthy that some pronouns: demonstrative (e.g. *naminuɽ*, etc. ‘this much more than one expects’), interrogative (*cimunuɽ*, etc. ‘how much’), and indefinite (*jaimunuɽ*, etc. ‘to any extent, whatever be’), when modifying an intensive adjective also acquire the infix (by way of a kind of ‘pleonastic agreement’):

- (22) *naminuɽ maraj hoɽo-ko ka=ñ lel-aka-d-ko-a.*
 this much big person-PL NEG=1SG:SUBJ see-CONT-TR-3PL:OBJ-IND
 ‘I have never seen such big person(s).’
- (23) *nam<p>inuɽ ma<pa>raj hoɽo-ko ka=ñ*
 this much<INTENS> big<INTENS> person-PL NEG=1SG:SUBJ
lel-aka-d-ko-a.
 see-CONT-TR-3PL:OBJ-IND
 ‘I have never seen all so big (but more than one’s expectation) person(s).’
- (24) *cim<p>unuɽ hu<pu>ɽij tai-ke-n-a.*
 how much<INTENS> small<INTENS> remain-COMPL-ITR-IND
 ‘How small was it?’
- (25) *jaim<p>unuɽ ji<pi>lij-re-o ka=ñ suku-a.*
 whatever<INTENS> long<INTENS>LOC-also NEG=1SG:SUBJ like-IND
 ‘Anything that is too long I don’t like.’

3.2.9 Adverbials

Adverbs can function independently as verbal complements. Expressions for adverbs of location can be made by constructing postpositional phrases. The number of adverbs is rather small.

Now I illustrate the following adverbs of time:

- gapa* ‘tomorrow’
tisiɽ ‘today’
meyaj ‘the day after tomorrow’
honder ‘some days ago’
naa? ‘now’

<i>nimir</i>	‘nowaday’
<i>kalom</i>	‘next year’
<i>satom</i>	‘two years later’

In addition to the adverbs given above, adverbs of time can be expressed by a postpositional phrase. For instance, *setaʔ-re* ‘in the morning’, *sombar-ate* ‘from Monday’, *etwar-jaked* ‘until Sunday’.

Adverbs of location must always be expressed by a postpositional phrase. For example, *Ranci-re* ‘in Ranchi’, *oʔaʔ-te* ‘to the house’, *hatu-ate* ‘from the village’. Several local semantic functions are expressed mainly by postpositional phrases.

The postposition *te* (instrumental) can be used for adverbs of manner; for example, *rasika-te* ‘joyfully’, *mani-te* ‘slowly’, *eskar-te* ‘alone’.

3.3 Verbal morphology

Langendoen (1967) tried to describe Mundari verb conjugation based on Chomsky’s standard theory. He confessed the following in a straightforward manner.

The reader who is convinced of the efficacy of morpheme order charts for displaying the facts of a complex morphological system are advised to attempt to formulate such a chart for the Mundari data presented in this paper. I am reasonably convinced that no such formulation will be a match for the generative-transformational statement given here (inadequate as it is at various points) for displaying the intricate interconnections among the various patterns found in the Mundari verb conjugation. And really it must be admitted that the morphology of the Mundari verbal form is not nearly as complex as that of many languages. (Langendoen 1967:57)

I aim to present in this chapter not a theory-oriented but a data-oriented description. The basic verbal structure in Mundari may be described in terms of an order element formula as given in Table 3.15.

A verbal base is formed by affixing to a verbal stem. Verbal bases can be simple or complex; complex bases are formed by reduplication or serializing of the verbal stem. Verbal stems may be either transitive, or intransitive, or labile (i.e. transitive-intransitive, like the English ‘break’). Intransitive verbs are few in number (here belong, *inuy* ‘to play’, *ay* ‘to dawn’, *donʔo* ‘to be foolish’, and the like). The intransitive or transitive use of labile verbs is distinguished by means of intransitive and transitive suffixes, *-n* and *-d*, respectively.

TABLE 3.15: MUNDARI VERB TEMPLATE

Verb base + (Aspect marker)	+ (<i>n</i>)	+ <i>a</i> (=Subj)
	+ (<i>d</i> ʔ) (+ Obj)	

Notes

- a *n*: intransitive marker, *d*: transitive marker, with variant ʔ.
- b the transitive marker and the intransitive marker only appear when an aspect marker is present.
- c the suffix *-a* (indicative marker) is used to indicate the main verb of the clause excepting certain imperative forms.

3.3.1 Subject

The subject and object agreement element can be marked only when the subject NP and object NP are classified as animate nouns. Table 3.16 lists personal pronominal suffixes which are used for subject–object agreement:

TABLE 3.16: MUNDARI AGREEMENT MARKERS

	SG.	DL	PL
1st			
Inclusive	<i>-ñ</i>	<i>-lay</i>	<i>-bu</i>
Exclusive		<i>-lij</i>	<i>-le</i>
2nd	<i>-m</i>	<i>-ben</i>	<i>-pe</i>
3rd	<i>-e/-il-e?/-i?</i>	<i>-kin</i>	<i>-ko</i>

The same forms are used for subject, object and indirect object, but occupy different slots. The subject agreement element is attached either to the end of the verb or as a clitic to the preverbal NP, which may be not only the subject but also a non-subject. For example,

- (26) *hon-ko=ko dub-aka-n-a.*
 child-PL=3PL:SUBJ sit-CONT-ITR-IND
 ‘The children have sat.’
- (27) *maṅḍi=ñ jom-ta-n-a.*
 food=-1SG:SUBJ eat-PROG-TR-IND
 ‘I am eating the food.’
- (28) *gapa=ko senog-a.*
 tomorrow=3PL:SUBJ go-IND
 ‘They will go tomorrow.’

Munda people belonging to the younger generation tend to place the subject agreement element at the end of the verb.

As I have pointed out earlier, the subject agreement element can be marked only when the subject NPs are classified as animate nouns. In addition to this principle, transitive subject NPs are basically animate nouns. Thus, the following sentence is not grammatical because the transitive subject is inanimate.

- (29) **maṅḍi hon-ko bisi-ja-d-ko-a.*
 food child-PL poison-INGR-TR-3PL:OBJ-IND
 ‘The food has poisoned the children.’

Instead, it is placed in instrumental function by adding the instrumental postposition *-te*, in a type of passive construction:

- (30) *maṅḍi-te hon-ko=ko bisi-ja-n-a.*
 food-by child-PL=3PL:SUBJ poison-INGR-ITR-IND
 ‘The children have been poisoned by the food.’

Some inanimate nouns denoting natural objects, which are capable of automatic locomotion, can occupy the transitive subject slot. In that case, the subject agreement

element gets marked on the verb like an animate transitive subject NP. This is called animatization. For example,

- (31) *gaṛa buru=i? bai-ke-d-a.*
 river mountain=3SG:SUBJ make-COMPL-TR-IND
 ‘The river made the mountain (by carrying the sands).’
- (32) *gaṛa hon-ko=e? idi-ke-d-ko-a.*
 river child-PL=3SG:SUBJ take-COMPL-TR-3PL:OBJ-IND
 ‘The river took away the children.’

Among inanimate nouns, only natural forces, for example, *hoyo* ‘wind’, *gama* ‘rain’, etc. can be animatized.

It is very easy to identify animate transitive subject NPs subjects because an animate subject is always marked on the verb. Apart from this principle, subjecthood and objecthood can be defined by a syntactic test, which I discuss in section 3.2.2.

3.3.2 Object types

As shown earlier, the same pronominal suffix is used for subject and object agreement. But the object agreement element occupies the slot just before the indicative marker *-a* in the indicative sentence or the slot just before the second person pronominal suffix in the imperative sentence or the slot just before the optative marker *-ka-*. For instance,

- (33) *Soma hon-ko=e? lel-ko-a.*
 Soma child-PL=3SG:SUBJ see-3PL:OBJ-IND
 ‘Soma will see (take care) the children.’
- (34) *Soma hon-ko lel-ko-me.*
 Soma child-PL see-3PL:OBJ-2SG
 ‘Soma, please see the children.’
- (35) *Soma hon-ko lel-ko-ka-e?*
 Soma child-PL see-3PL:OBJ-OPT-3SG:SUBJ
 ‘May Soma see the children.’

The above sentences are unmarked for aspect marker and transitive/intransitive marker. We can differentiate transitive sentences from intransitive sentence by the transitive/intransitive marker, as in the following:

- (36) *Soma hon-ko=e? dub-ke-d-ko-a.*
 Soma child-PL=3SG:SUBJ sit-COMPL-TR-3PL:OBJ-IND
 ‘Soma made the children sit.’
- (37) *hon-ko ote-re=ko dub-ke-n-a.*
 child-PL ground-LOC=3PL:SUBJ sit-COMPL-ITR-IND
 ‘The children sat on the ground.’
- (38) *pulis-ko kumbuṛu-kin=ko sab-ja-d-kin-a.*
 police-PL thief-DL =3PL:SUBJ catch-INGR-TR-3DL:OBJ-IND
 ‘The policemen have caught the two thieves.’

- (39) *kumbuꞗu-kin hola=kin sab-ja-n-a.*
 thief-DL yesterday=3DL:SUBJ catch-INGR-ITR-IND
 'Two thieves have been caught yesterday.'

Note that the verb *dub* 'to sit' acquires causativity in the transitive sentence, whereas the verb *sab* 'to catch' acquires a passive sense in the intransitive sentence.

In ditransitive sentences, the benefactive marker *-a* is used in the following:

- (40) *am seta-ko=ñ om-a-m-ta-n-a.*
 2SG dog-PL = 1SG:SUBJ give-BEN-2SG-PROG-ITR-IND
 'I am giving the dogs to you.'
- (41) *am seta-ko=ñ om-ke-d-ko-a.*
 2SG dog-PL = 1SG:SUBJ give-COMPL-TR-3PL:OBJ-IND
 'I gave the dogs to you.'

It is very interesting that only one object, for example, *am* 'you' or *seta-ko* 'dogs' can be cross-referenced. Thus the sentence (42a) is ungrammatical. Further, if you want to encode the beneficiary on the verb in the completive sentence the completion aspect marker *ke* should change to the 'cislocative' or suspended aspect marker *a* as in (42b). Thus,

- (42) (a) **am seta-ko=ñ om-a-m-ke-d-ko-a.*
 2SG dog-PL = 1SG:SUBJ give-BEN-2SG-COMPL-TR-3PL:OBJ-IND
 'I gave the dogs to you.'
- (b) *am seta-ko=ñ om-a-d-me-a.*
 2SG dog-P = 1SG:SUBJ give-SUS-TR-2SG:OBJ -IND
 'I gave the dogs to you.'

As I mentioned in the previous section, inanimate NPs cannot normally occupy the transitive subject slot. But with some verbs, inanimate NPs can do this. The semantic range of these verbs is restricted to the following:⁶

- (a) Sensory and mental experiences
 (b) Emotional experiences.
 (c) Physical and biological experiences.

I call these experiential verbs. This semantic range almost corresponds interestingly with dative subject predicates in Indo-Aryan (Klaiman 1986). There are two types of sentence in experiential verbal constructions in Mundari, as follows, in one, the experiencer is the subject (43, 45) while in the others it is the object (44, 46):

- (43) *balbal-te=ñ sowan-ta-n-a.*
 sweat-by=1SG:SUBJ smell-PROG-ITR-IND
 'I am experiencing a smell of sweat.'
- (44) *ne baa maja sowan-ja-ʔ-ñ-a.*
 this flower good smell-INGR-TR-1SG:OBJ -IND
 'This flower has made me experience a good smell.'

(45) *sida samae susun-te=ko rasika-le-n-a.*
 previous time dance-by=3PL:SUBJ be.joyful-ANT-ITR-IND
 ‘In the previous time they had experienced joy through dancing.’

(46) *susun bese rasika-ke-d-ko-a.*
 dance very be.joyful-COMPL-TR-3PL:OBJ -IND
 ‘The dance made them experience joy.’

Recall that transitive subject NPs are basically animate nouns. If we assign the NP to grammatical relations by the morphological marking system, as for (44) and (46), the subjects are either nothing, or inanimate nouns and the objects are *ñ* ‘me’ in (44) and *ko* ‘them’ in (46). As I have discussed the grammatical relations above, these criteria will be kept throughout this chapter. I give here a new analysis of experiential verbal constructions.

The new analysis adopts the term ‘experiencer’ and ‘stimulus’. The experiencer denotes the human experiencer of sensory, mental, emotional, physical, and biological states expressed by the experiential verbs, whereas the stimulus is the source or cause of experience.

Further, I adopt the notion of ‘experiencer-subject’ from Croft (1991,1993) for a typological analysis of mental verbs, where experiential verbs assign the experiencer to the subject position in (43) and (45). In (44) and (46), on the other hand, the experiencer is assigned to the object position. I, therefore, consider it as the experiencer–object construction. This analysis is useful, because it enables us to keep the criteria for subject–object assignment.

Now I give a new analysis of the experiential construction. In general I regard an experiential verb as an intransitive verb (see (43) and (45)). Then I consider the same experiential verb as a causativized form of an intransitive verb in the object-experiencer construction (see (44) and (46)). This analysis fits Croft (1991:215)’s cross-linguistic findings that ‘experiencer-object verbs are causative’.

The stimulus occurs with the instrumental postposition *-te* in the experiencer–subject construction shown in (43) and (45), and as a subject in the experiencer–object construction shown in (44) and (46). Are the NPs *ne baa* ‘this flower’ in (44) and *susun* ‘dance’ in (46) really subjects? And are the NPs *ñ* in (44) and *ko* in (46) really objects? We can make a syntactic test for subjecthood and objecthood by using relativization, as in the following:

(47) *maja sowan-le-n-baa goso?-ja-n-a.*
 good smell-ANT-ITR-flower wither-INGR-ITR-IND
 ‘The good-smelling flower has withered.’

(48) *ne baa maja sowan-le-d-(hoꝝo)=ko seno?-ja-n-a.*
 this flower good smell-ANT-TR-(people)=3PL:SUBJ GO-INGR-ITR-IND
 ‘The people whom this flower made experience a good smell have gone.’

(49) *bese rasika-ke-n-susun nimir caba-ja-n-a.*
 very be.joyful-COMPL-ITR-dance recently finish-INGR-ITR-IND
 ‘The very joyful dance has finished recently.’

(50) *susun bese rasika-ke-d-(hoꝝo)=ko maraꝝ-ja-n-a.*
 dance very be.joyful-COMPL-TR-(people)=3PL:SUBJ grow-INGR-ITR-IND
 ‘The people whom the dance made joyful have grown up.’

As the intransitive marker *-n-* appears in (47) and (49), the NPs *baa* and *susun* are subjects. On the other hand, the transitive marker *-d-* occurs in (48) and (50) the NPs (*hofo*)-*ko* are objects. My new analysis can be supported by this syntactic test of relativization.

I add an important note here. The stimulus is always an inanimate noun. In other words, transitive subject NPs in the object–experiencer construction are inanimate nouns. For instance, the following sentences are not acceptable:

- (51) **ne kuři maja=e?* *sowan-ja-?-ñ-a.*
 this girl good=3SG:SUBJ smell-INGR-TR-1SG:OBJ-IND
 ‘This girl made me experience a good smell.’
- (52) **susun kuři bese=?* *rasika-ke-d-ko-a.*
 dance girl very=3SG:SUBJ be.joyful-COMPL-TR-3PL:OBJ-IND
 ‘The dancing girl made them experience joy.’

In order to say the equivalent sentences of (51) and (52) in Mundari, these should be replaced by (53) and (54), respectively.

- (53) *ne kuři-a?* *sowan* *maja* *sowan-ja-?-ñ-a.*
 this girl-GEN fragrance good smell-INGR-TR-1SG:OBJ-IND
 ‘This girl’s fragrance made me experience a good smell.’
- (54) *susun kuři=ko* *lel-ki-?-i-ci* *bese*
 dance girl=3SG:SUBJ see-COMPL-TR-3SG:OBJ-CONJ very
rasika-ke-d-ko-a.
 be.joyful-COMPL-TR-3PL:OBJ-IND
 ‘As they saw the dancing girl she made them experience a joy.’

I, therefore, rewrite the constraint for subject–object assignment here.

- (55) Transitive subject NPs are animate nouns except for the object–experiencer construction.

As illustrated above, an experiential verb in Mundari is considered as an intransitive. That is to say, in the experiencer–subject constructions, only the experiencer assigned to intransitive subject occurs, and the stimulus occurs with the instrumental postposition *-te*. However only a few experiential verbs, for example, *suku* ‘to feel happy, to like’, *kairao* ‘to feel angry, to get an angry’, *giu?* ‘to feel ashamed, to shame’, can act as transitive verbs. In that case, the animate NPs can occupy the transitive object slot not as stimuli, but as beneficiaries. I show the general benefactive construction in (56) and the experiencer–subject and beneficiary–object construction in (57) and (58).

- (56) *manđi am=iñ* *om-a-m-ta-n-a.*
 food you=1SG:SUBJ give-BEN-2SG:OBJ-PROG-ITR-IND
 ‘I am giving you the food.’
- (57) *am=iñ* *giu?-a-m-ta-n-a.*
 you=1SG:SUBJ shame-BEN-2SG:OBJ-PROG-ITR-IND
 ‘I am feeling shame at you.’

- (58) *ne kuṛi-kin=iñ suku-aka-n-a.*
 this girl-DL=1SG:SUBJ like-CONT-ITR-IND
 'I like these girls.'

Examples (57) and (58) have two-place constructions, but include the intransitive marker *-n*. We, therefore, consider them semi-transitive; that is, they are two-place, but intransitive, they have a subject and an indirect object. Further, there are no experiencer–object constructions paired with (57) and (58) in Mundari. This is a great difference between Indo-Aryan and Mundari with respect to experiential constructions.

In sum, object types in Mundari are three; patient–object in the unmarked transitive construction, experiencer–object in the experiential construction, and experiencer–indirect object in the benefactive construction.

3.3.3–3.3.4 Tense and aspect

The tense system is divided into future (unmarked) and non-future (marked).

The future tense implies the habitual aspect like *used to* in English. In this case the iterative verbal base is usually used. Further, the future tense also indicates universal truth. For example,

- (59) *uri?-jilu ka=le jo-jom-a.*
 cattle-meat NEG=1PL.EX:SUBJ eat-ITER-IND
 'We (excl.) never eat beef.'

- (60) *seta? siggi=? rakab-a.*
 morning the Sun=3SG:SUBJ rise-IND
 'The sun rises in the morning.'

Present and past tenses are expressed by the aspect marker following the transitive or intransitive marker. Thus,

(i) Present

- (61) *maṅḍi=iñ jom-ta-n-a.*
 food=1SG:SUBJ eat-PROG-ITR-IND
 'I am eating the food.'

(ii) Past

- (62) *maṅḍi=iñ jom-ke-d-a.*
 food=1SG:SUBJ eat-COMPL-TR-IND
 'I ate the food.'

The aspect markers are classified into perfective and imperfective sets. We will look at the perfective first. The following perfective aspect markers, for example, *a*, *ke*, *le*, *ja* are involved in the verb morphology of Mundari.

(i) *a*.

This is not frequently used. Nobody except Munda takes this aspect marker into consideration. Munda regarded it as a 'cislocative' aspect which implies 'an action which is completed and suspended for an indefinite period of time' (Munda 1971:29). The 'cislocative' *a* can be followed by both the intransitive marker *n* and

the transitive marker *d*. The following examples indicate the relationship between *a-n* and *a-d*:

- (63) *diku=ñ* *itu-a-d-ko-a*.
 Hindi=1SG:SUBJ teach-SUS-TR-3PL:OBJ-IND
 'I have taught Hindi to them.'

- (64) *diku=ñ* *itu-a-n-a*.
 Hindi=1SG:SUBJ teach-SUS-ITR-IND
 'I have been taught Hindi; I have known Hindi.'

We have used the term suspended for the aspect marker *a*.

(ii) *ke*.

This aspect marker indicates the completion of an action without reference to any other action.

- (65) *mañdi=ñ* *jom-ke-a*.
 food=1SG:SUBJ eat-COMPL-IND
 'I will finish eating the food (without waiting for anything).'

- (66) *Ranci-te=ñ* *sen-ke-n-a*.
 Ranchi-to=1SG:SUBJ go-COMPL-ITR-IND
 'I went to Ranchi.'

(iii) *le*.

This aspect marker signifies the completion of an action in relation to some other action. We call it current relevance of anterior, or simply an anterior. We will compare *le* with *ke* as given below.

- (67) *duřum-le-n-a-e?*.
 sleep-ANT-ITR-IND-3SG:SUBJ
 'He/she had slept first (then has already got up).'

- (68) *duřum-ke-n-a-e?*.
 sleep-COMPL-ITR-IND-3SG:SUBJ
 'S/he slept.'

(iv) *ja*.

This aspect marker indicates the completion of an action which is relevant to a current situation; in fact it refers to its inception. According to Comrie (1976:19), 'the other perfect forms of the same verbs can in fact be used to indicate the beginning of a situation (ingressive meaning)'. In other words, this aspect marker is used for ingressive aspect.

- (69) *seno?-ja-n-a-ko*.
 go-INGR-ITR-IND-3PL:SUBJ
 'They have started going; they have just gone.'

- (70) *manḡi=ko jom-ja-d-a.*
 food=3PL:SUBJ eat-INGR-TR-IND
 ‘They have started eating the food; they are eating the food.’

Within the broad domain of imperfectivity ‘a distinction is made between the terms “progressive” and “continuous,” the former being a situation of the latter (progressiveness is the combination of continuousness with non-stativity)’ (Comrie 1976:12). The following imperfective aspect markers have been treated here:

- (v) *ta.*

This aspect marker is labelled ‘progressive’ based on the Comrie’s definition; that is, the combination of continuousness with non-stativity.

- (vi) *aka.*

This aspect marker is labelled ‘continuous’, on the other hand, slightly different from the progressive aspect according to the definition as follows: ‘durative without the habitual’ (Comrie 1976:26).

The distinction between the aspect markers *ta* and *aka* is based on the contrast between a telic and an atelic situation. Telic refers to an event where the activity has a clear terminal point while atelic, where the event has no such natural end-point. For instance,

- (71) *dub-ta-n-a-ko.*
 sit-PROG-ITR-IND-3PL:SUBJ
 ‘They are in the process of sitting.’

(This action will have a terminal point when they sit down.)

- (72) *dub-aka-n-a-ko.*
 sit-CONT-ITR-IND-3PL:SUBJ
 ‘They are sitting.’

(They have already sat down. They can stand up or continue sitting. It does not matter.)

Hence we will consider the aspect marker *aka* as continuous in an atelic situation. Unlike English, the stative verb in Mundari has both progressive and continuous forms. Furthermore, transitive verbs with a continuous aspect correspond to the experiential perfect in English.

- (73) *Ranci do=ñ lel-aka-d-a.*
 Ranchi top=1SG:SUBJ see-CONT-TR-IND
 ‘I have seen Ranchi; I have been to Ranchi.’

- (74) *Ranci-te=ko sen-aka-n-a.*
 Ranchi-to-3PL:SUBJ go-CONT-ITR-IND
 ‘They have gone to Ranchi (and have not yet come back).’

3.3.5 Mood

There are three moods in Mundari: indicative (unmarked), imperative, and optative. Imperative mood is marked by deleting the indicative marker *a* in an indicative

sentence. Negation of imperative, that is, prohibitive, is marked by preposing *alo* to a verbal base. Optative mood indicates the attitude of the speaker: it expresses wishes and is marked by *ka* which precedes the pronominal suffix. Although Munda (1971) called it subjunctive, we adopt the term optative as Hoffmann (1903), Cook (1965), and N.K. Sinha (1975) did. Negation of optative is marked by *alo + ka*, which are preposed to a verbal base. The modal intensifier *ko?* can be added to imperative and optative sentences. It implies politeness.

The basic structures are as follows:

(i) Imperative

	second	<i>-m/me</i>	singular
VB (+ AM) (+ OBJ) +	personal	<i>-ben</i>	dual
	suffix	<i>-pe</i>	plural

(ii) Optative

VB (+ AM) (+ OBJ) + NEG *ka* personal suffix

(iii) Prohibitive for second person

	second	<i>-m</i>	
<i>alo</i>	person	VB (+ AM) (+ OBJ) +	<i>-a</i>
	suffix	<i>-pe</i>	

(iv) Prohibitive for first and second persons

alo ka personal suffix VB (+ AM) (+ OBJ) + *-a*

The sentence examples are given below:

(75) *maṅḍi jom-e-me.*
 food eat-EPEN-2SG
 'Eat the food.'

(76) *maṅḍi jom-e-ka-ko.*
 food eat-EPEN-OPT-3PL
 'May they eat the food.'

(77) *alo=pe jom-a.*
 NEG=2PL eat-IND
 'Don't eat.'

(78) *alo-ka=ko jom-a.*
 NEG-OPT=3PL eat-IND
 'They must not eat.'

We note that all aspect markers except the suspended aspect marker *a* can follow a verbal base in the imperative sentence. The semantic difference between aspect markers can be shown below.

(i) Completion

(79) *jom-ke-m.*

‘Eat up; if you don’t eat it you won’t get a chance to eat later.’

(ii) Progressive

(80) *jom-ta-m.*

‘(You should) Eat (it); our life is not immortal. So you should eat it’.

(iii) Anterior

(81) *jom-le-m.*

‘Eat (it) first (and then).’

(iv) Ingressive

(82) *jom-ja-m.*

‘Eat (it) along (while on your way to doing something else).’

(v) Continuous

(83) *jom-aka-m.*

‘Eat (it) (continuously).’

3.3.6 Orientation/directionality

Not investigated in this study.

3.3.7 Valence/voice

There are three means of decreasing valency; that is, reflexive, reciprocal, and passive.

Reflexive is expressed by the suffix *-en* after consonants or *-n* after vowels. For example,

(84) *Soma=e? lel-en-ta-n-a.*

Soma=3SG:SUBJ see-RFLXV-PROG-ITR-IND

‘Soma is looking at himself.’

The reflexive *-en/-n* is highly productive but the following verbs do not take the reflexive suffix: *dub* ‘to sit’, *kami* ‘to work’, *giti?* ‘to lie down’, *ajom* ‘to feed’. The reflexive refers to ‘a verb where the subject and the object relate to the same entity’ (Crystal 2003). Causative verbs cannot be reflexivized; for example, *ajom* ‘to feed’. Interestingly, the class of intransitive verbs which can be causativized is also not

reflexivized. Thus, *dub* ‘to sit’ is intransitive but can take the transitive marker to add causativity, as in the following:

- (85) *dub-aka-n-a-e?*
sit-CONT-ITR-IND-3SG:SUBJ
‘He has sat, that is, he is still sitting.’
- (86) *hon-ko=e? dub-aka-d-ko-a.*
child-PL=3SG:SUBJ sit-CONT-TR-3PL:OBJ-IND
‘S/he has caused the children to sit down.’

The reciprocal marker <*pV*> is monosemous and never attached to non-verbal stems. The reciprocal decreases verb valency. Thus the reciprocal verbal base takes only the intransitive marker *-n* even with ditransitive verbs. For instance,

- (87) *Soma seta hon-ko=e? om-ki-?-i-a.*
Soma dog child-PL=3SG:SUBJ give-COMPL-TR-3SG:OBJ-IND
‘Soma gave the dog to the children.’
- (88) *seta-ko=le o<po>m-ta-n-a.*
dog-PL=1PL.EX:SUBJ give<RECIP>give-PROG-ITR-IND
‘We are giving the dogs to each other.’

Passive verbal bases can be formed by suffixing *-o?* to a verbal stem. The passive suffix can be attached to either transitive or intransitive verbs. The passive may imply the sense of possibility, that is, a ‘passive potential’, as is common in other Munda languages.

- (89) (a) *ayum* ‘to hear’ → *ayum-o?* ‘to be audible’
(b) *lel* ‘to see’ → *lel-o?* ‘to be visible’
(c) *duřum* ‘to sleep’ → *duřum-o?* ‘to feel sleepy’

Further, passivity in Mundari implies non-volitionality. Unlike Mundari, passive in Hindi (Pandharipande 1978) and Bengali (Klainman 1986) makes crucial reference to the semantic notion of volitionality. For instance, we may look at the following Mundari sentences:

- (90) *duřum-o?-ta-n-a-e?*
sleep-PASS-PROG-ITR-IND-3SG:SUBJ
‘S/he is feeling sleepy (by a non-volitional cause).’
- (91) *kug-o?-ta-n-a-e?*
cough-PASS-PROG-ITR-3SG:SUBJ
‘He is beginning to cough.’

Thus, the following verbs cannot be passivized due to their implication of volitionality. For example, *co?* ‘to kiss’ → **cog-o?*, *dulařa* ‘to love’ → **dulařa-o?*.

The means of increasing valency are causative, conjugation change, and benefactive.

Causative is expressed by the unproductive prefix *a-*:

- (92) (a) *jom* ‘to eat’ → *a-jom* ‘to feed’
(b) *nu* ‘to drink’ → *a-nu* ‘to give to drink’

Conjugation change affects labile verbs. For convenience we shall consider the transitive use of labile stems as causativization as I have shown above in (85) and (86). Thus,

- (93) (a) *dub* i. 'to sit' (with the intransitive marker *-n*)
 (b) ii. 'to cause to sit' (with the transitive marker *-d*)

(93a=85) *dub-aka-n-a-e?*
 sit-CONT-ITR-IND-3SG:SUBJ
 'He has sat, that is, he is still sitting.'

(93b=86) *hon-ko=e?* *dub-aka-d-ko-a.*
 child-PL=3SG:SUBJ sit-CONT-TR-3PL:OBJ-IND
 'S/he has caused the children to sit down.'

The benefactive suffix *-a* (always followed by the beneficiary agreement marker) indicates not only a beneficiary argument added to two-place transitive but also (optionally) the indirect object of ditransitives, and in this case both forms may coincide. Although the benefactive suffix increases valency when added to a two-place transitive, the benefactive verbal base takes the intransitive marker *-n* only. This may be the reason why reciprocals cannot be derived from the benefactive. Compare:

- (94) (a) *daru=m* *ma?-ke-d-a.*
 tree=2SG:SUBJ cut-COMPL-TR-IND
 'You cut the tree.'
 (b) *daru=m* *mag-a-ñ-ke-n-a.*
 tree=2SG:SUBJ cut-BEN-1SG-COMPL-ITR-IND
 'You cut the tree for me.'

3.3.8 (Non-)finiteness

Non-finite verb forms can be made by the deletion of the indicative marker *a*, and are used in the formation of relative clauses. They are of the following type:

Verbal Base (+ Aspect Marker + Transitive/Intransitive Marker).

The aspect markers and transitive/intransitive markers can be deleted in the future tense. Aspect markers are obligatorily followed by the transitive/intransitive marker *-d/-n* in relative clauses, although aspect markers without transitive/intransitive markers can be used in the finite verbal system. Thus,

(95) *jom-hoꝛo*
 eat-person
 'the man who will eat'

(96) *jom-ke-d-hoꝛo*
 eat-COMPL-TR-person
 'the man who ate it'

**jom-ke-hoꝛo* is not grammatical, although *jom-ke-a=e?* (eat-COMPL-IND=3SG:SUBJ) 'He/she will have eaten the food' is completely acceptable.

The non-finite form can be followed by a noun or pronoun in a relative clause and by a postposition in a subordinate clause. I discuss these clauses in sections 4.4 and 4.5.

3.3.9 Negation

The negative markers in Mundari are *ka* and *alo*.

ka is highly productive for lexical and sentence negation in indicative sentences. It is a morphologically bound form. *alo* is used for the negation of imperative or optative sentences.

Basically *ka* is 'No' in polar (yes–no) questions. It also functions as a sentence negation marker, and is then fixed in preverbal position, followed by the subject agreement element. If the subject is inanimate the negation marker *ka* is put in preverbal position as a free form. For instance,

(97) *Ranci-te=m seno?-ta-n-a ci.*
 Ranchi-to=2SG:SUBJ go-PROG-ITR-IND Q
 'Are you going to Ranchi?'

(98) *ka, ka=ñ senog-a.*
 NEG NEG=1SG:SUBJ go-IND
 'No, I don't go.'

(99) *ne gari Ranci-te ka senog-a.*
 this car Ranchi-to NEG go-IND
 'This car will not go to Ranchi.'

As is shown in the section on Mood (section 3.2.5), *alo* functions as a prohibitive marker in imperative sentences, and indicates the negation of hope and desire in optative sentences.

In sentences with an auxiliary verb we use special forms for negation. These are *bano?* for inanimates, *baygai?* for first person singular and third person singular and *bay* for animate other than first person singular and third person singular. I discuss these in detail in section 3.2.11.

3.3.10 Derivation

Verbal derivation in Mundari is a type of partial reduplication. It is not productive, and is only applied to closed monosyllabic words. Beside partial reduplication, we have full reduplication. This formation is productive. For example,

Verb	Partial reduplication	Full reduplication
<i>sab</i> 'to catch'	<i>sa-sab</i>	<i>sab-sab</i>
<i>jom</i> 'to eat'	<i>jo-jom</i>	<i>jom-jom</i>
<i>tud</i> 'to pick'	<i>tu-tud</i>	<i>tud-tud</i>
<i>goe?</i> 'to kill'	<i>go-goe?</i>	<i>goe?-goe?</i>
<i>taygi</i> 'to wait'		<i>taygi-taygi</i>
<i>rakab</i> 'to rise'		<i>rakab-rakab</i>

The partial reduplication may indicate either:

- (i) Repetition, implying an emphatic action.
- (ii) Customary act, implying a universal fact.

For instance,

(100) *alo=m kumbuꞑu-a da-dal-a-ko.*
 NEG=2SG thief-IND strike-ITER-IND-3PL:SUBJ
 ‘Don’t steal it. They strike you repeatedly.’

(101=59) *uriꞑ-jilu ka=le jo-jom-a.*
 cattle-meat NEG=1PL.EX:SUBJ eat-ITER-IND
 ‘We (EX.) never eat beef.’

The following partial reduplication is lexicalized:

(102) *tasad tu-tud-aka-n-a.*
 grass pick up-ITER-CONT-ITR-IND
 ‘The grass is ready to pick up.’

(103) *uli-joo jo-jom-aka-n-a.*
 mango-fruit eat-ITER-CONT-ITR-IND
 ‘The mango fruit is ready to eat up.’

In addition to the above meaning, we have the meaning of action of limited duration, just in full reduplication.

(104) *tanꞑi-tanꞑi-ke-n-a-le.*
 wait-wait-COMPL-ITR-IND-1PL.EX
 ‘We (excl) waited for a little while.’

I describe full reduplication formation later, in the section on serial verb constructions (section 3.2.12).

3.3.11 Copula

The equivalent of the verb ‘to be’ has two forms in Mundari:

- (i) The existential copula *menaꞑ*. This refers to the subject’s location in space.
- (ii) The identity copula *tan*. This refers to the subject’s identity.

Both are irregular verbs. The animate subject of the existential copula *menaꞑ* is encoded in the verb morphology while the inanimate subject is not encoded. Thus,

(105) *Soma oꞑaꞑ-re menaꞑ-i-a.*
 Soma house-LOC COP-3SG-IND
 ‘Soma is in the house.’

(106) *Soma tan-iꞑ.*
 Soma COP-3SG:SUBJ
 ‘It is Soma.’

Both the copula *mena?* and *tan* are defective verbs, occurring only in the present. These forms merge into *tai* ‘to stay’ in the past and future (the past form *tai-ke-n*, the future form *tai-n*). For instance,

- (107) *Soma oʔaʔ-re=? tai-ke-n-a.*
 Soma house-LOC=3SG COP-COMPL-ITR-IND
 ‘Soma was in the house.’
- (108) *Soma=eʔ tai-ke-n-a.*
 Soma=3SG:SUBJ COP-COMPL-ITR-IND
 ‘It was Soma’ or ‘Soma was there.’
- (109) *Soma oʔaʔ-re=? tai-n-a.*
 Soma house-LOC=3SG:SUBJ COP-ITR-IND
 ‘Soma will be in the house.’
- (110) *Soma=eʔ tai-n-a.*
 Soma COP-ITR-IND
 ‘It is going to be Soma who will be there’ or
 ‘Soma will be there.’

Negation in copula sentences is slightly more complicated. The negative of *mena?* has three variants. For example,

- (111) *Soma oʔaʔ-re baŋgaiʔ-i-a*
 Soma house-in COP NEG-3SG-IND
 ‘Soma is not in the house.’
- (112) *parkom oʔaʔ-re banoʔ-a.*
 bed house-LOC COP NEG-IND
 ‘A bedstead is not in the house.’
- (113) *hon-ko oʔaʔ-re baŋ-ko-a.*
 child-PL house-LOC COP NEG-3PL-IND
 ‘Children are not in the house.’

The negative of *tan* is formed by just adding to the negative marker *ka* before *tan* as in a regular verb.

- (114) *Soma ka tan-iʔ.*
 Soma NEG COP-3SG
 ‘It is not Soma.’

The identity copula *tan* may be related to the progressive aspect marker *ta*. The distinction between the existential copula and the identity copula may be considered to be an Indian areal feature.⁷

3.3.12 Auxiliary or serial verb constructions

Serial verb constructions have attracted the attention of linguists who are concerned with South Asian languages. Ever since Masica (1976) considered compound verbs with the ‘explicator verb’ as a typological areal feature, studies on compound verbs in South Asian context have been seriously attempted.

In Indo-Aryan languages, there are two types of verb–verb constructions; one is a verb with a conjunctive participle and another is a compound verb with an ‘explicator’ verb. For example, in Hindi a conjunctive participle is the so-called *-kar* construction whereas a compound verb is the combination of a main (polar) verb and an explicator (vector, operator, or intensifier) which indicates completion, passivity, permission, etc. of the action or process expressed by the main verb. Thus, we consider this type as an auxiliary verb construction.

In Mundari there are two types of the serial verb construction:

- (i) Main verb + Main verb. Its meanings are similar to the conjunctive participle construction in Hindi, that is, they are perfective and simultaneous.
- (ii) Main verb + Explicator. Unlike other South Asian languages, most of the operators retain their original meaning.

I illustrate the following examples in (i) here:

(a) Pair action

- (115) *hatu-re=liḡ jom-nu-ke-d-a.*
 village-in=1DL.EX eat-drink-COMPL-TR-IND
 ‘We two (excl) ate and drank in the village; we two took dinner in the village.’

These are pair actions like *seno?-hiju?* ‘to keep company with’, *isin-basaḡ* ‘to cook and boil, that is, to make food’, etc.

(b) Sequential action

- (116) *ne saan saḡima-cetaḡ-te=bu dondo-rakab-e-a.*
 this firewood roof-over-to-1PL.INC lift-go up-it-IND
 ‘We will lift this firewood and take it up to the roof.’

This class of serial verbs are *go?-ader* ‘to carry (something) on the shoulder and take it into the house’, *dul-pere?* ‘to pour and fill up’, etc.

(c) Result

- (117) *ne hon=le asul-maraḡ-ki-?-i-a.*
 this child=1PL.EX:SUBJ feed-grow-COMPL-TR-3SG:OBJ-IND
 ‘We (excl) fed and raised this child up.’

Examples of this type of serial verb are: *sen-nam* ‘to go and meet (somebody)’, *nir-laga* ‘to run and get tired’, etc.

(d) Cause

- (118) *añ-a? hon-kin banda-re=kin ḡumbui?-goe?-ja-n-a.*
 my child-DL pond-LOC=3DL:SUBJ be drown-die-INGR-ITR-IND
 ‘My two children have died by drowning in a pond.’

Other examples are: *ma?-goe?* ‘to kill with an axe’, *haka-goe?* ‘die by hanging’, etc.

(e) Simultaneous action or events

- (119) *ne gaṛa poṭopoṭia-te=ko har-parom-ke-d-a.*
 this river motorbike-by=3PL:SUBJ drive-CROSS-COMPL-TR-IND
 'They drove the motorbike and crossed the river.'

This class of serial verbs are: *dub-hape* 'while sitting to keep a silence', *duray-au* 'to come along singing', etc.

The second type (ii) Main verb + Explicator is a common feature in South Asian languages. The verbs meaning 'come' and 'go' are common explicators in Indo-Aryan and Dravidian languages (Kachru and Pandharipande 1980:115). In Mundari the equivalents of these verbs are never used as explicators, and never appear as the second member of serial verbs either. U.N. Singh *et al.* (1986) listed the equivalents of the following eighteen vectors (=explicators) for their analysis for classifying polar verbs in selected South Asian languages.

TAKE, GIVE, GO, DIE, RISE, SIT, COME, THROW,
 KEEP, MOVE, SEND, SEE, KILL, COME-OUT,
 BRING-OUT, HOLD, BRING.

Among them the following verbs can be considered the second member of serial verb constructions in Mundari.

English gloss	Mundari verbs	Meaning as a second verb of the series	Serial verbs
TAKE	<i>idi</i>	motion onward or away continuation	<i>nir-idi</i> 'to run away' <i>kuli-idi</i> 'to go on asking a question'
DIE	<i>goe?</i>	to the last degree	<i>rasika-goe?</i> 'to rejoice excessively' <i>landa-goe?</i> 'to be convulsed with laughter'
RISE	<i>rakab</i>	motion up	<i>raca?-rakab</i> 'to pull up' <i>nir-rakab</i> 'to run up'
THROW	<i>giṛi</i>	to exceed	<i>laga-giṛi</i> 'to get tired excessively' <i>pere?-giṛi</i> 'to full to excess'
SEND	<i>kul</i>	to send	<i>ra?-kul</i> 'to send to call' <i>kiriṅ-kul</i> 'to send to buy' <i>apir-uṛuṅ</i> 'to fly out'
COME OUT	<i>uṛuṅ</i>	motion out of the place	<i>dul-uṛuṅ</i> 'to pour out'
BRING	<i>au</i>	motion from a given point towards the speaker (to and fro)	<i>nir-au</i> 'to come running this way' <i>jom-au</i> 'to go for taking food, and then come back'

We will illustrate below the second verbs of the series whose meaning is slightly different from their basic meaning as main verbs, that is, the ones that have been (partially) grammaticalized.

Basic meaning	Mundari verbs	Semantic function	Serial/Auxiliary verbs
EAT	<i>jom</i>	self-benefactive	<i>giti?-jom</i> 'to lie down for one's benefit' <i>nam-jom</i> 'to get for oneself'
FULL	<i>pere?</i>	completion	<i>om-pere?</i> 'to give completely' <i>leka-pere?</i> 'to count completely'
MAKE	<i>bai</i>	carefully	<i>lel-bai</i> 'to look carefully/attentively' <i>ayum-bai</i> 'to listen carefully'
PLUCK (a fruit)	<i>god</i>	for a moment	<i>ayum-god</i> 'to hear for a moment' <i>lel-god</i> 'to see for a moment'
PLUCK (a potherb)	<i>sid</i>	to stop	<i>jagar-sid</i> 'to stop talking' <i>sayad-sid</i> 'to stop breathing'

The following verbs, when used as second verbs of the series, can act not only as modal and aspectual auxiliaries but also as adverbials:

Basic meaning	Mundari verbs	Semantic function	Serial verbs
FINISH	<i>caba</i>	perfective	<i>om-caba</i> 'to finish giving' <i>lel-caba</i> 'to finish seeing'
BEGIN	<i>e?e?</i>	inchoative	<i>jom-e?e?</i> 'to start eating' <i>ol-e?e?</i> 'to start writing'
WIN	<i>da?i</i>	can	<i>ol-da?i</i> 'to be able to write' <i>lel-da?i</i> 'to be able to see'
DO	<i>rika</i>	causative	<i>ol-rika</i> 'to cause to write' <i>bai-rika</i> 'to cause to make'
RETURN	<i>ru?a</i>	again, back	<i>lel-ru?a</i> 'to see again' <i>jom-ru?a</i> 'to eat again'
EXCEED	<i>laa?</i>	excessively	<i>du?um-laa?</i> 'to sleep excessively' <i>jom-laa?</i> 'to eat excessively'
FRONT	<i>ayar</i>	ahead	<i>seno?-ayar</i> 'to go ahead' <i>nir-ayar</i> 'to run ahead'
BACK	<i>tayom</i>	later	<i>jom-tayom</i> 'to eat later' <i>sen-tayom</i> 'to go later'

Every serial verb can be marked by affixation for mood on the basis of semantic and grammatical constraints operable on the second verb. In the Main verb + Main verb construction, reciprocal infixation of <pV> is applied to each of the main verbs; for example, *dondo-rakab* 'to lift and go up', *do<po>ndo-ra<pa>kab* 'to lift and go up each other'.

Verbal intensifiers follow a verbal base. This construction is similar to a serial verb construction, but the second element is not a free form but rather a bound form. Unlike Indo-Aryan, the system of verbal intensifiers is very rich in Mundari. For instance,

- (i) V-*ba?ltab* 'V quickly'
- (120) *ma?di jom-ba?-e-me.*
 food eat-quickly-it-2SG
 'Eat the food quickly.'
- (ii) V-*bapad/goroe?* 'V by all means'
- (121) *en ku?i au-bapad-i-me.* (EM)
 that girl bring-by all means-3SG-2SG
 'Marry the woman by all means.'

- (iii) *V-baṛa/balaṅ* ‘V here and there’
- (122) *aṛandi-naṅgen=ko sen-baṛa-ta-n-a.*
 marriage-for=3PL:SUBJ go-here and there-PROG-ITR-IND
 ‘They are going here and there for a marriage.’
- (iv) *V-coṭe?* ‘almost V’
- (123) *kaji-coṭe?-ke-d-ci=ko seno?-ja-n-a*
 say-almost-COMPL-TR-CONJ=3PL:SUBJ go-INGR-ITR-IND
 ‘They began saying and stopped in the middle, then have gone.’
- (v) *V-garaṅ* ‘intensified V’
- (124) *kakala-garaṅ-i?-me.*
 shout-loudly-3SG-2SG
 ‘Shout at him/her loudly.’
- (vi) *V-hantaṛa* ‘engage in V’
- (125) *jom-hantaṛa-e-me.*
 eat-engage in-it-2SG
 ‘Be engaged in eating it.’
- (vii) *V-katelkuca* ‘V repeatedly’
- (126) *maṅḍi=ko jom-kuca-ke-d-a.*
 food=3SG:SUBJ eat-repeatedly-COMPL-TR-IND-3PL
 ‘They ate it repeatedly.’
- (viii) *V-no?* ‘V a little (while)’
- (127) *saman go?-no?-le-m.*
 luggage carry on the shoulder-a little-ANT-2SG
 ‘Carry it on the shoulder for a while.’
- (ix) *V-torsa* ‘V along’
- (128) *hiju?-torsa-me.*
 come-along-2SG
 ‘Come back at once.’
- (x) *V-tuka* ‘V and return’⁸
- (129) *idi-tuka-ñ-ka-e?*
 take-return-1SG-OPT-3SG
 ‘He/she may take me and return.’
- (xi) *V-utar* ‘V entirely’
- (130) *Ranci-te=ko seno?-utar-ja-n-a.*
 Ranchi-to=3PL:SUBJ go-entirely-INGR-ITR-IND-3SG
 ‘They went away to Ranchi for ever.’

3.4 Expressives

Mundari has a rich system of expressives. The term ‘expressive’ was suggested by Diffloth (1976:263–264) and adopted by Emeneau (1980:7) in the South Asian context in the following:

‘(E)xpressive’ is the most inclusive term for a form class with semantic symbolism and distinct morphosyntactic properties; ‘ideophones’ are a subclass in which the symbolism is phonological; ‘onomatopoeics’ are ideophones in which the reference of the symbolism is acoustic (i.e. imitative of sounds). Since the ideophones may have reference not only to sounds, but to any other objects of sense, including internal feelings as well as external perceptions (sight, taste, smell, etc.), and since the Indo-Aryan/Dravidian items already examined have this very wide type of reference, the broadest term ‘expressives’ seems appropriate.

I have already written about Mundari expressives in my grammar (Osada 1992:140–144). However, I could not touch the syntactic and semantic properties of expressives. Thus I will discuss here (1) morphology (2) syntax, and (3) semantics of expressives.

3.4.1 Morphology of expressives

Expressives can be divided into the following types on the basis of their word formation pattern:

- Full reduplication
- Partial reduplication
- Vowel mutation

3.4.1.1 Full reduplication

This type of expressive should be distinguished from verbal reduplication, which is clearly derived from the verbal base. It is a salient feature that a basic unit of the reduplicational element has no meaning. Thus,

Expressive form	Meaning
<i>cakob cakob</i>	‘to eat noisily’
<i>lugum lugum</i>	‘to mumble (something)’
<i>hayam hayam</i>	‘to talk in whispers’
<i>gusu gusu</i>	‘an inactive character’
<i>suyuy suyuy</i>	‘lean and small (person)’
<i>kase kase</i>	‘to look askance at (a person)’
<i>mondor mondor</i>	‘a smell of rice beer’
<i>mogo mogo</i>	‘a smell of flowers’
<i>kata kata</i>	‘to roar with laugh by many people’
<i>mugui? mugui?</i>	‘smiling cheerful’

3.4.1.2 Partial reduplication

Partial reduplication can be formed by two elements. The second element is a partial reduplication of the first element. We can subcategorize this type according to

the exact formal pattern. So far it has been impossible to find common meanings associated with each partial formal type:

(i) CVX pVX

Expressive form	Meaning
<i>riti piti</i>	'very small leaves as those of tamarind'
<i>risuṛi pisuṛi</i>	'the act of showing the teeth again and again'
<i>rasa pasa</i>	'a continuous rustle of dry leaves, paper, or straw as produced by the gliding of a snake or the passage of a rat or other small animals'
<i>laṭa paṭa</i>	'to make a stew thick, pasty'
<i>laṭar paṭar</i>	'a mixture of truth and lies wherein one does not know what to believe'
<i>ledeṅ pedeṅ</i>	'so fat that in walking he has difficulty'
<i>leco peco</i>	'connoting several recurvations, or twisting of the mouth to the right and left'
<i>loso poso</i>	'a loosely limbed body with soft or flabby muscles'

(ii) CVX bVX

Expressive form	Meaning
<i>kau bau</i>	'to do uncomfortably or uneasily'
<i>kered bered</i>	'a quarrelling and fighting disposition'
<i>cere bere</i>	'chattering and twittering of numerous birds'
<i>cali bali</i>	'trickiness'
<i>lada bada</i>	'the thuds of things soft, as mud, falling in succession'
<i>ladi badi</i>	'to put things in a disorderly manner, more or less one over another'
<i>sador bador</i>	'the act of letting bits fall while eating of strewing bits all around by pecking'
<i>rada bada</i>	'onomatopoeia of hail, dry fruit or other hard and dry things falling all about in rapid succession, also of rather numerous drops of water falling all about'

(iii) CVX mVX

Expressive form	Meaning
<i>celoṅ meloṅ</i>	'naughty boy'
<i>ceṅgol meṅgol</i>	'shamelessness'
<i>jaka maka</i>	'shining with a flashy dress (sari with gold)'
<i>jiki miki</i>	'shining with leather'
<i>rigi migi</i>	'a cloth variegated with parallel lines or squashes of various colour'
<i>keoṅ meoṅ</i>	'a feeling of loneliness and fear in the middle of the forest'
<i>kaṅḍu maṅḍu</i>	'indigestion and pain in the pit of the stomach after eating or drinking something acid or sour or unripe'
<i>seled meled</i>	'mixture of different kinds of grain, etc.'
<i>gero mero</i>	'a shamed face or a crying face'

(iv) CVX kVX

Expressive form	Meaning
<i>dale kale</i>	‘negligent (of taking care)’
<i>haŋi kuŋi</i>	‘to do the opposite action’

(v) CVX gVX

Expressive form	Meaning
<i>rain gain</i>	‘good or bad principles of conduct’ (EM)
<i>mane gane</i>	‘want of punctuality in starting dilatoriness’ (EM)

(vi) CVX eVX

Expressive form	Meaning
<i>repo cepo</i>	‘shrivelled’
<i>dukur cukur</i>	‘uneasiness of mind’

(vii) CVX jVX

Expressive form	Meaning
<i>reŋge jeŋge</i>	‘the condition of getting bothered or being subjected to trouble or annoyance’
<i>haufu jauŋu</i>	‘desultory talk or conversation, passing from one subject to another without order or natural connection’
<i>runu junu</i>	‘to go or walk with difficulty due to a handicap’

(viii) CVX dVX

Expressive form	Meaning
<i>rawa dawa</i>	‘opportunity to do something reprehensible, because there is nobody to interfere’

(ix) CVX tVX

Expressive form	Meaning
<i>ribui? tibui?</i>	‘the act of fat people, walking with the buttocks rubbing against each other’
<i>roka toka</i>	‘quickly’

(x) CVX sVX

Expressive form	Meaning
<i>rahan sahan</i>	‘the use of dress, furniture, plate and utensils by more civilized people’
<i>boro soro</i>	‘cowardice’

(xi) CVX rVX

Expressive form	Meaning
<i>tiri riri</i>	‘the sound of a flute’

(i), (ii), and (iii) are very common.

3.4.1.3 Vowel mutation

This type of expressives are divided into six on the basis of vowel mutational patterns.

(i) (C)aC[(C)a(C)] (C)uC[(C)u(C)]

Expressive form	Meaning
<i>ɖala ɖulu</i>	'a fat and short person'
<i>lada ludu</i>	'a fat child'
<i>ladar ludur</i>	'a wrinkled old person'
<i>aŋgar uŋgur</i>	'to look around restlessly'
<i>ba? bu?</i>	'holes here and there'
<i>caba? cubu?</i>	'the splashing sound made by repeated poking with a stick in water or mud'
<i>ɖapa? ɖupu?</i>	'baby tries to walk'
<i>tagam tugum</i>	'a fat person who cannot walk swiftly'
<i>lada ludu</i>	'a fat baby'

(ii) CaC[a(C)(a)] CoC[o(C)(o)]

Expressive form	Meaning
<i>sar sor</i>	'to eat away with a savage appetite'
<i>karae? koroe?</i>	'a gurgling breathing of one being strangled'
<i>kal kol</i>	'a buzzing of the ears'
<i>ragaɾa rogoɾo</i>	'soil mixed with stones so large that it cannot be ploughed'
<i>rakaɾa rokoɾo</i>	'the rattling of something in a box or in a bottle or the like'
<i>ɖaŋ ɖoŋ</i>	'a deep and big hole'
<i>paɖad pogod</i>	'a swollen state of the whole body'
<i>ɖarad ɖorod</i>	'a sound of frog'

(iii) CaC[aC] CiC[iC]

Expressive form	Meaning
<i>palad pilid</i>	'the act of shining in various places'
<i>par pir</i>	'the act of dispersing'

(iv) CaC[(C)aC] CeC[(C)eC]

Expressive form	Meaning
<i>paŋɖad peŋɖed</i>	'a glitter of light appearing and disappearing now here, then there'
<i>ca? ce?</i>	'used for the cry of babies' (EM)

(v) CiCa(C) CoCo(C)

Expressive form	Meaning
<i>kidar kodor</i>	'a rooster with a long upright comb and long wavy feathers on the neck and tail' (EM)
<i>kiɾaŋ koɾoŋ</i>	'a tall and lean person'
<i>gida godo</i>	'semi-liquid things'
<i>pica poco</i>	'to empty a soft or pasty substance by compression'

(vi) CiC CoC

Expressive form	Meaning
<i>bir bor</i>	'tall and straight'
<i>lir lor</i>	'a long and weak sapling'

The formal analysis of expressives has been done.

3.4.2 Syntax of expressives

The syntax of expressives has never been described. Expressives can occupy any place, that is, in a predicate, complement, or argument slot. As the head of a predicate, expressives can take derivational suffixes, for example, passive, reflexive, benefactive, and aspect markers. Expressives can also form serial verb constructions. Thus,

- (131) *busu?re seta-hon=e utul-putul-ta-n-a.*
 straw-LOC dog-child=3SG:SUBJ EXPR-PROG-ITR-IND
 ‘The puppy is playing in the straw then the straw is shaking.’

- (132) *nir-nir-te=? anggor-sanggor-giri-aka-n-a.*
 run-run-to=3SG:SUBJ EXPR-throw away-CONT-ITR-IND
 ‘S/he is running and running then s/he is totally getting out of breathe.’

Some expressives require an experiencer object like in the experiential constructions. For instance,

- (133) *rua-te alae-balae-ki-?-ñ-a.*
 fever-to EXPR-COMPL-TR-1SG:OBJ-IND
 ‘I got a trouble by a fever.’

An expressive alone or an expressive with the progressive aspect marker *ta* and the intransitive marker *-n* can occupy in the complement slot as an adverbial phrase in the following:

- (134) *kata-kata=e landa-ta-n-a.*
 EXPR=3SG:SUBJ smile-PROG-ITR-IND
 ‘S/he is laughing uproariously.’
- (135) *iri?-iri?-ta-n=(e)-m landa-ta-n-a.*
 EXPR-PROG-ITR=EPEN-2SG:SUBJ smile-PROG-ITR-IND
 ‘You are smiling like you are mocking somebody.’

An expressive can occupy in the argument slot to modify a noun or noun phrase. For example,

- (136) *ini? do janao ako?-bako? ho?o-ge.*
 that person TOP always EXPR person-EMPH
 ‘S/he is always a stupid person.’

An expressive can occupy in the head of noun phrase in the following instance:

- (137) *ini?-a? isiri-siki?i ka=ñ suku-a.*
 that person-GEN EXPR NEG=1SG like-IND
 ‘I don’t like her coquettish laughing.’

As is seen above, expressives have a reduplicated form. Although the single form has usually no meaning, some single forms which are followed by the completive aspect

marker *ke* and intransitive marker *n* occupy the complement slot as an adverbial phrase:

- (138) *tii=e?* *cadʒa-cadʒa-ke-d-a.*
 hand=3SG:SUBJ clap:EXPR-COMPL-TR-IND
 ‘S/he clapped her/his hand.’
- (139) *cadʒa-ke-n=e?* *tabʒi-li-ʔ-i-a.*
 clap-COMPL-ITR=3SG:SUBJ slap-ANT-TR-3SG:OBJ-IND
 ‘S/he slapped him/her like clapping.’

3.4.3 Semantics of expressives

Nobody has ever described the semantics of expressives in Mundari. Hoffmann has just described the several expressive forms as variants in EM. For example, the following thirteen forms are the sole entry for ‘a smile to smile etc.’:

mogoeʔ, mogoeʔ-mogoeʔ, mergoeʔ, mergoeʔ-mergoeʔ, merloŋ, merloŋ-merloŋ, mirluŋ, mirluŋ-mirluŋ, moeʔ-moeʔ, muguiʔ, muguiʔ-muguiʔ, musuiʔ, musuiʔ-musuiʔ.

According to my informants some forms such as *mogoeʔ-mogoeʔ, mirluŋ-mirluŋ, moeʔ-moeʔ* are not known by them because of dialectal differences. They, however, can differentiate meanings in the following:

mergoeʔ mergoeʔ ‘smiling in mouth’
merloŋ merloŋ ‘smiling by children or aged-persons who have no teeth’
muguiʔ muguiʔ ‘smiling cheerful’
musuiʔ musuiʔ ‘smiling in eyes shyly’

Apart from these, there are a lot of expressives to express the action of laughing, etc. I demonstrate the semantic field of laughing, smiling and chuckling below.⁹

hada hada ‘to roar with laughter successively’
kata kata ‘to roar with laughter (less than *hada-hada*) by many people’
kaʔ kaʔ ‘to laugh like a hen’s clucking’
keʔ keʔ ‘to laugh like a jackal’s howling’
keteʔ keteʔ ‘to laugh innocently (by children)’
kōē kōē ‘to laugh without sound’
kere kere ‘to laugh while talking’
isiʔi isiʔi ‘to ridicule one’s action or talk’
isiʔi sikiʔi ‘to laugh coquettishly’
iriʔ iriʔ ‘to laugh like a mock at’

I give another example of expressive for light reflection in the following:

jaka jaka ‘shining with gold’
jaka maka ‘shining with a flashy dress (sari with gold)’
jiki miki ‘shining with leather’
caka maka ‘shining with steel or silver’
jili mili ‘shining with building’
jilib jilib ‘dazzle with electric light’

<i>bijir bijir</i>	‘lighting’
<i>jilab jolob</i>	‘glimmering with a firefly’
<i>jolob jolob</i>	‘glimmering with many fireflies’
<i>jaran jaran</i>	‘glittering in the sun’
<i>pangad pegged</i>	‘a glitter of light appearing and disappearing now here, then there’
<i>piřid piřid</i>	‘glimmering on the sand’
<i>palad pilid</i>	‘the act of shining in various places’
<i>pilid pilid</i>	‘twinkling with stars’

3.4.3.1 *Sound symbolism*

As far as sound symbolism is concerned, ‘it is often said that if vowel quality is used for size symbolism, [i] will symbolize smallness, and the lower vowels, especially [a], will symbolize largeness, with degrees in between’ (Diffloth 1994:107). Diffloth, however, has suggested a counter-example (*i*: big, *a*: small) from Bahnar, which also belongs to the Austroasiatic language family.

In Mundari, it seems to me that *i* symbolize smallness while *a* symbolize largeness in the following:

<i>sařa sařa</i>	‘a passing rain for a long time’
<i>siři siři</i>	‘a passing rain’
<i>jařam jařam</i>	‘a heavy rain (the water in the river is full)’
<i>jiřim jiřim</i>	‘a heavy rain (the water in the rice-field is full)’
<i>kaca kaca</i>	‘to scold somebody with action’
<i>kici kici</i>	‘to scold somebody only by mouth’

The following cases should be taken into consideration in our future study:

<i>baya baya</i>	‘to act lazily’
<i>buyu buyu</i>	‘to act, especially walk lazily (more lazy than <i>baya-baya</i>)’
<i>pisir pisir</i>	‘to drizzle (not enough to get wet, even without an umbrella)’
<i>pusur pusur</i>	‘to drizzle (but to get wet)’

4 SYNTAX

4.1 Syntax of the simple sentence

As we have seen in section 3.2.2, the subject and object of a sentence are determined by word order. The unmarked word order is as follows: S + O + Verb.

The word order is not fixed for subject NP and object NP. Subject and object agreement, therefore, is very important for the signalling of grammatical relations. But in some cases ambiguity cannot be excluded. When the subject NP and object NP have the same person and number, the sentence is ambiguous. For instance,

- (140) *Soma seta=e?* *hwa-ki-?-i-a.*
 Soma dog=3SG:SUBJ bite-COMPL-TR-3SG:OBJ-IND
 (a) ‘Soma bit the dog.’
 (b) ‘The dog bit Soma.’

On pragmatic grounds, meaning (a) may be less likely. But if *pusi* ‘cat’ is placed in the first position instead of Soma, the sentence is totally ambiguous.

- (141) *pusi seta=e?* *hua-ki-?-i-a.*
 cat dog=3SG:SUBJ bite-COMPL-TR-3SG:OBJ-IND
 (a) ‘The cat bit the dog.’
 (b) ‘The dog bit the cat.’

One construction in which subjects and objects are distinguishable is in relative clause constructions. The head NP in relative clauses can be a subject or an object in Mundari. But the transitive marker and intransitive marker can be used to distinguish the subject head NP from the object head NP: the intransitive marker is used when the pivot is the subject. Thus,

- (142) (a) *pusi hua-ke-n-seta=e?* *goe?-ja-n-a.*
 cat bite-COMPL-ITR-dog=3SG:SUB die-INGR-ITR-IND
 ‘The dog who bit the cat has died.’
 (b) *pusi hua-ke-d-seta=e?* *goe?-ja-n-a.*
 cat bite-COMPL-TR-dog=3SG:SUB die-INGR-ITR-IND
 ‘The dog whom the cat bit has died.’

The subject and object agreement elements can be marked only when the subject NP and object NP are classified as animate nouns.

In addition to two arguments, a postpositional phrase or adverb denoting location or time can be inserted into any position before the verb. In this case the orders of NPs and PP are rather free except the last position, which is reserved for the verb. We illustrate this in (143).

- (143) (a) *seta?-re seta-ko maṅḍi=ko jom-ke-d-a.*
 morning-LOC dog-PL food=3PL:SUBJ eat-COMPL-TR-IND
 (b) *maṅḍi seta-ko seta?-re=ko jom-ke-d-a.*
 food dog-PL morning-LOC=3PL:SUBJ eat-COMPL-TR-IND
 (c) *maṅḍi seta?-re seta-ko=ko jom-ke-d-a.*
 food morning-LOC dog-PL=3PL:SUBJ eat-COMPL-TR-IND
 ‘In the morning the dogs ate the food.’

4.2 Typological features

I list the constituent order below:¹⁰

- (i) S + O + V

- (144) *Soma maṅḍi=? jom-ke-d-a.*
 Soma food=3SG:SUBJ eat-COMPL-TR-IND
 ‘Soma ate the food.’

- (ii) S + O + V or O + S + V

- (145=141) *pusi seta=e?* *hua-ki-?-i-a.*
 cat dog=3SG:SUBJ bite-COMPL-TR-3SG:OBJ-IND
 (a) ‘The cat bit the dog.’
 (b) ‘The dog bit the cat.’

(iii) Oblique + S + O + V or S + Oblique + O + V or S + O + Oblique + V

- (146=143) (a) *seta?re seta-ko maᅇᅇi=ko jom-ke-d-a.*
 morning-LOC dog-PL food=3PL:SUBJ eat-COMPL-TR-IND
 (b) *maᅇᅇi seta-ko seta?re=ko jom-ke-d-a.*
 food dog-PL morning-LOC=3PL:SUBJ eat-COMPL-TR-IND
 (c) *maᅇᅇi seta?re seta-ko=ko jom-ke-d-a.*
 food morning-LOC dog-PL=3PL:SUBJ eat-COMPL-TR-IND
 ‘In the morning the dogs ate the food.’

(iv) Noun Phrase + Postposition
 For example, *oᅇa?re* ‘at home’

(v) Genitive + Noun Phrase
 For example *diri-rea? oᅇa?* ‘a stone house’

(vi) Adjective + Noun
 For example *maray oᅇa?* ‘a big house’

(vii) Demonstrative + Noun
 For example, *ne oᅇa?* ‘this house’

(viii) Numeral + Noun
 For example, *baria oᅇa?* ‘two houses’

(ix) Relational clause + Head Noun

- (147=142) (a) *pusi hua-ke-n-seta=e? goe?-ja-n-a.*
 cat bite-COMPL-ITR-dog=3SG:SUBJ die-INGR-ITR-IND
 ‘The dog who bit the cat has died.’
 (b) *pusi hua-ke-d-seta=e? goe?-ja-n-a.*
 cat bite-COMPL-TR-dog=3SG:SUBJ die-INGR-ITR-IND
 ‘The dog whom the cat bit has died.’

(x) Degree word + Adjective
 For example, *bese maray* ‘very big’

(xi) Final position of polar question particle

- (148) *Soma maᅇᅇi=? jom-ke-d-a ci.*
 Soma food=3SG:SUBJ eat-COMPL-TR-IND Q
 ‘Did Soma eat the food?’

(xii) First position of interrogative in content questions

- (149) *oko-e hiju?-aka-n-a*
 who come-CONT-ITR-IND
 ‘Who has come?’

(xiii) Clause + Adverbial subordinator

- (150) *maᅇᅇi jom-tayom-te=ko senog-a.*
 food eat-after-LOC=3PL:SUBJ go-IND
 ‘After eating the food they will go.’

4.3 Complex sentence structure

4.3.1 Coordination

Coordination is expressed by the following particles:

(i) *oʔoʔlad* ‘and’

This coordinating conjunction can conjoin not only noun phrases but also clauses.

(a) Noun Phrases

- (151) *araʔ sim oʔoʔ hende merom*
 red fowl and black goat
 ‘red fowl and black goat’

(b) Clauses

- (152) *jom-ke-d-a-eʔ ad=eʔ senoʔ-ja-n-a.*
 eat-COMPL-TR-IND-3SG and=3SG:SUBJ go-INGR-ITR-IND
 ‘He/She ate and went away.’

(ii) *ci* ‘or’

This coordinating conjunction can connect not only noun phrases but also clauses.

(a) Noun Phrases

- (153) *araʔ sim ci hende merom*
 red fowl or black goat
 ‘red fowl or black goat’.

(b) Clauses

- (154) *haga-m Ranci-te-ʔ senoʔ-ja-n-a ci hatu-re menaʔ-i-a.*
 brother-your Ranchi-to-3SG go-INGR-ITR-IND or village-LOC COP-3SG-IND
 ‘Your brother has gone to Ranchi or he is at home.’

(iii) *ca* ‘or’

This coordinating conjunction is not used to connect two noun phrases but to connect two clauses.

- (155) *Soma hijuʔ ca=eʔ senoʔ ka-ñ itu-a-n-a.*
 Soma come or=3SG:SUBJ go NEG-LSG:SUBJ know-SUS-ITR-IND
 ‘I don’t know whether Soma comes or goes.’

(iv) *ci-aʔci* ‘because’

According to Hoffmann in EM, ‘this conjunction was introduced into the translation of the Bible made by the first Lutheran Missionaries’ (p. 843). It has been made by the calque of Hindi *kyōki* or *cūki*.

- (156) *Ranci-te nida=le tebaʔ-ke-d-a. ci-aʔci bas*
 Ranchi-to night=1PL.EX:SUBJ reach-COMPL-TR-IND because bus
bagrao-le-n-a.
 be.broken-ANT-TR-IND
 ‘We (excl.) arrived at Ranchi at night, because the bus was out of order.’

(v) *mendo* ‘but’

This coordinating disjunction *mendo* can be analyzed into *men* ‘to say’ and the particle *do*. This can connect two sentences.

- (157) *maṅḍi jom-mone-ja-ʔ-ñ-tai-ke-n-a. mendo ka-ñ*
 food eat-want-INGR-TR-1SG-COP-COMPL-ITR-IND but NEG-1SG
nam-ke-d-a.
 get-COMPL-TR-IND
 ‘I wanted to eat the food, but I have not got it.’

(vi) *karedo* ‘otherwise’

This consists of the negator *ka* the postposition *-re* and the particle *do*. This connects two sentences, especially an indicative sentence.

- (158) *maṅḍi jom-le-m. karedo loyoṅ-te=bu senog-a.*
 food eat-first-2SG otherwise rice field-to=1PL.INC go-IND
 ‘Eat the food first, or we will go to the rice-field.’

4.3.2 *Relative-type clauses*

I have already illustrated the non-finite form in Mundari. The non-finite form can be followed by a noun or pronoun in a relative clause. For instance,

- (159) *Ranci-te sen-ke-n-hoṛo=eʔ hijuʔ-ruṛa-ja-n-a.*
 Ranchi-to go-COMPL-ITR-person=3SG:SUBJ come-return-INGR-ITR-IND
 ‘The person who went to Ranchi has just returned.’
- (160) *Ranci-te sen-ke-n=iʔ hijuʔ-ruṛa-ja-n-a.*
 Ranchi-to go-COMPL-ITR=3SG:SUBJ come-return-INGR-ITR-IND
 ‘The one who went to Ranchi has just returned.’
- (161) *abu jom-ke-d-sim-do=eʔ sibil-ge-tai-ke-n-a.*
 1PL.INC eat-COMPL-TR-chicken-TOP=3SG:SUBJ tasty-EMPH-COP-COMPL-ITR-IND
 ‘The chicken that we (INC) ate was tasty’.

In a relative clause the following points can be noted:

- (i) The head noun can be the object (161) or the subject (159) of the non-finite verb. It is very clear that the object follows the transitive marker *d* and the subject follows the intransitive marker *n*. As was shown above, this is a syntactic test to distinguish objecthood from subjecthood.
- (ii) The head noun can be omitted; in this case the third person singular form which follows is *=iʔ* rather than *=eʔ*.
- (iii) The personal pronoun other than the third person cannot be allocated in the postverbal position as a head noun in a relative clause. Thus, **jom-ke-d-pe* ‘you who ate it’ is ungrammatical (*-pe* is second personal plural suffix). If we change ‘one’ to ‘you’ in the sentence (160), we should paraphrase it with two sentences.

- (162) *Ranci-te=m sen-ke-n-a. ena-te=m*
 Ranchi-to=2SG:SUBJ go-COMPL-ITR-IND that-by=2SG:SUBJ
hijuʔ-ruṛa-ja-n-a.
 come-return-COMPL-ITR-IND
 ‘You went to Ranchi and (then) you have just returned.’

4.3.3 Subordinate clauses

The non-finite form can be followed by a postposition or a question marker in a subordinate clause. The postposition *re* and the question marker *ci* are most frequently used. The difference between *re* and *ci* is complicated. The basic meaning *re* is ‘in that time’ while *ci* is ‘while’. The usage of the aspect marker is also complicated. I give an example of each below:

- (163) *maŋdi jom-ta-n-ci=ko jagar-ta-n-a.*
 food eat-PROG-ITR-CONJ=3PL:SUBJ talk-PROG-ITR-IND
 ‘While they are eating the food, they are talking simultaneously.’
- (164) *maŋdi jom-ta-n-re=ko jagar-ta-n-a.*
 food eat-PROG-ITR-LOC=3PL:SUBJ talk-PROG-ITR-IND
 ‘When they are eating the food they are talking at that time.’
- (165) *hoŋo-ko ukuŋa jagar-ta-n-ci=ko*
 people-PL backbiting talk-PROG-ITR-CONJ=3PL:SUBJ
ayum-ke-d-liŋ-a.
 hear-COMPL-TR-1DL.EX:OBJ-IND
 ‘While we two were backbiting, people heard us.’
- (166) *maŋdi jom-ta-n-re hoŋo?-re=ñ di?-ja-n-a.*
 food eat-PROG-ITR-LOC throat-LOC=1SG:SUBJ stick-PROG-ITR-IND
 ‘When I was eating the food it stuck in my throat.’
- (167) *paisa nam-aka-d-ci rasika-ja-?-ñ-a.*
 money get-CONT-TR-CONJ joy-INGR-TR-1SG:OBJ-IND
 ‘I am joyful because I have got money.’
- (168) *paisa nam-aka-d-re uri?=le kiriy-li-?-i-a.*
 money get-CONT-TR-LOC cattle=1PL.EXC:SUBJ buy-ANT-TR-3SG:OBJ-IND
 ‘By the time we got money, we had already bought the cattle.’

As for the personal marking, the subject marking in the subordinate clause can be omitted when the same subject occurs in the main clause in the following:

- (169) *hon-ko maŋdi jom-ke-d-ci=ko seno?-ja-n-a.*
 child-PL food eat-COMPL-TR-CONJ=3PL:SUBJ go-INGR-ITR-IND
 ‘When the children had eaten the food they went.’
- (170) *hon-ko maŋdi=ko jom-ke-d-ci*
 child-PL food=3PL:SUBJ eat-COMPL-TR-CONJ=3PL:SUBJ
kami-te=ñ seno?-ja-n-a.
 work-to=1SG:SUBJ go-INGR-ITR-IND
 ‘As the children ate the food I have gone to work.’

The *-re* clause also means conditional ‘if’. The combination *-re-do* (locative and topic marker) is more common for the conditional clause:

- (171) *bisi jom-ja-n-re go-goe?-a.*
 poison eat-INGR-ITR-LOC die-ITER-IND
 ‘If the poison will be taken everybody shall die.’

- (172) *maŋdi=m jom-ke-d-re-do piiŋi-te=laŋ senog-a.*
 food=2SG:SUBJ eat-COMPL-TR-LOC-TOP market-to=1DL.INC:SUBJ go-IND
 ‘If you finish eating the food let us go to the market.’

Instead of *re*, *sida-re* ‘before’ and *tayom-te* can be used in the subordinate clauses:

- (173) *maŋdi jom-sida-re=ko seno?-ja-n-a.*
 food eat-before-LOC=3PL:SUBJ go-PROG-ITR-IND
 ‘Before eating the food they have gone.’
- (174) *maŋdi jom-tayom-te=ko senog-a.*
 food eat-after-LOC=3PL:SUBJ go-IND
 ‘After eating the food they will go.’

The following examples denote supposed reality. In this sentence *te-ra?* or *honaŋ* should be added:

- (175) *ini?-lo? aŋandi-ja-n-re-do naa?-lo?-do=ñ*
 that person-with marry-INGR-ITR-LOC-top now-with-top=1SG:SUBJ
buŋia-ja-n-te-ra?
 old-INGR-ITR-INS-GEN
 ‘If I would marry him I should get old (I am still young because I didn’t marry him).’
- (176) *añ-honaŋ seta bai-ja-n-re-do cadlom*
 1SG-if dog become-INGR-ITR-LOC-TOP tail
peŋe-peŋe-baŋa-i-a.
 sway-repeatedly-EPEN-IND
 ‘If I were a dog I may sway my tail repeatedly (I flatter my master).’

Another postposition *lo?* can follow the non-finite form in a subordinate clause:

- (177) *maŋdi jom-ta-n-lo?=ko duraŋ-ta-n-a.*
 food eat-PROG-ITR-LOC=3PL:SUBJ sing-PROG-ITR-IND
 ‘While they are eating the food they are singing.’

It is very similar to (163). But *jom-ta-n-lo?* is used in a wider range of contexts than *jom-ta-n-ci*.

The postposition *ate* ‘from’ cannot follow the non-finite form directly. But *ke-ate* in the indicative sentence and *le-ate* in the imperative sentence are very common when denoting successive action:

- (178) *maŋdi jom-ke-ate=ko seno?-ja-n-a.*
 food eat-COMPL-from=3PL:SUBJ go-INGR-ITR-IND
 ‘After taking the food they have gone.’
- (179) *maŋdi jom-le-ate seno?-me.*
 food eat-ANT-from go-IND
 ‘When they are eating the food they are talking in that time.’

In addition to *ke-ate* and *le-ate* there are *ta-n-ate*, *ke-n-ate*,¹¹ *ja-n-ate*, *le-n-ate*, and *aka-n-ate*, as in the following:

- (180) *maŋɖi jom-ta-n-ate=m paɾao-ta-n-a.*
 food eat-PROG-ITR-from=2SG:SUBJ study-PROG-ITR-IND
 ‘You are still eating but starting to study.’
- (181) *maŋɖi jom-ke-n-ate=ko aboŋ-en-ja-n-a.*
 food eat-COMPL-ITR-from=3PL:SUBJ wash-RFLXV-INGR-ITR-IND
 ‘After taking the food they have washed their hands’.
- (182) *eŋga-m goeʔ-ja-n-ate bar-sirma hoba-ja-n-a.*
 mother-your die-INGR-ITR-from two-year happen-INGR-ITR-IND
 ‘It has taken two years after your mother’s death.’
- (183) *dub-aka-n-ate=ñ jom-ta-n-a.*
 sit-CONT-ITR-from=1SG:SUBJ eat-PROG-ITR-IND
 ‘I am eating while sitting on the ground.’

The concessive clauses are as follows:

- (184) *jom-e-ca ka=eʔ jom-e maŋɖi bai-ta-m.*
 eat-it-or NEG=3SG:SUBJ eat-it food make-PROG-2SG
 ‘Whether s/he eats or not, you must be ready for the food.’
- (185) *senoʔ-re-o ka=m senoʔ-re-o kami-do calao-a.*
 go-LOC-also NEG=2SG:SUBJ go-LOC-also work-TOP go well-IND
 ‘Whether you go or not, the work goes well.’

A negative clause is illustrated in the following:

- (186) *ne-hoʔo auri nu-bairi-do ka=eʔ jagar-a.*
 this-person yet drink-only-TOP NEG=3SG:SUBJ talk-IND
 ‘He doesn’t talk until he drinks liquor.’

4.4 Switch reference

According to Anderson and Boyle (2002:48), based solely on some of the very limited data in (Osada 1992), it was speculated that *ci* in Mundari might be a same subject marker. Thus,

- (187) *jom-ke-d-ci=ko senoʔ-ja-n-a.*
 eat-COMPL-TR-=3PL:SUBJ go-INGR-ITR-IND
 ‘They went away as soon as they had eaten.’

In our observation it is not the same subject marker as is shown below.

- (188) *añ jom-ke-d-ci=ko senoʔ-ja-n-a.*
 1SG eat-COMPL-TR-CONJ=3PL:SUBJ go-INGR-ITR-IND
 ‘They went away as soon as I had eaten.’

5 SEMANTICS/DISCOURSE

5.1 Semantics

There are many verbs in Mundari which are equivalent to the meaning of ‘cut’ in English.

- (i) *had*
‘to cut with sawing motion by a knife, a saw, an axe, etc.’
- (ii) *ged*
‘to cut meat by *bāiṭi* (a large meat cutter)’
- (iii) *laṭab*
‘to cut a paper, hair, etc. by scissors’
- (iv) *ma?*
‘to cut a tree with a striking motion by an axe’
- (v) *od*
‘to cut through and through a tree and so fell it’
- (vi) *sama?*
‘to cut something (meat bone, jackfruit, etc.) in small pieces with a *koṅḍe* (axe)’
- (vii) *ir*
‘to cut the stalk of grain by *ḍatrom* (sickle) for reaping’
- (viii) *gaṅḍui?*
‘to cut the long things (rope, etc.) in pieces of a given length’
- (ix) *ṭona*
‘to cut timber into log’
- (x) *ḍula*
‘to cut the top portion of trees with *duṅu* (long axe)’
- (xi) *paṭa?*
‘to cut the long things (firewood, etc.) into two portions’

According to Suwilai (2002), there are numerous verbs in the same semantic field in Kham, which belongs to the Austroasiatic language family spoken in Cambodia.

5.2 Discourse

As far as discourse is concerned, there are the topic marker and emphatic marker in Mundari. I describe these here.

The particle *do* follows the noun phrase or postpositional phrase that it marks as the topic in discourse. It singles out the element about which the comment is made in the sentence. The following elements can be marked by the topic marker *do*:

(i) Subject

- (189) *añ do=ñ senog-a.*
 1SG TOP=1SG:SUBJ GO-IND
 ‘I will go.’

(ii) Direct object

- (190) *maŋɖi do=ñ jom-ke-d-a.*
 food TOP=1SG eat-COMPL-TR-IND
 'I ate the food.'

(iii) Location

- (191) *Ranci-re do=ñ tai-ke-n-a.*
 Ranchi-LOC TOP=1SG:SUBJ live-COMPL-ITR-IND
 'I lived in Ranchi.'

(iv) Source

- (192) *Ranci-ate do=m hiju?-aka-n-a.*
 Ranchi-from TOP=2SG:SUBJ come-CONT-ITR-IND
 'You have come from Ranchi.'

(v) Instrumental

- (193) *ne daru hake-te do=pe ma?-ke-d-a.*
 this tree axe-by TOP=2PL:SUBJ cut-COMPL-TR-IND
 'You cut this wood with an axe.'

(vi) Benefactive

- (194) *Soma-nanjen do naki?=ñ kiriŋ-a-i?-a.*
 Soma-for TOP comb=1SG:SUBJ buy-BEN-3SG-IND
 'I will buy the comb for Soma.'

(vii) Comitative

- (195) *añ-lo? do han-te=laj sen-ke-n-a.*
 1SG-with TOP yonder-to=1PL.INC:SUBJ go-COMPL-ITR-IND
 'We two (inc) went there together.'

The possessive is not marked for the topic by the particle *do*. For example, **diri-rea? lra? do oŋa?* (*diri* 'stone', *-rea?l-ra?* GEN *oŋa?* 'house') is ungrammatical. While the independent possessive is marked for the topic. Thus, *añ-ag-a? do mena?*. (*añ-ag-a?* 'mine', *mena?* COP) 'There is mine.'

In addition to the topic of a sentence, *do* in Mundari marks the contrast as well. For instance,

- (196) *am do Ranci-te, añ do Kuŋti-te=ñ sen-ke-n-a.*
 2SG TOP Ranchi-to 1SG TOP Khunti-to=1SG:SUBJ go-COMPL-ITR-IND
 'You (went) to Ranchi, but I went to Khunti.'

Further, as we have mentioned in the indefinite *oko* can be followed by the topic mark *do* as in (198) but the interrogative *oko* cannot.

- (197) *oko-e hiju?-aka-n-a*
 who come-CONT-ITR-IND
 'Who has come?'
- (198) *oko-e do hiju?-aka-n-a.*
 someone TOP come-CONT-ITR-IND
 'Someone has come, (but not all).'

The particle *ge* may function as an emphatic marker in discourse. The following elements can be marked by the emphatic marker *ge*:

(i) Subject

- (199) *añ ge=ñ senog-a.*
 1SG EMPH=1SG:SUBJ GO-IND
 'It is I who will go.'

(ii) Direct Object

- (200) *mañdi ge=ñ jom-ke-n-a.*
 food EMPH=1SG:SUBJ eat-COMPL-ITR-IND
 'It is the food that I ate.'

(iii) Benefactive

- (201) *Soma-nanngen ge naki?=ñ kirij-a-i?-a.*
 Soma-for EMPH comb=1SG:SUBJ buy-BEN-3SG-IND
 'For Soma, I will buy the comb.'

(iv) Source

- (202) *Ranchi-ate ge=m hiju?-aka-n-a.*
 Ranchi-from EMPH=2SG:SUBJ come-CONT-ITR-IND
 'From Ranchi, you have come.'

(v) Instrumental

- (203) *ne daru hake-te ge=pe ma?-ke-n-a.*
 this wood axe-with EMPH=2PL:SUBJ cut-COMPL-ITR-IND
 'With axe, you cut this wood.'

(vi) Independent possessive

- (204) *nea do añ-ag-a? ge mena?.*
 this TOP 1SG-GEN-GEN EMPH COP
 'This is mine (not any other persons).'

While the topic marker *do* is never allocated in the postverbal position, the emphatic marker *ge* can be used for the verbal phrase in postverbal position. Thus,

- (205) *mañdi=ko jom-ta-n-ge-a.*
 food=3SG:SUBJ eat-PROG-ITR-EMPH-IND
 'They are taking food indeed.'

As seen above, the interrogative *oko* cannot be followed by the topic marker *do* as in (198) but it can be followed by the emphatic marker *ge*. The indefinite *oko*, on the other hand, cannot be followed by the emphatic marker *ge*. For instance,

- (206) *okoe ge her-le-d-a mani do*
 who EMPH SOW-ANT-TR-IND mustard TOP
 'Who has sown the mustard seed indeed?'

6 LEXICON

6.1 Austroasiatic/Munda components

It is easy to list the cognated words as Schmidt (1906) and Pinnow (1959) have proposed. I list here some words with Pinnow's index number.

Mundari meaning	Pinnow's index number
<i>baba</i> 'paddy'	V-4
<i>bir</i> 'forest'	V-321
<i>buru</i> 'mountain, hill'	V-278
<i>da?</i> 'water'	V-2
<i>hon</i> 'child'	V-205
<i>jaŋ</i> 'seed, bone'	V-7
<i>japid</i> 'to shut eyes'	V-93
<i>kula</i> 'tiger'	V-281
<i>lutur</i> 'ear'	V-147
<i>lu?</i> 'to ladle out'	V-400
<i>mui?</i> 'ant'	V-130
<i>mid/miyad, mod/moyod</i> 'one'	K-338
<i>nu</i> 'to drink'	V-112
<i>ra?</i> 'to call, cry'	V-57
<i>sim</i> 'fowl'	V-314
<i>ti</i> 'hand'	K-160

6.2 Loan strata

There are many loanwords in Mundari mainly from the adjoining Indo-Aryan languages, notably Sadani. I note the following points:

- (i) As for nouns, Mundari-speaking area is in a boundary of vowel variant *a/o* in the Indo-Aryan. So some word forms coexist in the following: *man/mɔn* 'mind' < Indo-Aryan *man/mɔn*, *badnam/bodnam* 'infamous' < Indo-Aryan *badnām/bodnam*, etc.
- (ii) As for verbs, those with the endings *-ao* in Mundari are borrowed from the Indo-Aryan languages; For example, *bujao* 'to understand', *dekaao* 'to show', *hatao* 'to remove', etc.
- (iii) In some dialects, the dative marker *ke* has been introduced from Sadani. Thus, *ape-ke joar* (2PL-Dative greeting) 'Joar to you'.
- (iv) From a phonetic point of view, the final *b* in some loanwords is realized as a so-called checked consonant; For example, */kitab/* 'book' [kitaʔb̥].

7 BRIEF ANALYSED TEXTS

7.1 Riddles

- (i) Q. *aʀakaʀa palad-pilid*.
rafter flickering:EXPR
'Over the rafters of the roof flickering here and there.'
- A. *kaʀea* 'mouse'

- (ii) Q. *aṭa-maṭa bir-ko tala-re boṅga hon-ko*
dense jungle-PL middle-LOC spirit son-PL
caṭu=ko harub-ta-d-a.
the earthen pot=3PL:SUBJ COVER-PROG-TR-IND
'The sons of spirits are covering the earthen pot in the middle of jungle.'
- A. *ud* 'mushroom'
- (iii) Q. *aṭa-maṭa bir-ko tala-re boṅga hon-ko*
dense jungle-PL middle-LOC spirit child-PL
ub=ko raṅa-ta-d-a.
hair=3PL:SUBJ untie-PROG-TR-IND
'The sons of spirit are untying their hair in the middle of jungle.'
- A. *badcom.* 'a kind of grass which looks like a child who untied his hair'
- (iv) Q. *catoma-re oṛa? lo-ta-n-a.*
Catoma (Village name and umbrella)-LOC house fire-PROG-ITR-IND
gagara-re=ko ra?-ta-n-a.
Gagara (Village name and brasspot)-LOC=3PL:SUBJ cry-PROG-ITR-IND
'The house in the umbrella is on fire while they are crying in the brasspot.'
- A. *huka* 'hookkah (a kind of pipe for smoking)'
- (v) Q. *da ta-ñ koko? ḍaṅḍa?*
give me GEN-1SG hockey stick
'Give me my hockey stick.'
- A. *seta cadlom.* 'dog's tail'
- (vi) Q. *dub-me daru ḍemka daru de?-a.*
sit-2SG tree dwarf tree climb-IND
'Sit down please tree, dwarf tree will climb.'
- A. *haṛad boo?* 'a yam of the jungles'
- (vii) Q. *edel daru rau-ta-n-a*
cotton tree felt DOWN-PROG-ITR-IND
bajuṅia-ko hiju?-ta-n-a.
musician-PL come-PROG-ITR-IND
'The cotton tree is falling down then the musicians are coming.'
- A. *i?* 'excrement' (the raw cotton tree smells. the musician means a fly.)
- (viii) Q. *gaṛa-gaṛa-te ledera atu-ta-n-a.*
river-river-to cloth float-PROG-ITR-IND
'The clothes are floating towards rivers.'
- A. *geded* 'algae'
- (ix) Q. *gaṛa-gaṛa-te laka koṛa kapi go?-aka-d-a.*
river-river-to thickset boy axe carry-CONT-TR-IND
'The thickset fellow has carried the axe on shoulder towards rivers.'
- A. *kaṛakom* 'crab'
- (x) Q. *gaṛa-gaṛa-te laka koṛa paciri tapa?-ta-n-a.*
river-river-to thickset boy wall raise-PROG-ITR-IND
'The thickset fellow is raising the walls towards rivers.'
- A. *kaṛakom* 'crab'

7.2 Kande's story

This is a part of *seña jumbaṛa* 'Tree of intelligence in the Munda way' written by the late Kande Munda who was a good writer and musician.

- (i) *kalam ku<nu>li utrao-re=ñ met-a-i-ta-n-a.*
 pen ask-Nom-ask rise-LOC=1SG:SUBJ say-BEN-3SG-PROG-ITR-IND
 'When the question rises why you are writing this I am saying to him/her.'
- (ii) *jumbaṛa mane ḍoṛo? daru-ko-ra?*
 a bunch of branch and leaf mean sponge gourd tree-PL-GEN
koto-ko-re sakam ar sakam-koto-ge
 branch-PL-LOC leaf and leaf-branch-EMPH
jumbaṛa-giṛi-aka-n-jumbaṛa.
 dense-totally-CONT-ITR-a bunch of branch
 'Jumbaṛa, a dense bunch of branches and leaves means jumbaṛa which is a dense bunch of flowers in the branch of the tree.'
- (iii) *jaan jetan bir-ko-re ana?-mana? naṅi-ko,*
 any any forest-PL-LOC various:EXPR climber-PL
daru-cupad-ko-te juuṅṅu-giṛi-aka-n-jumbaṛa
 tree-bush-PL-to bush-totally-CONT-ITR-a bunch of branch
 'In any forest jumbaṛa in which various creepers overgrow in the trees and bushes.'
- (iv) *jumbaṛa tuṅ hiju?-ta-n-a.*
 a bunch of branch meaning come-PROG-ITR-IND
 'The meaning of jumbaṛa is coming clear.'
- (v) *ne vacak-re mosa-te-ge bariya oroto hambud-aka-n-a.*
 this significance-LOC once-by-EMPH two meaning embrace-CONT-ITR-IND
 'This significance embraces two meanings.'
miyad-do hulduh-hulduh daru jumbaṛa-ko-re
 one-TOP dense:EXPR tree a bunch of branch-PL-LOC
nana-boron ceṅe-cipurub=ko dub-aka-n-ci
 various:EXPR bird-feather=3PL:SUBJ sit-CONT-ITR-that
niral cere-bere, kere-ḍore-ta-n=ko jagar-a.
 clean chatting-EXPR screeching:EXPR-PROG-ITR=3PL:SUBJ talk-IND
 'First, when the various birds are sitting in the dense branch they are talking clearly to each other.'
- (vi) *ne jagar-ta-n-ko-re ako-a? duku-suku-ko-ra?*
 this talk-PROG-ITR-PL-LOC 3PL-GEN grief-joy-PL-GEN
jagar-ko-ge dayṅ jaa?=ko ja<pa>gar-ta-n-a.
 talk-PL-EMPH indeed perhaps=3PL:SUBJ talk<RECIP>PROG-ITR-IND
 'They are perhaps talking each other's grief and joy in their talk.'
- (vii) *en leka-ge ne puti paṛao-ko-yo.*
 that like-EMPH this book read-PL-also
 'To read this book is also in the same way like a bird.'

- (viii) *jidān-jumbaṛa-re=ko* *dub-aka-n-a.*
 life-a bunch of branch-LOC=3PL:SUBJ sit-CONT-ITR-IND
 ‘They are sitting in the dense bunch of branch.’
- (ix) *ad duku-suku-ra? heṛem-haṛad oncar-re=ko tain-a.*
 and grief-joy-GEN sweet-bitter care-LOC=3PL:SUBJ remain-IND
 ‘And their desire remains a sweet or bitter memory in their suffering and joyful life.’
- (x) *neya seṇa-jumbaṛa cae*
 this intelligence-jumbaṛa or
duku-suku-jumbaṛa men-ke-re-yo bai-o?-a.
 grief-joy-jumbaṛa say-COMPL-LOC-also make-PASS-IND
 ‘You can call it the tree of intelligence or the tree of grief and joy.’
- (xi) *paṛao-ta-n-re gaṭa-gaṭa-ko-re med-da?*
 study-PROG-ITR-LOC various:EXPR-PL-LOC ear-water(=tear)
joro=e leka-n-kaji-ko nam-og-a.
 shed=3SG:SUBJ like-RFLXV-word-PL meet-PASS-IND
 ‘S/he is faced with the critical words because of which s/he sheds tears.’
- (xii) *ar lai? capu-capu-te landa=e leka-n-a?-o.*
 and belly touch-touch-INS laugh=3SG:SUBJ like-RFLXV-GEN-also
 ‘And s/he is also convulsed with laughter.’
- (xiii) *ne puti seṇa-jumbaṛa duku-suku-jumbaṛa*
 this book intelligence- jumbaṛa grief-joy-jumbaṛa
ar prem-priṭ-jumbaṛa=bu men-ta-?
 and love-love-jumbaṛa=1PL.INC say-PROG-TR
 ‘This book is the tree of intelligence or the tree of grief-joy or the tree of love.’
- (xiv) *ol-o?-do neya haya-te-o mone-ja-n-a.*
 write-PASS-TOP this desire-INS-also desire-INGR-ITR-IND
 ‘I am writing because of the following intention.’
- (xv) *juug badli-idi-badli-idi-ta-n-a.*
 time change-continue-change-continue-PROG-ITR-IND
 ‘The time is changing and changing continuously.’
- (xvi) *muṇḍa daṅḡa-daṅḡi-ko isukul-kolej-ko-re*
 Munda boy-girl-PL school-college-PL-LOC
isu=ko paṛao-n-ja-n-a.
 very=3PL:SUBJ study-RFLXV-INGR-ITR-IND
 ‘The Munda boys and girls have studied at school and college.’
- (xvii) *isu=ko paṛao-n-ta-n-a ar*
 very=3PL:SUBJ study-RFLXV-PROG-ITR-IND and
ayar-te-o=ko paṛao-n-ge-a.
 future-INS-also=3PL:SUBJ study-RFLXV-EMPH-IND
 ‘They are studying very much and will study in the future.’

- (xviii) *paṛao-seṇa-ja-n-isu-hoṛo-ko* *jagar-ta-n* *ayum-le-ko-re*
 study-intelligent-INGR-ITR-many-person-PL talk-PROG-ITR hear-ANT-PL-LOC
enḡa-a? *jagar muṇḡari* *hirum-hirum-te-do=ko*
 mother-GEN talk Mundari co-wife-co-wife-INS-TOP=3PL:SUBJ
lusuṇḡu-giṛi-ja-ṛ-i-leka=ko *ayum-o?*
 despondent-totally-INGR-TR-3SG-like=3PL hear-PASS-IND.
 ‘When educated persons speak and hear in their mother tongue, Mundari, they hear totally in a despondent way to face their co-wives.’
- (xix) *sen-hora seṇaṇḡe*, *paidal cae moṭor-ko-re*,
 go-road go-about walk or motorcycle-PL-LOC
 ‘Along the path and travelling way by walk or by motorcycle,’
- (xx) *buru piṭi-piṛi-ko-re*, *jeta hoṭel-ko-re*,
 festival market-place-PL-LOC any hotel-PL-LOC
 ‘in the festival and weekly market place, or in any hotel,’
- (xxi) *ne samae huṛiṇ-huṛiṇ paṛao-aka-n-muṇḡa-danḡa-danḡi-ko*
 this time little-little study-CONT-ITR-Munda-boy-girl-PL
 ‘The Munda boys and girls who educated a little bit recently.’
- (xxii) *na<pa>m-ja-n-lo?-ge* *hindi-te-ge*
 meet<RECIP>INGR-ITR-COMIT-EMPH Hindi-INS-EMPH
joar-jagar=ko *eṭe?-torsa-e-a*.
 greeting-talk=3PL:SUBJ begin-quickly-EPEN-IND
 ‘(They) are quickly greeting and talking in Hindi.’
- (xxiii) *Buṇḡu-re mid sirma sarna-samaj-ra?* *du<nu>b taḡarao-le-n-a*.
 Bundu-LOC one year Sarna society-GEN sit<NMLZ> open-ANT-ITR-IND
 ‘One day there was a meeting of the Sarna society in Bundu.’
- (xxiv) *kaji-do muṇḡa-ko-a?* *du<nu>b*
 word-TOP Munda-PL-GEN sit<NMLZ>
 ‘The name is for the Mundas’ meeting.’
- (xxv) *en-re seṇa-garaṇ-aka-n-leka-n* *muṇḡa-danḡaṛa-ko-ge=ko*
 there intelligent-highly-CONT-ITR-like-ITR Munda-boy-PL-EMPH=3PL:SUBJ
jama-le-n-a.
 gather-ANT-ITR-IND
 ‘The highly intelligent Munda boys have gathered there.’
- (xxvi) *aliṇ-o ḡaḡḡar Ram Dayal Muṇḡa-lo?=liṇ*
 1DL.EX-TOO Dr Ram Dayal Munda-with=1DL:SUBJ
sen-aka-n-tai-ke-n-a.
 go-CONT-ITR-COP-COMPL-ITR-IND
 ‘I had been with Dr Ram Dayal Munda there.’

- (xxvii) *en-re je jagar-ta-n-ko hindi-te-ge*
 there all talk-PROG-ITR-3PL Hindi-INS-EMPH
jagar-la<pa>?-ta-n-ko ayum-o?-ta-n-a.
 talk-excess<RECIP>PROG-ITR-3PL hear-PASS-PROG-ITR-IND
 ‘All speakers speak and compete to talk in Hindi.’
- (xxviii) *ḍakṭar sayob-a? jagar paṛi teba?-le-n-a,*
 Doctor Lord-GEN talk turn reach-ANT-ITR-IND
 ‘It had been a time to speak, Dr Munda’s turn.’
- (xxix) *do=e? met-a-ko-ta-n-a.*
 then=3SG:SUBJ say-BEN-3PL-PROG-ITR-IND
 ‘Then he is saying something for them.’
- (xxx) *hela, neya-do muṇḍa-ko-a? du<nu>b tana?*
 alas this-TOP Munda-PL-GEN meeting COP
 ‘Alas! This is the Mundas’ meeting.’
- (xxxi) *ar diku-te-ge soben=pe jagar-ja-d-a.*
 and Hindi-INS-EMPH all=2PL:SUBJ talk-INGR-TR-IND
 ‘And everyone speaks in Hindi.’
- (xxxii) *neya-do muṇḍa du<nu>b tana?*
 this-TOP Munda meeting COP
 ‘This is the Mundas’ meeting.’
- (xxxiii) *are=ñ ayum-ja-d-pe-a, hindi-ra? ṭekano bano?*
 well=1SG:SUBJ hear-INGR-TR-2PL:OBJ-IND Hindi-GEN certainty NEG.COP
 ‘I am listening to you but your Hindi is not good.’
- (xxxiv) *ni? neka=e ṭokoe?-le-d-ko-a-te*
 this person like this=3SG:SUBJ scold-ANT-TR-3PL:OBJ-IND-to
 ‘This person had scolded them like this.’
- (xxxv) *muṇḍa-te jagar-ja<pa>gar-se-sen-ja-n-a.*
 Munda-INS talk-talk<RECIP>GO-ITER-INGR-ITR-IND
 ‘Then people start talking in Mundari.’

NOTES

* I would like to express my sincere thanks to the editor, Gregory D.S. Anderson who encouraged me to complete my paper and to Nicholas Evans and Yoshiharu Takahashi who gave me useful comments on the early version of my paper. I am also indebted to my wife, Maki who gave me a lot of data as a native speaker of Mundari. Naturally, the description in this chapter will overlap with my previous work to some extent.

1 According to Gregory Anderson, his Ho informants in Mayurbhanj claim only about 80–85% intelligibility with Mundari which is only a little higher than that for Santali. In my own experience, however, when I spoke to Ho speakers in Mundari in the Chaibasa area they considered me to be speaking in Ho. Thus, the Ho of Chaibasa is obviously closer in the dialect/language continuum to Mundari than the Ho of Mayurbhanj.

- 2 According to Anderson and Zide (2002:55), ‘a minimally bimoraic shape appears to have been obligatory for free-forms of nouns in Proto-Munda and many of its daughter languages’. Our data on Mundari supports this.
- 3 According to Nagaraja (1999), he never reported a tonal contrast in Korku.
- 4 We use the upper bar \bar{V} for the pitch accent here.
- 5 According to our definition, we classified into four types among languages proposed to lack a noun–verb distinction to distinguish morphological from syntactic evidence and to leave open the possibility that word classes distinguishable by morphological criteria could be indistinguishable by syntactic criteria; that is, (i) Omnipredicative languages (all major word classes are able to function directly as predicates without derivation, and with no change of meaning), (ii) Precategorial languages (we will restrict ‘precategorial’ to the case where – as in omnipredicative languages – open-class lexemes can occur in any syntactic position. However, in precategorial languages, it is not possible to state a predicate-type meaning for the lexeme directly; rather there is an increment that is made, according to the functional position it is plugged into), (iii) Broschartian languages (the semantic result of placing lexemes in referring or predicating environments depends not on a high-level word class category like noun or verb, but rather is sensitive to much more specific semantic categories, each characterized by their own particular pattern of semantic incrementation), and (iv) Rampant zero conversion languages (the vast majority of lexical items of a given form may appear in both predicating and referring syntactic environments with no formal signalling of conversion, but unlike in a Broschartian language, the semantic effects of syntactic environment are far less predictable). It is only monoprecategorial languages, from among the types above, that can truly be claimed to lack a noun–verb distinction. As the space of my chapter is limited in this book I cannot touch on the details of this issue. Please see (Evans and Osada 2005a), three commentaries (Croft 2005, Hengeveld and Rijkhoff 2005, Peterson 2005) and our response (Evans and Osada 2005b).
- 6 These categories are due to Klaiman (1986). I give a list of verbs in Osada (1999).
- 7 See my paper (Osada 1991) for details.
- 8 *tuka* may be derived from the noun *tuka* ‘nest’.
- 9 I do not repeat the above-mentioned expressives here.
- 10 This list owes to Haspelmath, Dryer, Gil and Comrie (eds) (2005) *WALS*.
- 11 *ke-ate* and *ke-n-ate* are slightly different. The former is the successive action while the latter is the reverse action. It means they must wash their hands before taking the food.

REFERENCES

- Anderson, Gregory D.S. and John P. Boyle (2002) ‘Switch reference in South Munda’, in M.A. Macken (ed.) *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society 2000*, Tempe: Program for Southeast Asian Studies Monograph Series, Arizona State University, 39–54.
- Anderson, Gregory D.S. and Norman H. Zide (2002) ‘Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint’, in M.A. Macken (ed.) *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society 2000*. Tempe: Program for Southeast Asian Studies Monograph Series, Arizona State University, 55–74.
- Bhaduri, M.B. (1931) *A Mundari–English dictionary*, Calcutta: Calcutta University Press.
- Census of India 1991 (2004) *Language Atlas of India 1991*, General direction by Jayant Kumar Bantia. Delhi: Controller of Publications, Civil Lines.

- Comrie, Bernard (1976) *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*, Cambridge: Cambridge University Press.
- Cook, Walter A. (1965) *A Descriptive Analysis of Mundari: A Study of the Structure of the Mundari Language According to the Methods of Linguistic Science*, PhD Dissertation, Georgetown University, Ann Arbor: Microfilms International.
- Croft, William (1991) *Syntactic Categories and Grammatical Relations: The Cognition of Information*, Chicago: The University of Chicago University.
- (1993) ‘Case marking and the semantics of mental verbs’, in J. Pustejovsky (ed.) *Semantics and the Lexicon*, Dordrecht: Kulwer Academic, 55–72.
- (2005) ‘Word classes, part of speech and syntactic argumentation’, *Linguistic Typology*, 9: 431–441.
- Crystal, David (2003) *A Dictionary of Linguistics and Phonetics*, London: Blackwell.
- Diffloth, Gérard (1976) ‘Expressives in Semai’, in P.N. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austroasiatic Studies Part 1*, Honolulu: The University Press of Hawaii, 249–264.
- (1994) ‘i: big, a: small’, in L. Hinton, J. Nichols and J. Ohala (eds) *Sound Symbolism*, Cambridge: Cambridge University Press, 107–114.
- Emeneau, Murray B. (1980) *Language and Linguistic Area*, selected and introduced by Anwar S. Dil, Stanford: Stanford University Press.
- Evans, Nicholas and Toshiki Osada (2005a) ‘Mundari: The myth of a language without word classes’, *Linguistic Typology*, 9:351–390.
- (2005b) ‘Author’s response: Mundari and argumentation in word-class analysis’, *Linguistic Typology*, 9:442–557.
- Gumperz, John in collaboration with H.S. Biligiri (1957) ‘Notes on the phonology of Mundari’, *Indian Linguistics Taraporewala Memorial Volume*: 6–15.
- Haldar, Rakhal Das (1871) ‘An introduction to the Mundari language’, *Journal of Asiatic Society of Bengal*, 40(1): 46–67.
- Haspelmath, Martin, Matthew S. Dryer, David Gil and Bernard Comrie (eds) (2005) *The World Atlas of Language Structure (=WALS)*, Oxford: Oxford University Press.
- Hengeveld, Kees and Jan Rijkhoff (2005) ‘Mundari as a flexible language’, *Linguistic Typology*, 9: 406–431.
- Hoffmann, John (1903) *Mundari Grammar (=MG)*, Calcutta: Government Press.
- (1930–1978) *Encyclopaedia Mundarica (=EM)*, Patna: Government Press.
- Kachru, Yamuna and Rajeshwari Pandharipande (1980) ‘Toward a typology of compound verbs in South Asian languages’, *Studies in the Linguistic Sciences*, 10(1): 113–124.
- Klaiman, M.H. (1986) ‘Semantic parameters and the South Asian linguistic area’, in Bh. Krishnamurti (ed.) *South Asian Languages Structure, Convergence and Diglossia*, Delhi: Motilal Banarsidass, 179–194.
- Langendoen, D. Terence (1963) *Mundari Phonology*. Unpublished paper.
- (1966) ‘The copula in Mundari’, in J.W.M. Verhaar (ed.) *The Verb ‘Be’ and Its Synonyms*, Dordrecht: Reidel, 75–100.
- (1967) ‘Mundari verb conjugation’, *Linguistics*, 32: 39–57.
- Masica, Colin (1976) *Defining a Linguistic Area: South Asia*, Chicago: University of Chicago Press.
- Munda, Ram Dayal (1971) ‘Aspects of Mundari verb’, *Indian Linguistics*, 32: 27–49.
- (1980) *muṇḍārī-vyākaraṇ* (Mundari Grammar in Hindi), Ranchi: Muṇḍārī sāhitya pariśad.
- Mundu, M.M. (1995) *muṇḍārī ṭuṅkoḍhāri* (Mundari-Hindi-English dictionary), Ranchi: Catholic Press.
- Nagaraja (1999) *Korku Language: Grammar, Texts, and Vocabulary*, Tokyo: Institute for Study of Languages and Cultures of Asia and Africa.
- Nottrott, A. (1882) *Grammatik der Kohl-Sprache*, Gütersloh.

- Osada, Toshiki (1991) 'Notes on convergence features in the Chotanagpur area', in S. Basu-Mullick (ed.) *Unity in Diversity in Chotanagpur*, Delhi: Uppal, 99–119.
- (1992) *A Reference Grammar of Mundari*, Tokyo: Institute for Study of Languages and Cultures of Asia and Africa.
- (1999) 'Experiential constructions in Mundari', in *Gengo Kenkyu: Journal of Japanese Linguistic Society*, 115: 51–76.
- (2007) 'Reciprocals in Mundari', in V. Nedjalkov (ed.) *Reciprocal Construction*, Philadelphia: John Benjamin, 1575–1590.
- Pandharipande, Rajeshwari (1978) 'Exceptions and rule government: The case of the passive rule in Hindi', *Studies in the Linguistic Sciences*, 8(1): 153–173.
- Peterson, John (2005) 'There's a grain of truth in every "myth," or, why the discussion of lexical classes in Mundari isn't quite over yet', *Linguistic Typology*, 9: 391–405.
- Pinnow, Heinz-Jürgen (1959) *Versuch einer historischen Lautlehre der Kharia Sprache*. Wiesbaden: Otto Harrassowitz.
- Prasad, Swarnalata (1973) *munḍārī-hindī śabdakoṣa*. (Mundari–Hindi Dictionary in Devanagari Script), Ranchi: Bihar Tribal Research Centre.
- (1976) *hindī-munḍārī śabdakoṣa*. (Hindi–Mundari Dictionary in Devanagari Script), Ranchi: Bihar Tribal Research Center.
- Schachter, Paul (1985) 'Parts of Speech System', in Timothy Shopen (ed.) *Language Typology and Syntactic Description*, Cambridge, 1:3–61.
- Schmidt, Wilhelm (1906) Die Mon-Khmer Völker, ein Bindeglied Zwischen Völkern Zentralasiens und Austronesian. *Archiv für Anthropologie*, 5:59–109.
- Singh, Udaya Narayan, K. V. Subbarao and S. K. Bandyopadhyay (1986) 'Classification of polar verbs in selected South Asian languages', in Bh. Krishnamurti (ed.) *South Asian Languages Structure, Convergence and Diglossia*, Delhi: Motilal Banarsidass, 244–269.
- Sinha, N. K. (1974) *Mundari Phonetic Reader*, Mysore: Central Institute for Indian Languages.
- (1975) *Mundari Grammar*, Mysore: Central Institute for Indian Languages.
- Suwilai Premsrirat (2002) *Thesaurus of Kham Dialects in Southeast Asia*, Nakhon Pathom: Institute of Language and Culture for Rural Development, Mahidol University at Salay.
- Whitley, J. C. (1873) *Mundari Primer*, Calcutta: Government Printing Office.
- Zide, Norman H. (1960) *Korku Phonology and Morphophonemics*, PhD Dissertation, University of Pennsylvania.
- (ed.) (1966) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton.
- (1978) *Studies in the Munda Numerals*, Mysore: Central Institute of Indian Languages.

KERA? MUNDARI

*Masato Kobayashi and Ganesh Murmu**

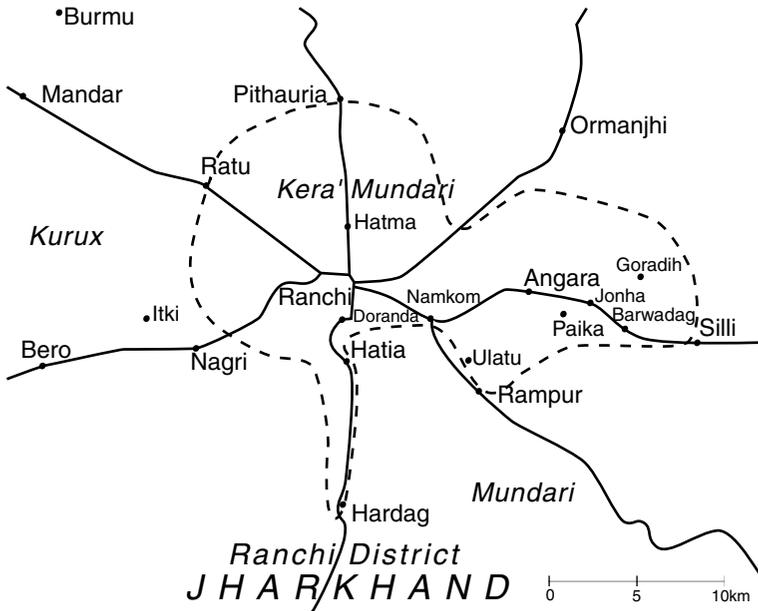
1 INTRODUCTION

1.1 Speakers

Kera? Mundari is a variety of Mundari spoken by the Oraons residing in and around Ranchi, the capital of the new state of Jharkhand. According to their oral tradition, the Oraons once lived in the area between the Son and the Ganges.¹ When they left Rohtas under the attack of invaders,² they moved south along the North Koel river, settled in the Chota Nagpur plateau, mainly in the northwestern part of it. In Gumla, Lohardaga, and Latehar districts, Kurux (Kurukh), the Oraons' original language belonging to the Dravidian family, continues to be in active daily use. In the eastern part of Ranchi district, on the other hand, the Oraons have completely switched from Kurux to Mundari, which was, and to some extent still is, the dominant language of the area. However, their Mundari has unique characteristics and constitutes a distinct regional as well as ethnic dialect.³

It is not clear how long ago it was that Oraons in this area gave up Kurux (if they ever spoke it), and shifted to Mundari.⁴ According to the tradition of the Nagbanshi Raja family, the erstwhile kings of Chota Nagpur, Oraons as well as Mundas were present at the coronation of Phani Mukut Rai, the adopted heir of Madra Munda, manki of Sutiambe.⁵ The implication of this story is that Kera? Mundari-speaking Oraons lived for a long time in an area where the Mundas were the dominant body of residents and where the royal family and its administrative officers spoke a non-Munda language (Sadani or Nagpuri). Although the Oraons now greatly outnumber the Mundas in the areas where they settled, they continue to use Kera? Mundari and usually do not know Kurux. The Oraons living in the neighboring blocks, namely Ratu, Mandar, Bero, and Burmu Blocks of Ranchi district, speak Kurux, but Sadani (Sadri) or Hindi is also commonly used in these areas and serve as a link language between the two groups of Oraons.⁶

Map 4.1 shows the approximate area where Kera? Mundari is spoken, based on the village names which our consultants gave us as Kera? Mundari areas. Except on the northern side where steep slopes separate the Ranchi plateau from Hazaribagh District, there are no geographical boundaries around the area where Kera? Mundari is spoken. On the southern side of the Kera? Mundari area, the Mundas and the Oraons live in separate villages and it seems possible to draw a clear demarcation line, but on the western part, there might be a transitional area of Kera? Mundari-speaking and Kurux-speaking Oraons which remains to be studied in future fieldwork.



MAP 4.1 APPROXIMATE AREA WHERE KERA? MUNDARI IS SPOKEN

The Census of India does not provide a separate entry for Kera? Mundari speakers, and it is difficult to get an exact number of them. According to the 1991 Census of India, the total number of the Oraons in Ranchi district is 1,054,064. By subtracting from it 158,447, which is the number of the people who returned Kurux as their mother tongue, we get 895,617 as the possible maximum number, but this figure undoubtedly includes a large number of Oraons who speak Hindi or Sadani as their first language. Considering that the number of total Mundari speakers in Jharkhand is 667,872, the number of Kera? Mundari speakers probably does not exceed a third or quarter of that.

Kera? Mundari-speaking Oraons are mostly settled agriculturists engaged in paddy and pulse cultivation, or vegetable farming where underground water supplies are abundant. Some of those in urban and suburban settings are employed workers in the public and various other sectors.

1.2 Name of the language

According to a generally accepted etymology,⁷ the name of the language comes from the past suffix *-kera* (pronounced *-ker?a* in careful speech), one of its salient features distinguishing it from other dialects of Mundari which have *-keda* instead.

1.3 History of research

To the best of the authors' knowledge, Kera? Mundari has never been systematically described. Oraon villages near Ranchi are being absorbed by the growing suburbs

of Ranchi, which have been rapidly urbanizing for the past century and especially in recent years. Due to the impact of urbanization on village life, and due also to the low recognition of the language even by the Oraons themselves, Kera? Mundari may fall out of use in a few generations. In this report, we will describe the basic structure of Kera? Mundari, focusing on the differences from Hasada? Mundari⁸ and also on the influence from neighboring languages.

1.4 Taxonomy

Kera? Mundari shares a large part of its basic vocabulary with other dialects of Mundari, and in that respect it can be called a dialect of Mundari. But its verbal morphology, both derivation and inflection, is significantly different from that of the Hasada? dialect, in that it lacks infixation, intricate marking of indirect and direct objects, or noun incorporation, while having suffixes not used in other dialects. Among the dialects of Mundari, it seems to be closer to the Naguri than to the Tamaria or Hasada? dialects. For example, it shares with Naguri the past suffix *-?en-* and genitive suffix *-ra?*. Major features which distinguish Kera? from other dialects of Mundari are the simpler verbal morphology such as the use of *-o?* in both an intransitive and reflexive sense (sections 3.2, 3.2.6 and 3.2.10), the lack of the transitive marker *-d-* (section 2.6), negative forms in *ka-* such as *kano* ‘do not’ or *ka?ale* ‘there is no ...’ (section 3.2.9) and the existential verb *dahin*. Interestingly, it has some features in common with Kherwarian languages other than Hasada? Mundari, such as the possessive formation with *-t-*, for example, *hon-t-?na* ‘own child’, cf. Santali *hon-t-ijn* ‘my child’, or words such as *rama* ‘nail’, cf. Santali *rama* : Hasada? *sarsar*.⁹

Since Kurux is considered to be the original language of the Oraons, we naturally expected to see substratum influence of Kurux on Kera? Mundari. We have not conducted an extensive comparison of their vocabulary yet, but with respect to grammar there are relatively few common features with Kurux, and none of them serves as conclusive evidence of Kurux influence.

- The perfect suffix *-kera*, for example (section 3.2.3), resembles Kurux *kera*, past third singular non-masculine form of the root *kama:* ‘to go’, which is often added to a past form of a Kurux verb to emphasize completion of an action. However, since Kera? Mundari has other perfect forms with *k-*, such as *-kia* (third person singular animate object) and *-kuka* (plural), borrowing from Kurux is rather unlikely.
- While other dialects have a distinct personal pronoun *ae?* ‘he, she’ for the third person singular, Kera? Mundari uses a demonstrative pronoun *ini(?)* instead (section 3.1.6). The same holds true of Kurux *as* ‘he, that man’, but also of Indo-Aryan like Hindi *vah* ‘that, he, she’.
- While a participial clause serves as a relative-type construction in Hasada? Mundari, Kera? Mundari uses neither a participial clause nor a relative pronoun of Munda origin, and simply repeats a demonstrative pronoun. Kurux and Malto also have the same type of relative construction (section 4.2.3).
- Neither second dual nor plural forms, which are used in an honorific sense in Ho and Hasada? Mundari respectively, are used for a singular entity in Kera? Mundari and Kurux (section 3.2.1).
- Kera? Mundari almost completely lacks prefixation (such as the Hasada? causative *a-*) and infixation (such as the Hasada? reciprocal *-pV-*) in its derivational morphology

(section 3.1.10, 3.2 and 3.2.10). Kurux, as well as Indo-Aryan languages like Hindi, also uses only suffixation in derivation.

- As in Kurux or Hindi, Kera? Mundari does not show flexible use of nouns as verbs as is found in Hasada? dialect (section 3.2.10).
- While the glottal stop occurs only after a vowel in Hasada?, in Kera? Mundari /ʔ/ tends to occur between a consonant and a vowel (section 2.3). This is a possible context of /ʔ/ also in Kurux.

2 PHONOLOGY

2.1 Vowels

Kera? Mundari has five vowels, /a/, /e/, /i/, /o/ and /u/ as does Hasada?. Although there is no phonemic contrast in vowel length, lengthening occurs in monosyllabic words, for example, *ub* ‘hair’ or *bir* ‘forest’.¹⁰ There are vowel sequences as in Hasada?, for example, *au-* ‘bring’, *nia* ‘this’, but they do not pattern as phonemic diphthongs. Nasalized vowels are found both in Indo-Aryan loanwords such as *kūāsa* ‘fog’ and in Munda words such as *mej*⁷ *mūhaɽ* ‘face’.

/a/ is sometimes fronted and raised to /e/ or /i/ before /ɲ/: ¹¹ /aɲ-aʔ/ ‘my’ > *epaʔ*, *inaʔ*, /sen-kan-a-ɲ/ ‘I am going/have been to’ > *senkanɛɲ*, /lel-ku-a-ɲ/ ‘I saw them’ > *lelkuiɲ*. /e/ and /o/ are also often raised, -*ge* ~ -*gi* (emphatic particle), *nia* ‘this’ vs. Hasada? *nea*, -*ko* ~ -*ku* (plural marker).

2.2 Suprasegmentals

Kera? Mundari has no phonemic tone. For stress and intonation, see section 2.5.

2.3 Consonants

Kera? Mundari has the same set of consonants as Hasada? (Osada, this volume) except that the former has a checked palatal stop /jʔ/ which the latter lacks, for example, *mej*⁷ ‘eye’ (Hasada? *med*⁷), *mūj*⁷ ‘ant’ (Hasada? *mui*⁷). However, it may not necessarily be an archaism, for Kera? Mundari /jʔ/ does not seem to show a consistent correspondence to other Mundari dialects or Kherwarian languages:

	Kera?	Hasada?	Ho	Santali
‘eye’	<i>mej</i> ⁷	<i>med</i> ⁷	<i>med</i> ⁷	<i>met</i> ⁷
‘come’	<i>hej</i> ⁷	<i>hij-uʔ</i>	<i>huj</i>	<i>hec</i> ⁷
‘ant’	<i>mūj</i> ⁷	<i>mui</i> ⁷	<i>muʔi</i>	<i>mūc</i> ⁷ , <i>muɽ</i> ⁷

An archaism Kera? consonants show is that intervocalic /h/ is not lost as in Hasada?. This is common to Kera?, Naguri, and Santali. For example, Hasada? *baa* ‘flower’: Kera?, Naguri *baha* : Santali *baha*.

In Hasada? Mundari, /ʔ/ is found only in post-vocalic, syllable-final position, often accompanied by an echo vowel following it. On the contrary, Kera? prefers /ʔ/ to be post-consonantal and syllable-initial; for example, Kera? *setaʔa* ‘morning’ : Hasada? *setaʔ*; Kera? *aɲ-ʔa* ‘my’ : Hasada? *aɲ-aʔ*; Kera? *torʔe* ‘ash’ : Hasada? *toroeʔ*; Kera? *nahʔa* ‘now’ : Hasada? *naʔ*; Kera? *lahʔi* ‘belly’ : Hasada? *laiʔ*; Kera? *bohʔo* ‘head’ : Hasada? *boʔ*, cf. Santali *bohok*⁷.

Note that /-hʔ-/ is a possible cluster in Kurux, for example, *cahʔoi* ‘you will want’. Change of /h/ in loanwords to /ʔ/ seems to have occurred in both Kera? Mundari and Kurux for the Indo-Aryan verb *rah-* ‘be’ (Sadani *rah-*), Kera? *raʔ-(a)e* ‘is’, Kurux *raʔi*: present non-masculine 3SG. ‘is’. Etymological /ʔ/ is often lost in Kera? Mundari, for example, Kera? *ini(?)* ‘he’ : Hasada? *iniʔ* ‘that person’, or *-ki(?)a* (past suffix with 3SG. animate object) : Hasada? *-kiʔia*, while in careful speech adventitious /ʔ/ is freely inserted to make a morpheme boundary clear.

While the distinction between /d/ and /t/ is neutralized in Hasada?, it is preserved in Kera? as well as in Tamaria and Naguri Mundari, for example, *gaʔa* ‘river’, *oʔoʔ* ‘furthermore’, *huʔij* ‘small, little’ (Osada 1992:27).

Double /r/ is often pronounced single as in other dialects, for example, /sahar-re/> *sahare* ‘in a city’.

2.4 Syllable structure and phonotactics

(C)V and (C)VC are the most typical shapes of a syllable, and a maximal syllable is CV₁V₁C.

2.5 Intonation/stress

As in other languages of Chota Nagpur, word stress is nondistinctive in Kera? Mundari. Although it is an impressionistic observation and is not confirmed by an acoustical analysis yet, iterative rhythm of high and low pitch is not as conspicuous as in Hasada? Mundari, and sounds somewhat flat as in Santali.

That the vowel of a monosyllabic word is lengthened (section 2.1) suggests an existence of a foot structure. Typically, a nominal or a verbal phrase constitutes its own stress unit, which often begins low unless the initial syllable is a verbal root, and becomes high in the penultimate syllable, or ends high if the word is disyllabic. We have not noticed any sentential stress other than the rising tone at the end of yes-no questions.

2.6 Morphophonology

At a morpheme boundary, morpheme-final /n/ is sometimes deleted when it is followed by /ʔ/, for example, *kanʔale* ‘is not’ ~ *kaʔale*, /sen-ʔena/ ‘went’ > *seʔena*. Root-final stops of some verbs are also replaced by /ʔ/, for example, /dub-akan-a-el/ ‘is sitting’ > *duʔakanae*, /sab-ku-m/ ‘catch them!’ > *saʔkum*.¹²

Vowel harmony is not an active phonological process in Kera? Mundari. It is found in lexicalized items such as *kuri* ‘girl’ : *kora* ‘boy’. In a few verbal forms, umlaut is observed between a suffix vowel and a vowel of an immediately preceding root, such as *lil-i-m* see-OBJ.3SG.INAN-2SG.IMP ‘look’ : *lel-ker-a* see-PST-IND ‘looked’, or *hinʔ-i-a* be-3SG.AN-IND ‘he is’ : *henʔ-a* be-IND.INAN ‘it is’. When there is a variation between high- and mid-vowels within a paradigm, a high one tends to be generalized by paradigmatic leveling, for example, *nii~nia* ‘this’ (Hasada? *nea*), plural marker *-ku* (Hasada? *-kul-ko* as in *si-ku* ‘lice’ : *hon-ko* ‘children’). Such heightening is apparently a postlexical phenomenon, for it does not trigger further umlaut,

for example, *lel-li-a* ‘saw him’, *em-ki-a-ku* ‘they gave to him/her’, or *lel-e-a-ku* ‘they see’.

Kera? Mundari has the intransitive marker /n/ as in Hasada?, but transitive marker /d/ in Hasada? does not find its equivalent in Kera?, possibly except in /r/ in the past endings *-kera* and *-tara*, which are used when there is no explicit object. The /r/ in these sequences might be a result of flapping of /d/ in original **-ke-d-a* or **-ta-d-a*, which was then grammaticalized as *-ker-a* or *-tar-a* because transitive /d/ is suppressed elsewhere in verbal inflection. The past suffix *-ke-* becomes *-ki-* when the object is third person animate, *-ki-n-* for first singular object, etc. (section 3.2.2), and there is no /ʔ/ which Hasada? has probably as a slot holder of the elided transitive suffix *-d-* (cf. Osada 1992:100f.).¹³ Note that etymologically unexplainable /r/ also appears in the genitive suffix *-ra?* (vs. Hasada? *-a?*, cf. section 3.1.2).

3 MORPHOLOGY

3.1 Nominal morphology

3.1.1 Number

As in Hasada? Mundari, Kera? has three numbers, singular (unmarked), dual (*-kin*), and plural (*-kul-ko*). The number marker is often unattached both after human and non-human nouns, especially when number is expressed elsewhere as in (1):

- (1) *tin tʰo: hon*
 three CLSSFR son
 ‘three sons’
- (2) *baria seta pusi-ke sa?ker-a-e*
 two dog cat-ACC catch-PST-IND-3SG
 ‘Two dogs caught a cat.’
- (3) *ro-ko seγγελ-re boloy-en-a-e*
 fly-PL fire-LOC enter-PST-IND-3SG
 ‘Flies flew into the fire.’

For non-human entities as in the last two cases, it is ungrammatical to use plural verb forms according to the judgment of some of our consultants (cf. example 60 in 3.2.9).

3.1.2 Case

In Kera? Mundari, postpositions are not always preceded by oblique (i.e. genitive) forms of the words they follow, and there is no clear distinction between postpositions and case endings. Kera? Mundari shares three case suffixes, *-(r)a?*, *-re*, and *-loy* with other dialects, and other suffixes such as Hasada? *-sa?* ‘toward’, *-ta?* ‘near’, or *-ate* ‘from’ were not used by our consultants.

The genitive suffix *-a?* as in Hasada? is found only in pronominal forms such as *an-?a* ‘my’, and *-ra?* is used for nouns in general. *-ra?* is used in the Naguri dialect as well (Cook 1965:204). It might come from *-re-a?*, or from *-a?* with secondary /r/. In any case, the initial /r/ of this suffix effectively keeps the boundary of the stem-final syllable intact, and makes *-ra?* more like a postposition than a suffix. *-re* is a locative suffix as in Hasada?, for example, *seγγελ-re* ‘in the fire’. The comitative suffix *-loy* as

in *ini-loŋ* ‘with him’ or *reŋgao-loŋ* ‘walking’ is a Kera? equivalent of Hasada? *-lo?*. One of our consultants also returned *-ta* as a dative case ending, as in

- (4) *nia kicri am-ta-ŋ au-tar-a*
 this cloth YOU-DAT-1SG bring-PRF-IND
 ‘I brought this cloth for you.’

Other case functions are expressed with postpositions of which multisyllabic ones form separate stress units. While accusative is not marked on the substantive in Hasada?, Kera? has borrowed Sadani *-ke* (*-ko* in Hindi), used when the referent of the nominal is animate.

3.1.3 Person

Three persons are distinguished in Kera? Mundari pronouns and verbal inflection. There are inclusive and exclusive forms in the first person dual and plural depending on whether the speaker is included in the reference.

3.1.4 Definiteness

Kera? Mundari has no marker of definiteness. The only morpheme related to definiteness is the suffix *-t-* which introduces a possessor pronominal suffix, for example, *hon-t-aŋa* ‘my/own child’. This suffix is also used in Santali, but more extensively than in Kera? Mundari, as in Santali *hon-t-iŋ* ‘my child’, *oŋak²-ta-le* ‘our house’, etc.

3.1.5 Class/gender

In section 3.1.1 we noted that plural subject suffixes are not used to refer to multiple non-human entities in Kera? Mundari. In Hasada? Mundari, on the contrary, all multiple animate entities take plural forms. On the other hand, Mundari makes no grammatical distinction between male and female-neuter genders as does Kurux. As far as concord with verbs is concerned, nouns are grouped into the following classes:

	<i>Subject marking</i>	<i>Object marking</i>
Hasada? Mundari	animate/inanimate	animate/inanimate
Kera? Mundari	human/nonhuman	human/nonhuman
Kurux	masculine/non-masculine	(female/nonhuman)

3.1.6 Pronouns

For the third person singular, Hasada? has a personal pronoun *ae?* ‘he, she’ (Table 4.1). Kera? uses demonstrative pronouns *ni(?)* ‘this one’, *ini(?)* ‘that one’, *ini hoŋo* ‘that person’, etc. and never *ae?*. This might be due to influence from Kurux or Indo-Aryan languages such as Sadani, which both use the demonstrative as third-person pronouns. The person/number suffixes of Kera? Mundari are the same as in Hasada?, except that the glottal stop of the third person singular *-e?* is often dropped and the third person plural is usually *-ku* with a high vowel: 1sg. *-ŋ*, 2sg. *-m*, 3sg. *-e(?)*, 1du.incl. *-laŋ*, 1du.excl. *-liŋ*, 2du. *-ben*, 3du. *-kin*, 1pl.incl. *-bu*, 1pl.excl. *-le*, 2pl. *-pe*, 3pl. *-ku(l-ko)*.

TABLE 4.1: PERSONAL PRONOUNS

	Singular	Dual		Plural	
		Inclusive	Exclusive	Inclusive	Exclusive
First person	<i>aŋ</i>	<i>alay</i>	<i>aliŋ</i>	<i>abu</i>	<i>ale</i>
Second	<i>am</i>	<i>aben</i>		<i>ape</i>	
Third	<i>ini(?)/ni(?)/ ani(?) etc.</i>	<i>inkin</i>		<i>inku</i>	

Interrogatives:

- ciŋo, cilo* ‘when’ (Hasada? *cimtaŋ*).¹⁴
oŋ^hai ‘where’ (Hasada? *kore, kote*)
okoi ‘who’ (Hasada? *okoe*)
cikan ‘what’ (Hasada? *cana*)
oka, cikan ‘which’ (Hasada? *okoa*)
ciŋa, cikan le ‘why’ (Hasada? *cilikate*)
cimti, ciŋi ‘how much, how many’ (Hasada? *ciminaŋ*)
cilka ‘how’ (Hasada? *cilika*)

Relative and indefinite pronouns. Aside from Indo-Aryan relative pronouns, Kera? Mundari uses demonstrative pronouns in the relative construction (see section 4.2.3), and not interrogative pronouns as in other dialects of Mundari. Indefinite pronouns are *jeta, jetan* ‘some’, and *jetae* ‘somebody’ as in Hasada?, for example.

- (5) *jetan ka?a-le-a*
 some NEG.COP-SG-IND
 ‘There is nothing.’

3.1.7 Demonstratives and other deictics

While Hasada? Mundari has different forms for demonstrative pronouns and adjectives, for example, *en* adj. ‘that’: *ena* inanimate pron. ‘that’, Kera? uses forms with *-a* both as an adjective and a pronoun. The distinction between animate and inanimate classes is not strict (Table 4.2). In animate forms without a final glottal stop, the vowel is lengthened.

Examples of demonstrative pronouns are as follows: *ina kahani* ‘that story’ ~ *ini hoŋo* ‘that person, he’, *nia gaŋa* ‘this river’ ~ *ni: kisan* ‘this farmer’.

Deictic adverbs are formed from the stems *ni-* ‘this’, *en-* ‘that’ and *(h)an-* ‘that’ (remote). Place deixis: *neman, neta?i, neŋe, noŋe* ‘here’, *eman, enta?i, eŋŋe* ‘there’ (Hasada? *neta?, henta?*), *(h)aman, (h)anta?i* ‘yonder’. When a speaker is calling someone’s attention, *nilim* ‘look, here’, *inilim* ‘look there’, *(h)anlim* ‘look yonder’ are used, for example, *anlim seno?-ra?-e* ‘Look! There he goes.’

Direction: *neter* ‘this way’, *enter* ‘that way’, *(h)anter* ‘that way yonder’ (*nesa?, ensa?, hensa?*). Degree: *niti?* ‘this much’, *enti?* ‘that much’. Comparison: *ninka* ‘like this’, *enka* ‘like that’.

TABLE 4.2: DEMONSTRATIVES

	Proximate		Middle		Remote	
	Animate	Inanimate	Animate	Inanimate	Animate	Inanimate
Adjectives/pronouns	<i>ni(?)</i>	<i>nia</i>	<i>ini(?)</i>	<i>ina</i>	<i>ani(?)</i>	<i>ana</i>
Interjections	<i>nilim</i>	<i>nilim</i>	<i>inilim</i>	<i>inilim</i>	<i>anlim</i>	<i>anlim</i>

3.1.8 Numerals

While original Munda cardinals are preserved in other dialects relatively well, our consultants of Kera? Mundari returned only *miyad* ‘one’ and *baria* ‘two’ of Mundari origin, while all other cardinals and ordinals are loanwords from Hindi. For collectives, we could find only *barno* ‘both (people)’.

Kera? Mundari uses of Indo-Aryan (Sadani) classifiers, *tʰo:* for human as well as non-human and inanimate entities, and never *jan*, a classifier for humans used in neighboring languages: *tin tʰo:* *daru* ‘three trees’, *tin tʰo:* *hon* ‘three sons’. *go:* is used for animate entities: *du: go:* *kaṭea* ‘two mice’.

3.1.9 Adpositions

As we noted in section 3.1.2, the distinction between a case ending and a postposition is not always clear in Kera? Mundari (or many other Munda languages for that matter); in this section, we will give those which are of Indo-Aryan origin or form their own stress domain.

Preposition: We have found only one preposition, *bin*, *bina:* ‘without’ (< Indo-Aryan).

- (6) *q̃n-ke bin kaji ge ini se-ʔen-a-e*
 I-ACC without talk EMPH she go-PST-IND-3SG
 ‘She went without talking to me.’

Postposition

ke: animate object marker (< Sadani, = Hindi *ko*)

- (7) *q̃n ini-ke ka-ʃ sari-ʃ*
 I he-ACC not-1SG know-1SG
 ‘I don’t know him.’

Note the double subject marking in the form in (7); similar doubled subject forms in emphatic discourse (often with negatives) are found in Mayurbhanj Ho (Anderson *et al.*, this volume) as well.

le ‘for’ (< Sadani): *am-a?* *le* ‘for you’, *kiriṅ le* ‘in order to buy’.
lagin ‘for’ (< Sadani).

Dative use of the quotative particle *mente* as in Hasada? was not observed (cf. ex.(4)).

- (8) *nia kicri? am-a? lagin au-tar-a*
 this cloth you-GEN sake bring-PRF-IND
 ‘(I) brought this cloth for you.’

se, te ‘from, since; than’: *seta? se* ‘since morning’, *aŋ-a? te maraŋ* ‘older than I’.
segin ‘from’

- (9) *aŋ miɖil iskul aŋ-ʔa hatu segin paʔ^hao-en-ij*
 I middle school I-GEN village from study-PST-1SG
 ‘I went to the middle school from my village.’
- (10) *ini apne segin koʔa nam-ki-a-e*
 she self from boy get-PST.OBJ.3SG.AN-IND-3SG
 ‘She found a husband herself.’

3.1.10 Derivation

In Hasada? Mundari, the infix *-nV-* derives a noun from a verbal root, for example, *rakab* ‘climb’ ~ *ranakab* ‘slope’. This formation is not active in Kera? Mundari, but there are lexicalized *-nV-* derivatives like *jono?* ‘broom’ to *jo?* ‘sweep’. Infixation is not found in verbal morphology either (section 3.2).

Compound nouns and adjectives are formed by juxtaposing two nouns in normal syntactic order:

<i>hatu hoʔo-ko</i>	village people-PL = ‘villagers’
<i>haku utu</i>	fish curry = ‘fish curry’
<i>mid loʔa da?</i>	one cup water = ‘a cup of water’

Words of similar meaning are sometimes combined (section 6):

<i>boɖoj’ ili</i>	rice.beer rice.beer = ‘rice beer’
<i>deota boŋga</i>	god(Indo-Aryan) spirit = ‘god’
<i>parab buru</i>	festival(Indo-Aryan) festival = ‘festival’ (possibly to make clear it is not <i>buru</i> for ‘mountain’)

However, Kera? Mundari does not make use of the verb+noun compounds typical of many South Munda languages (Anderson 2007).

3.1.11 Adjectives

An adjective takes the indicative suffix *-a* when it is used as a predicate, for example, attributive: *marəŋ daru* ‘a big tree’ vs. predicative:

- (11) *nia daru gapa marəŋ-a*
 this tree tomorrow big-IND
 ‘This tree will be big tomorrow.’ [not ‘become big’]

Adjectives cannot be used as verbs meaning ‘become ...’, ‘make ...’ as in Hasada?; so for the verb meaning ‘to become big’, a different etymon, *harao*, is used, and not *marəŋ-o?* as in Hasada?.

Comparison is expressed by an adjective preceded by what is to be compared, plus the postposition *-sel-te* ‘from’, for example, *aŋ-a? te marəŋ* ‘older than I’. For superlatives, an Indo-Aryan phrase *sob kui se* ‘than anyone (else)’ is used:

- (12) *ini sob kui se marəŋ hin-ʔia* (or *marəŋ-a-e?*)
 he all someone than big be-3SG.AN big-IND-3SG
 ‘He is the biggest/eldest one.’

-ren ‘belonging to ...’ derives an adjective from a noun, for example, *hatu-ren munḍa* ‘village master’.

Verbs can also be used as verbal adjectives as in *hejo? hapta* ‘next (lit. coming) week’, or *hejo?-ra? hapta* ‘id.’ with genitive suffix *-ra?* functioning as an adjectivizer. They can also combine with the Indo-Aryan suffix *-wala*:

- (13) *da? hejo? wala hen-a?*
 water come ADJV be-3SG.INAN
 ‘It is about to rain.’

3.1.12 Adverb(ial)s

Most adverbs the authors have come across are Indo-Aryan loanwords, such as *bahut* ‘much, very’, *ḡyada* ‘much, very’, *ḡaldi* ‘fast’, *aste* ‘slowly’, *acanak* ‘suddenly’, etc. Deictic adverbs are mostly of Munda origin, for example, *ena le* ‘therefore’, *neter* ‘this way’, etc. (section 3.1.7). An adverb can be derived by adding *te* ‘from’, for example,

- (14) *ṡ^hi:k te ayum-e-a-pe*
 good from listen-OBJ-IND-IMP.2PL
 ‘You guys listen carefully.’

By adding *-re*, as in *pahle (-re)* ‘before’, *ba:d (-re)* ‘after’ both used with genitive forms of verbs in the formation of adverbial (temporal) subordinate clauses:

- (15) *aṡ do siṡḡi dubuc²-ra? pahle-re uri?-ku ader-ku-a-ṡ*
 I FOC sun sink-GEN before-LOC COW-PL bring.in-OBJ.3PL-IND-1SG
 ‘I drove the cattle in before the sun set (i.e. “before the sinking [of the] sun”).’
- (16) *rā:ci hejo?-ra? ba:d*
 Ranchi come-GEN after
 ‘after coming to Ranchi’

The suffix *-loṡ* is a converbal marker denoting simultaneity, for example, *nir* ‘run’: *nir-loṡ* ‘running’.

3.2 Verbal morphology

In Hasada? Mundari, the possible maximal structure of a verb is

(prefix *a-* + (root + infix *-pV-*)) + voice {-*en*, *-o?*} + (benefactive *-a* + indirect object) + (tense + object) + mood + subject. (cf. Cook 1965:140)

In comparison with this, Kera? Mundari has a simpler structure with fewer slots and affixes:

(root + *o?*) + {_A(benefactive *-a* + indirect object) or _B(tense + object)}¹⁵ + mood + subject

The present authors have not come across any verb form in which the bracketed elements A and B co-occur. Thus when a present form with a benefactive suffix is turned into a past form, the past suffix takes over the second slot and the benefactive suffix is suppressed. For example,

isin-a-m-a-ṡ COOK-BEN-2SG-IND-1SG ‘I will cook for you’ vs. *am-a? le isin-ker-a-ṡ*
 you-GEN for COOK-PST-IND-1SG ‘I cooked for you’.

Verbal roots in Kera? Mundari do not take the causative prefix *a-* or the reciprocal infix-*pV* (section 4.6).

	Hasada?	Kera?
<i>jom</i> ‘eat’	<i>a-jom</i> ‘feed’	<i>k^hiao</i> (<Indo-Aryan)
<i>dal</i> ‘hit’	<i>da-pa-l</i> ‘hit each other’	<i>dal-o?</i> (see section 4.6 for <i>-o?</i>)

The suffix *-o?* forms an intransitive verb. For example, the root *uɖuŋ* ‘take out’ is used in a transitive sense in *uɖuŋ-ata-ŋ-me* ‘take ... out for me!’, with *-o?*, it takes on an intransitive meaning, for example, *uɖuŋ-o? ra?e* ‘(It) is coming out’. *-reta-* is attached to form a causative verb, for example (73) *uɖuŋ-reta-ku-iŋ* ‘I make them take out’.

3.2.1 Subject

The person and the number of the subject are either marked at the end of a finite verb, at the end of a word immediately preceding a finite verb, or often at both.¹⁶

- (17) *hatu-ren muŋɖa ar pahan hec²-len-a-kin*
 village-ADJ village.master and priest come-ANT-IND-3DL
 ‘The village master and the pahan have come.’
- (18) *am maŋɖi-m k^hiao-ki-a*
 you food-2SG feed-PST.OBJ.3SG.AN-IND
 ‘You made (her) eat food.’
- (19) *ini-ku lel-ki-a-ku*
 he/she-3pl see-PST.OBJ.3SG.AN-IND-3PL
 ‘They saw him/her.’
- (20) *aŋ a:lu ka-ŋ jom-e-ŋ*
 I potato not-1SG eat-IND-1SG
 ‘I don’t (=can’t) eat potatoes.’

In addition to the same subject markers as in Hasada? (section 3.1.6), *-le* seems to be used as a third person singular suffix: *kaji-ke-le* ‘he says’, *ka?-a(n)-le-a* ‘there is no ...’.

Non-human dual or plural subjects (and objects) are usually not marked as such on the finite verb (7):

- (21) *baria seta pusi-ke sa?-ker-a-e*
 two dog cat-ACC catch-PST.INAN-IND-3SG
 ‘Two dogs caught a cat.’

In Hasada?, this would be *baria seta-kin pusi sab-ki?a-kin*, with the dual ending *-kin* and with no accusative ending, at least with older speakers (Osada 2000:46f.). In Kera?, on the other hand, the object is explicitly marked by the postposition *ke*, and no ambiguity is incurred by not using a dual ending (cf. section 3.1.1) or animate object marker *-?-*.

Kera? Mundari does not use plural forms for a single person in an honorific sense, as is often done in Hasada? Mundari and in some Indo-Aryan languages, nor dual forms as in Ho. Among neighboring languages, Kurux also lacks the honorific plural.

Some culturally important natural phenomena are treated as animate as in Hasada? Mundari, and not as inanimate as in Kurux.

- (22) *siŋgi dʒuŋguc²-en-a-e*
 sun sink-PST-IND-3SG
 'The sun has set.'

3.2.2 Object types

Hasada? Mundari, at least in the speech of the older generation, lacks dative and accusative case marking in nominal phrases, and personal suffixes on the verb function as markers of those cases. Kera? Mundari uses *-ke*, a dative/accusative case-marker of Indo-Aryan origin, and does not use the verbal indirect object marker *-a-* except in non-past forms such as (55) *bano-a-m-e-ŋ* 'I will make for you'.¹⁷

First person:

- (23) *mi^hai em-ki-ŋ-a-e?*
 sweets give-PST-OBJ.1SG-IND-3SG
 'He gave me sweets.'

In Hasada?, the benefactive object 'me' (*-a-ŋ-*) comes before the past suffix *-ked-*, that is, *om-a-ŋ-ked-a-e?* 'he gave ... to me'.

Second person:

- (24) *ini aben dono-ke pasand-ke-b(e)n-a-e*
 he you.DL both-ACC like-PST-OBJ.2DL-IND-3SG
 'He liked both of you.'

A corresponding Hasada? form is *suku-a-ben-tan-a-e?* like-BEN-2DL-CONT-IND-3SG.

Third person:

In principle, Kera? Mundari marks the third person singular animate object with *-i-*, which is fused with past suffix *-ke* into *-ki-*.

- (25) *k^hiao-ki-a-ŋ* (Hasada? *ajom-ki-ʔ-i-a-ŋ*)
 feed-PST.OBJ.3SG.AN-IND-1SG
 'I let him/her eat.'
- (26) *em-ki-i-ŋ*
 give-PST.OBJ.3SG.AN-IND-1SG
 'I gave him/her something.'
- (27) *lel-ki-a-ku*
 see-PST.OBJ.3SG.AN-IND-3PL
 'They saw him/her.'
- (28) *lil-i-m* (Hasada? *lil-i²-me*)
 look-OBJ.3SG.AN-2SG
 'Look at him/her!'

In Hasada? *ajom-ki-ʔ-i-a-ŋ* (25), the transitive suffix *-d-* is lost and */ʔ/* is inserted in its place. In Kera? Mundari, */ʔ/* is not observed in this position unless pronounced carefully. For the possibility that Kera? Mundari originally lacked the *-d-* suffix, see section 3.2.3.

The element *-e-* is used for an inanimate object (cf. Osada 1992:87 *jom-e-a-e?*):

- (29) *hon-ku ʃi:vi-ku lel-e-a-ku*
 child-PL TV-PL see-OBJ-IND-3PL
 ‘Children watch TV.’
- (30) *am jyada alo jom-e-am*
 you too.much don’t eat-OBJ-IMP.2SG
 ‘Don’t eat so much.’

But *-e-* is also found in verbs referring to an animate object:

- (31) *ini hoʒo cinha-e-a-m*
 this person know-OBJ-IND-IMP.2SG
 ‘Do you know him?’

The third person dual is marked with *-kɲ-*, and the plural with *-ku-* or *-uk-* (see section 3.2.3):

- (32) *sob hon-ke pā:c rupia em-ku-m*
 all child-ACC five rupee give-OBJ.3PL-IMP.2SG
 ‘Give five rupees to each child.’
- (33) *caca caci-ra? beʒa-ko-ke lel-ku-iɲ*
 uncle aunt-GEN son-PL-ACC see-OBJ.3PL-1SG
 ‘I look after my uncle and aunt’s sons.’
- (34) *sipahi baria cor-ke saʔ-ki-kɲ-a-e* (Hasada? *sab-ke-d-kiɲ-a-e?*)
 policeman two thief-ACC catch-PST-OBJ.3DL-IND-3SG
 ‘The policeman caught two thieves.’

The following is an inflectional table (Table 4.3) of object marking for the verb *saʔ* ‘catch’ and the past suffix *-ke*. Since dual objects are relatively rare, singular and plural forms are often found in the place of the dual forms. Subject markers are added either after the final indicative suffix *-a* or after the preceding word (section 3.2.1). For the morpheme boundary of the third person forms, see section 3.2.3.

3.2.3–3.2.7 Tense, aspect, mood, and voice

Tense. The future (or imperfective) tense is expressed by not attaching any aspect suffix, as in Hasada?:

- (35) *ini sen-a-e*
 he GO-IND-3SG
 ‘He will go.’

TABLE 4.3: OBJECT PARADIGM

Object	Singular	Dual		Plural	
		Inclusive	Exclusive	Inclusive	Exclusive
First person	<i>saʔ-ki-ɲ-a</i>	<i>saʔ-ke-lɲ-a</i>	<i>saʔ-ki-lɲ-a</i>	<i>saʔ-ku-bu-a</i>	<i>saʔ-ke-le-a</i>
Second person	<i>saʔ-ke-m-a</i>	<i>saʔ-ke-b(e)n-a</i>		<i>saʔ-ke-pe-a</i>	
Third human	<i>saʔ-ki-a</i>	<i>saʔ-ki-kɲ-a</i>		<i>saʔ-kuk-a</i>	
non-human	<i>saʔ-ker-a</i>	<i>saʔ-ker-a</i>		<i>saʔ-ker-a</i>	

It also functions as habitual present: (29) *hon-ku tivü-ku lel-e-a-ku* ‘Children watch TV.’ The sequence *-e-a* OBJ-IND OR NPST-IND also refers to a non-past action. Whether it is different from just the indicative suffix *-a* in function remains an open question.¹⁸

- (36) *inkin dauṛ suru-e-a-kin*
 3.DL run begin-NPST-IND-3DL
 ‘They start running.’

Aspect markers. *-kan-l-akan-* are equivalent to Hasada? *-akan-*, which Osada (1992:96) describes as having continuous and atelic aspect.¹⁹ They are used only for intransitive verbs in Kera? Mundari, and *ra?* is attached to transitive verbs instead (section 3.2.12).

- (37) *eik ʔho: seta kuṭa-re tol-kan-a-e*
 one CLSSFR dog post-LOC tie-CONT-IND-3SG
 ‘A dog is tied to the post.’

- (38) *ini seta? se tiṅgu-akan-a-e*
 he morning from stand-CONT-IND-3SG
 ‘He has been standing since morning.’

As in Hasada? *-akan-*, *-kan-* is used for experience in the past:

- (39) *am kab^{hi}: raxci: hej[?]-kan-a-m?* (Hasada? *hiju[?]-akan-a-m*)
 you ever Ranchi come-CONT-IND-2SG
 ‘Have you come to Ranchi before?’

- (40) *aṇ kalkatta: kab^{hi}: ka-ṇ sen-kan-e-ṇ*
 I Calcutta ever not-1SG go-CONT-IND-1SG
 ‘I have never been to Calcutta.’

- (41) *sahar-wala kab^{hi}: kab^{hi}: hej[?]-ka[?]-a-ku* (/hej-kan-a-ku/)
 city-ADJV sometimes sometimes come-CONT-IND-3PL
 ‘City people sometimes come.’

-ta-n- and *-ta-r-* are used in perfect or past sense, unlike Hasada? *-ta-n-l-ta-d-* which is used in progressive sense:

- (42) *aṇ maṇḍi ka-ṇ bana-tar-e-ṇ*
 I food not-1SG make-PF-IND-1SG
 ‘I did not prepare food.’

-ke- is a past tense suffix, although it might have the connotation of ‘completion of an action’ as in Hasada? (Osada 1992:95).

- (43) *dulan-ra? ente?-ke ga: em-ki-a-ku* (Hasada? *om-ke-d-a-ko*)
 bride-GEN mother-ACC cow give-PST.OBJ.3SG.AN-IND-3PL
 ‘They gave cows to the bride’s mother.’

- (44) *sukanburu lel-ker-a-ṇ*
 Sukanburu see-PST.INAN-IND-1SG
 ‘I have seen Mt Sukanburu.’

- (45) *ini sen-akan dahin-ke-n-a*
 he go-CONT be-PST.ITR-ITR-IND
 ‘He was going.’

-ka- is probably an allomorph of *-ke-*, for our consultants gave *saʔkukae* as a plural equivalent of *saʔki-a-e*. If *-ku-* in this form refers to the object, it is the only case where an object marker comes before a tense suffix (cf. section 3.2.2).

- (46) *sipahi do cor-ke saʔku-ka-e*²⁰
 policeman two thief-ACC catch-3PL-PST-3SG (tentative analysis)
 ‘The police caught two thieves.’
 (Hasada? *sab-ke-d-kij-a-e?* catch-PST-TR-OBJ.3DL-IND-3SG)
- (47) *siŋgi ɖuŋguc²-ka ina-ra? ba:d-re*
 sun sink-PST that-GEN after-LOC
 ‘after the sun sets (= the sun set, after that)’

-ka- can also be used by itself to denote non-preterite perfect, as in *durum-ka-e* ‘he has/will have fallen asleep’, *hej²-ka-e* ‘he has come’. Or alternatively, we can consider the *-k-* part to be the perfect suffix, *-uk-*, *-i-*, *-e(-)r-* 3PL, 3SG animate, and inanimate object markers respectively, and *-a* the normal indicative marker. Although a third person plural object marker *-uk-* is unknown in other dialects, Kera? has another case of it in *kaʔa(n)luka* ‘there are no ...’ (3PL.), and there also *-uk-* can be analyzed to be the third person personal suffix, for the singular form is *kaʔa(n)le(a)*. Synchronically, we think it is best to interpret these forms as follows: The suffix is *-ke-*, of which *e* is deleted when a vowel-initial object suffix follows, for example,

/sab-ke-i-a/	/lel-ke-e(-)r-a/	/sab-ke-uk-a-e/
<i>saʔ-k-i-a</i>	<i>lel-ke(-)r-a</i>	<i>saʔ-k-uk-a-e</i>
catch-PST-OBJ.OBJ.3SG.AN-IND	see-PST-OBJ.3SG.INAN-IND	catch-PST-OBJ.3PL.AN-IND-3SG
‘caught him/her’	‘saw it’	‘he/she caught them’

-en- is a past marker found in Naguri as well as in Kera?:

- (22) *siŋgi ɖuŋguc²-en-a-e* ‘The sun set.’
- (48) *apu-m piŋia se-ʔ-en-a-e* (< /sen-en-a-e/)
 father-2SG market go-PST-IND-3SG
 ‘Did your father go to the market?’
- (49) *hej²-en-a-e*
 come-PST-IND-3SG
 ‘He came.’

-le- denotes an anterior past action:

- (50) *aŋ t^horasa man sunum kirij le piŋia sen-le-n-a-ŋ*
 I a little mustard oil buy for market go-ANT-ITR-IND-1SG
 ‘I first went to the market in order to buy some mustard oil.’
- (51) *ini hoŋ-ke bis baras pahle le-li-a-ŋ*
 that man-ACC twenty year before see-ANT.OBJ.3SG.AN-IND-1SG
 ‘I saw him twenty years ago.’
- (52) *aŋ tanisun ara:m-le-ŋ*
 I a little rest-ANT-1SG
 ‘I will first take a little rest.’

Mood. The only mood suffix we have found so far is the imperative marker of the second person singular *-?m/-me*: *tijgu-n-a?m* ‘stand!’, *sen-a?m* ‘go!’, *em-me* ‘give!’. Normal second person suffixes *-m*, *-ben* and *-pe* are also used to make an imperative form: *lil-i-m* ‘look!’.

Kera? Mundari shows an interesting reduplication of the last syllable of verbs in subordinate clauses when they have a subjunctive function, for example, *idi-di* from *idi* ‘take’, *lik^hae-k^hae* from *lik^h* ‘write’ (see section 4.2).²¹ When a subordinate clause has a connotation of unrealized condition, *hunaŋ* ‘would have ...’ is attached after the fully inflected verb, as in *kaji-ki-a hunaŋ* ‘(I) would have talked to him/her’.

Voice. While the intransitive voice is expressed by attaching *-n-* after a tense suffix as in Hasada?, transitive *-d-* is not found in Kera?. The */r/* in *-kera* or *-tara* might originate from a flapped *-d-*, which is suppressed elsewhere in inflection:

(23) *em-ki-ŋ-a-e* ‘he gave me (sweets)’ (Hasada? *om-a-ŋ-ke-d-a-e?*)

(42) *bana-ta(-)r-a-ŋ* ‘I prepared (food)’ (Hasada? *-ta-d-a-ŋ*)

(44) *lel-ke(-)r-a-ŋ* ‘I saw (the mountain)’ (Hasada? *-ke-d-a-ŋ*)

The suffixes *-kan-l-akan-*, which denote an atelic aspect, are used only for the intransitive voice, and suppletive forms with *ra?*, an auxiliary verb of Indo-Aryan origin, is used for a transitive verb (section 3.2.12).

3.2.8 Finiteness

As in Hasada? Mundari, finite verb forms are usually marked with the indicative suffix *-a*, which is often raised to */e/* or */i/*. In Kera?, verbs are sometimes found without the indicative suffix:

(7) *ka-ŋ sari-ŋ* ‘I do not know.’

(53) *aliŋ do barno uraon tan-liŋ* (Hasada? *tan-a-liŋ*)
 we.DL.EXCL FOC two.people Oraon be-1DL.EXCL
 ‘We two are Oraons.’

(54) *am acc^ha hoŋo tan-me*
 you good person be-2SG.
 ‘You are a good person.’

Verb forms without the indicative suffix *-a* are also found in subordinate clauses, as in the conditional clause of the following sentence:

(55) *am oŋa?-re dahin-ke-n do am-a? le piŋa bano-a-m-e-ŋ*
 you home-LOC be-PST-ITR FOC you-GEN for bread make-BEN-2SG.-IND-1SG
 ‘If you were home, I would have made roti (bread) for you.’

3.2.9 Negation

Kera? Mundari uses negative *ka* ‘not’ and negative imperative *alo* ‘do not’, which are used in Hasada? as well:

(56) *ka hoba-o-a*
 not happen-ITR-IND
 ‘(It) will not take place.’

- (57) *dura-ra?* *samne alo tiḡgu-n-a?m*
 door-GEN near don't stand-ITR-IMP.2SG.
 'Don't stand near the door.'

A unique negative form *kano*, which might be a cross-formation of negative *ka* and *bano?* found in Hasada?, is also used:

- (58) *am-ke jyada maṅḡi kano jo-jom rahen-a*
 you-ACC too.much food don't REDPL-eat should-IND
 'You should not eat that much food.'

ka?alile 'there is no ...', *ka?anlea* 'there is no ...' and *ka?anluka* 'there are no ...' are negative verbs of existence, equivalent to *bano*, *bankoa*, etc. in other dialects:

- (59) *Soma-ra?* *era do oḡa?-re ka?a-li*
 Soma-GEN wife FOC house-LOC NEG.COP-3SG
 'Soma's wife is not home.'
- (60) *b^husunḡi ka?an-luk-a*
 mosquito NEG.COP-3PL-IND
 'There are no mosquitos.'

3.2.10 Derivation

Mundari is known for the flexibility regarding the use of nouns and nominal phrases as verbs (Hoffmann 1903:xxi). Kera? Mundari does not share this feature with other dialects so much, and nouns are used as verbs only in the sense of 'do ...', with or without a verbalizing suffix *-ol-u*, for example, *aḡandi-u-a-le* wedding-VBLZ-IND-1PL 'we conduct wedding' (*aḡandi* 'wedding'), *sindri-u-a-le* vermilion-VBLZ-IND-1PL 'we put vermilion' (*sindri* 'vermilion'), *katam-a* end-IND 'finishes' (*katam* 'end'), etc.

With respect to extending the meaning of verbal roots with affixes, Kera? Mundari does not show nearly as much variety as Hasada? either. Where Hasada? extends a verb with affixation, Kera? almost always uses compound verbs or a different lexical item:

- (61) *ini gitic? seno? ra?-a-e* :: Hasada? *duḡum-o?-tan-a-e?*
 he sleep go be-IND-3SG sleep-RFLXV-CONT-IND-3SG
 'He is falling asleep (lit. going to sleep).'²²
- (62) *k^hiao-ki-a* :: Hasada? *a-jom-ki-?-i-a*
 feed-PST.OBJ.3SG.AN-IND CAUS-eat-PST-TR-OBJ.3SG.AN-IND
 'feed him/her'

In Kera? Mundari, the intransitive suffix *-o?* also has a reflexive function, just like Hasada? *-en*:

- (63) *ini hoḡo apne-ap ge dal-o?* *ra?-a-e* (Hasada? *dal-en-tan-a-e?*)
 that man himself EMPH hit-ITR be-IND-3SG
 'He hits himself.'
- (64) *ti: abum-o?-m*
 hand wash-ITR-IMP.2SG
 'Wash your hands.'

In Hasada? Mundari, reciprocity is expressed by the *-pV* infix. Here also, Kera? uses the intransitive suffix *-oʔ-* in a reciprocal sense, for example, *dal* ‘strike’, Hasada? *da-pa-l* ‘hit each other’, Kera? *dal-oʔ* ‘id.’.

3.2.11 Noun incorporation

While Hasada? Mundari shows a considerable flexibility with respect to incorporating nouns into a verbal form, the authors have been unable to find any such example in Kera? Mundari. Certainly a system of the type found in South Munda languages like Sora is lacking in Kera? Mundari.

3.2.12 Compound verb constructions

Here again, Kera? Mundari does not show as much variety as in Hasada?. Where a compound verb is used in Hasada?, Kera? has a tendency to use an adverb with a simplex verb, for example,

- (65) Hasada? *ciʃi ol-ruaʃ-me* :: Kera? *dobara ciʃ^{hi}i*
 letter write-return-IMP.2SG again letter
lik^ha-em.
 write-IMP.2SG
 ‘Write a letter again.’

The following are a few verbs and verbal modifiers found in Kera? Mundari:

daʃi ‘can’

- (66) *alua ka-ʃ jom daʃi-a*
 potato not-1SG eat can-IND
 ‘I cannot eat the potatoes.’
- (67) *giʃia ka-ʃ senoʔ daʃi-ker-a*
 market not-1SG go can-PST.INAN-IND
 ‘I could not go to the market.’

raʔ progressive: Kera? Mundari forms the progressive present by adding *raʔ*, an auxiliary verb of Indo-Aryan origin, to the root.

- (68) *ini g^humao -taʔ-a-e*
 he stroll -PROG-IND-3SG
 ‘He is strolling.’
- (69) *ini aŋda jom-e-raʔ-a-e*
 he egg eat-OBJ-PROG-IND-3SG
 ‘He is eating an egg.’
- (70) *naha das bajao-raʔ-a*
 now ten strike-PROG-IND
 ‘It is ten o’clock now.’

This is equivalent to Hindi ... *raha: hai*. In form, it looks similar to Sadani *karat rahe*, preterite imperfect ‘was doing’ (Jordan-Horstmann 1969:87), or Kurux *kamckii: raʔi:*, present perfect ‘she has made’ (Grignard 1924:74), but agrees with neither in meaning.

raʔen, rahen ‘should’ (originally ‘be, stay’)

(58) *am-ke kano jojom rahen-a* ‘You should not eat.’

- (71) *ale-aʔ loyoy-re sioʔ k^hojʔa raʔen*
 we-GEN paddy.field-LOC plowing begin should
 ‘We (GEN.) should start plowing our paddy field.’

-tuka is attached to a verbal root and denotes completion of an action:

- (72) *ini inku hejuʔ mente kaji-tuka-e*
 he they come QUOT say-CMPLT-3SG
 ‘He has told them to come.’

-reta forms a causative verb (Hasadaʔ *-rika*):

- (73) *paisa uduy-reta-ku-i-n*
 money bring-CAUS-OBJ.3PL-IND-1SG
 ‘I make some people withdraw money.’

-hataʔ denotes a continuative action:

- (74) *keʔa-hataʔ-ku-am*
 call-CONT-OBJ.3PL-IMP.2SG
 ‘You please keep calling them.’

k^hojʔa ‘start ...ing’ < Sadani *k^hoj-* ‘try, search’’: (71) *ale-aʔ loyoy-re sioʔ k^hojʔa raʔen*
 ‘(We) should start plowing our paddy field.’

cahao ‘would like to’ < Indo-Aryan *cah-* ‘want’:

- (75) *am citʔi piʔa idi cahao-am*
 you how.many bread take want-2SG
 ‘How many pieces of roti (bread) would you like to take?’

duk^ʔ ‘begin, start’

- (76) *inku hoʔo naha ge maŋʔi jom duk^ʔ-akan-a-ko*
 those person now EMPH foot eat begin-PF-IND-3PL
 ‘They have just started eating.’

3.3 Expressives

Echo-word construction, which is common in neighboring Indo-Aryan languages and in the Dravidian languages Kurux and Malto, is found in a few expressions such as *maŋʔi uŋʔi* ‘food and the like’ (*maŋʔi* ‘food’), *sabji-ubji* ‘vegetables and the like’ (*sabji* ‘vegetable’), *jaldi paldi* ‘quickly’ (< Hindi *jaldi*). Forming reduplicants with *u* is not unknown in other Kherwarian languages, such as Naguri Mundari *piʔa uʔa* ‘bread and the like’ (*piʔa* ‘roti bread’).

4 SYNTAX

4.1 Simple sentence

The following are examples of equative sentences:

- (77) *nia bahut ũca daru tan-aʔ*
 this very tall tree be-INAN
 ‘This is a very tall tree.’

- (78) *ini ram tan-i?*
 he Ram be-3SG.AN
 'He is Ram.'
- (79) *bugi ge men?-jn-a*
 fine EMPH exist-1SG-IND
 'I am fine.'²³

Yes–no questions can be introduced by a question marker *ici?am*, which is not obligatory.

- (80) *ici?am apu-m bajar se?-en-a-e*
 QM father-2SG market go-PST-IND-3SG
 'Did your father go to the market?'
- (81) *ni am-a? pustak tan-a?*
 this you-GEN book be-IND.INAN
 'Is this book yours?'

A possessive relationship is expressed by verbs of existence, with the possessor in the genitive or sometimes in the dative/accusative case in *-ke*.

- (82) *du: be?ta: tin be?ti: (kuri-hon) aŋ-a? mena?-ku-a*
 two son three daughter I-GEN exist-3PL-IND
 'I have two sons and three daughters.'
- (83) *ini-ra? miyad kuri-hon mena-i-a*
 he-GEN one daughter exist-3SG.AN-IND
 'He has a daughter.'
- (84) *aŋ-a? bahut kami mena-? = aŋ-ke bahut kami hen-a?*
 I-GEN much work exist-IND.INAN I-ACC exist-IND.INAN
 'I have a lot of work.'
- (85) *aŋ-a? du: kitab hen-a?, kalam do ka?a-le*
 I-GEN two book exist-IND.INAN pen FOC NEG.COP-3SG
 'I have two books, but I don't have a pen.'

4.2 Complex sentence

A main clause seems to come first in a complex sentence.

- (86) *am b^hula-en-a-m, aŋandi cito hoi-a (seki)*
 you forget-PST-IND-2SG wedding when be-IND COMP
 'Did you forget when the wedding will take place?'
- (87) *ram-ke kaji-m mia ci?hi*
 Ram-ACC say-IMP.2SG one letter
lik^ha-e-k^hae (See section 3.2.3 for *lik^ha-e-k^hae*)
 write-OBJ-REDPL
 'Tell Ram to write a letter (lit. that he should write one letter).'

4.2.1 Coordination

Coordinate conjunctions used in Kera? Mundari are *oḍo(?)* ‘and further’, *ar* ‘and’ and *baki* ‘but’ (< Urdu):

- (88) *ini piyaṅkar hin?-i-a baki bes hoḍo hin?-i-a*
 he drunkard be-3SG.AN-IND but good person be-3SG.AN-IND
 ‘He is a drunkard, but he is a good person.’

Converbs in *-kotor* or *-koto* are more common for clauses with temporal sequence:

- (89) *ina kahani-ke ayum kotor बहुत k^hus-le-n-a-ṅ*
 that story-ACC hear CV much be.happy-ANT-ITR-IND-1SG
 ‘I heard that story and became happy.’ (Hasada? *ayum-ked-ci*)
- (90) *ḡn-ke bi:n kaji kotor ge ini seḍ-en-a-e*
 I-ACC without talk CV EMPH she go-PST-IND-3SG
 ‘She went without talking to me.’

4.2.2 Complement clauses

We could not collect enough examples of complement clauses to draw any general rule. Kera? Mundari uses *seki* to introduce a complement clause, for example, (86) *am b^hula-en-a-m, aḗandi cito hoi-a seki* ‘Did you forget when the wedding will take place?’

4.2.3 Relative-type clauses

In Hasada? Mundari, relative clauses are made with interrogative pronouns such as *okoe* ‘who’, and are placed after main clauses. In Kera? Mundari, on the other hand, relative clauses are placed first, and demonstrative pronouns are used instead of interrogatives:

- (91) *ini korahon hola susun-ke-n-a ini*
 this boy yesterday dance-PST-ITR-IND he
 ‘He is the boy who danced yesterday.’
- (92) *ini hoḍo hola hec²-le-n-a-e cinha-e-a-m*
 that person yesterday come-ANT-ITR-IND-3SG know-OBJ-IND-2SG
 ‘Do you know the man who came yesterday?’ (Hasada? *hola hiju?-le-n hoḍo*)
- (93) *ini laḗki ba:re jagar-ke-n-a-le ini laḗki ka-e hej-kan-a-e*
 that girl about talk-PST-ITR-IND-1PL that girl not-3SG come-CONT-IND-3SG
 ‘That girl whom we were talking about has not come.’²⁴

Kurux (Grignard 1924:288, 292) and Malto (Droese 1884:40f.) have a similar relative construction, in addition to the one with interrogative pronouns, for example,

- (94) *Kurux* (Grignard)
endr is a:s-im malk-as okk-ar
 QM this.MASC.SG that.MASC.SG-EMPH NEG.COP-MASC.3SG sit.CV
temba:lagyas a:s
 beg.PST.PROG.MASC.3SG that.MASC.SG
 ‘Is not this the man who used to sit and beg for alms?’

Indo-Aryan relative pronouns are also used.

Another element found in complex clause structures in Kera? Mundari is *jesan*: ‘as, like’ (used here as a postposition):

- (95) *aŋ am-a? jesan jaldi-jaldi: nir ka-n daŋi-n*
 I you-GEN like fast run not-1SG can-1SG
 ‘I cannot run as fast as you do.’

jab: relative adverb, ‘when’:

- (96) *jab aŋ japan seŋera-e-n ci^hi ge lik^hao-i-n*
 when I Japan arrive-IND-1SG letter EMPH write-IND-1SG
 ‘When I arrive in Japan, I will write a letter.’

4.2.4 Subordinate clauses

In Hasada? Mundari, participial forms followed by the locative case suffix *-re*, optionally marked by the topic marker *do*, function as adverbial clauses of time and condition, for example, *jagar-ken-re-do* ‘while I was talking’. In Kera? Mundari, where simple coordination is preferred to a participial construction, clauses usually contain a finite verb.

Temporal clause:

- (97) *siŋgi duŋguc²-ka ina-ra? ba:d-re bahar alo sen-a?m*
 sun sink-PRF that-GEN after-LOC outside don’t go-IMP.2SG
 ‘Don’t go out after the sun has set.’
- (98) *seta? hui-a hulaleke baŋi-tuka-m*
 morning become-IND before leave-CMPLT-2SG
 ‘Leave before dawn.’
- (99) *abu jagar-ke-n-a-bu du ini hej²-en-a-e* (Hasada? *jagar-ken-re-do*)
 we.INCL talk-PST-ITR-IND-1PL FOC he come-PST-IND-3SG
 ‘He came while we were talking.’

Conditional clause:

- (100) *aŋar aŋ sari-ker-a-n ki am oŋa²-re dahin-a*
 if I know-PST-IND-1SG that you house-LOC be-IND
am-ta-n piŋa pako-a-m-i-n hunay
 you-for-1SG bread cook-BEN-2SG-IND-1SG would.have ...
 ‘If I knew you would be home, I would have baked roti (bread) for you.’

Equivalents of English clauses with ‘so that’ (Hindi *jisse*) are hard to find in Kera? Mundari, but the following sentences are close in meaning:

- (101) *aste aste jagar-me aŋ sam^hao sakabe*
 slowly slowly talk-IMP.2SG I understand will.be.able (I-A)
 ‘Please speak slowly so that I can understand.’ (Hasada? *bujao-leka*)
- (102) *eŋgen du²-en-a-e maŋdi jom-e-a-e*
 mother sit-PST-IND-3SG food eat-OBJ-IND-3SG
 ‘Mother sat down (so that) she will (be able to) eat.’

5 SEMANTICS/DISOURSE

Some verbs, like *jom* ‘eat’, are repeated, as in *jom jom samapt-ker-a-e* ‘He has finished eating.’ The first one in such repetitions seems to be a cognate object, like English *dream a dream*.

- (103) *dura-ra? samne-re tijgu? alo tijgu?-n-am* (~57)
 door-GEN proximity-LOC stand don’t stand-ITR-IMP.2SG
 ‘Don’t stand near the door (at all).’
- (104) *jagar alo jagar-am*
 talk don’t talk-IMP.2SG
 ‘Don’t talk (and keep silent).’

6 LEXICON

The present authors have not conducted any systematic study of Kera? Mundari lexicon except the list of 272 words in Kobayashi *et al.* (2003), but according to our limited research, Kera? Mundari lexicon consists of a Mundari core vocabulary and a part which is common with Santali, such as Kera? Mundari and Santali *rama* ‘nail’ vs. other Mundari dialects *sarsar*. Borrowing from Indo-Aryan languages like Sadani is more extensive than any of the other Mundari dialects. There are also compounds of Indo-Aryan and Munda words, like *deota boyga* ‘god’, *parab buru* ‘festival’ (section 3.11, Kobayashi *et al.* 2003:357). Although Kera? Mundari speakers identify themselves ethnically as Oraons, we found very few cases of Kurux words left in their language. Two possible examples are

Hasada? *mera* ‘rice boiled for fermenting’: Kera? *jagu* ‘id.’: Kurux *jagu* ‘id.’,
 Malto *ja:gu* ‘boiled rice’,
 Hasada? *sikinji* ‘mosquito’: Kera? *b^husunḡi* ‘id.’: Kurux *b^hūsḡi* ‘id.’

7 BRIEF ANALYZED TEXT

This is a spontaneous narrative about wedding and the first Sarhul festival after marriage, told by Ms Birsi Minj of Hatma, our main consultant.

- (i) *baplao-a-le aḡandi-u-a-le.*
 marry-IND-1PL wedding-VBLZ-IND-1PL
 ‘We conduct a wedding.’
- (ii) *aḡandi-ra? ba:d-re paḡi aḡe-kotor sunum sindri-u-a-le.*
 wedding-GEN after-LOC mat spread-CV oil vermilion-VBLZ-IND-1PL
maḡḡi khiao-a-le
 meal feed-IND-1PL
 ‘After the wedding we spread a mat and apply oil and sindur [on it]. [Then] we give meal.’
- (iii) *tab se-le ina-ra? ba:d-re boḡoj’-ili nu-a-le.*
 then from-1PL this-GEN after-LOC rice.beer-rice.beer drink-IND-1PL
 ‘Then after that we drink rice beer.’

- (iv) *neka nia-ke nia-ke sagʔ-a-le, duray-e-a-le.*
 this way this one-ACC this one-ACC hold-IND-1PL sing-OBJ-IND-1PL
jom-e-a-le enejʔ-a-le, susun-a-le.
 eat-OBJ-IND-1PL dance-IND-1PL dance-IND-1PL
 ‘This way we hold this and that person, sing, eat, dance and dance.’
- (v) *cilka nu-e-a-ko cilka jom-e-a-ko lel-e-a-m.*
 how drink-OBJ-IND-3PL how eat-OBJ-IND-3PL see-OBJ-IND-2SG
 ‘You see how they drink, how they eat?’
- (The new couple is invited to the first Sarhul of the bride’s village after their marriage.)
- (vi) *sarhul-re pahan-ke da? duli-a-le.*
 Sarhul-LOC Pahan-ACC water pour-IND-1PL
 ‘At Sarhul, we pour water on the Pahan priest.’
- (vii) *pahan dub-kan-ge-a iya-re, akʰara-re sarna-re.*
 Pahan sit-CONT-EMPH-IND somewhere-LOC Akhara-LOC Sarna-LOC
 ‘Pahan sits somewhere, Akhara or Sarna.’
- (viii) *to ini-ke sunum idi-di mid loʔa da? idi-koto*
 then he-ACC oil take-REDPL one metal.cup water bring-CV
sen-a-le.
 go-IND-1PL
 ‘Then we go after taking oil and one metal cup of water.’
- (ix) *da? dul-koto bad-re ini-ke tir dʰanuf saʔ-e-a-ko.*²⁵
 water pour-CV after-LOC he-ACC arrow bow hold-OBJ-IND-3PL
 ‘After they pour water on the Pahan, they hand over bow and arrow to him.’
- (x) *akʰara-re gʰumao-e-a-ko.*
 Akhara-LOC walk.around-OBJ-IND-3PL
 ‘They walk around in the Akhara.’
- (xi) *tab ini dusra hulai oʔo? hejʔ-a ini.*
 then he other day again come-IND he
 ‘Then he (the Pahan) comes again on another day.’
- (xii) *baha kʰosao hejʔ-a-e.*
 flower put.on.the.back.of.the.ear come-IND-3SG
 ‘He comes to put the Sal flower on the back of (our) ear.’
- (xiii) *nawa haʔa? ar mid haʔa? baha dahin-a.*
 new winnow and one winnow flower be-IND
 ‘There are a new winnow and a winnowful of flowers.’
- (xiv) *ini-ke apan oʔa? idi-a-ku.*
 he-ACC self house take-IND-3PL
 ‘The villagers take him to his own house.’
- (xv) *oʔo ini hejʔ-a-e dosar hulai.*
 and he come-IND-3SG second day
 ‘And he comes on the next day.’

- (xvi) *ini oḍo baha dura dura k^hosao-e-a,*
 he more flower door door put-OBJ-IND
 ‘He puts more flowers on every door.’
- (xvii) *ini-ke e:k rupia ḍe:ḍ^h rupia kaca em-a-i-a.*
 he-ACC one rupee one.and.half rupee money give-DAT-OBJ.3SG-IND
 ‘To him (they) give one rupee or one and a half rupees.’
- (xviii) *oḍo? sim puja-e-a-ko.*
 and chicken sacrifice-OBJ-IND-3PL
 ‘And they sacrifice a chicken.’
- (xix) *tab ina ge niti? maraṅ maraṅ sim puja-e-a-ko.*
 then this EMPH extent big big chicken sacrifice-OBJ-IND-3PL
 ‘They sacrifice a very big chicken like this.’
- (xx) *inku-ke rayḡ lagao-ku-a-le.*
 they-ACC color put-OBJ.3PL-IND-1PL
 ‘We put color (with soil) on them (=the bridegrooms).’
- (xxi) *ina-ra? ba:d-re inku-ke saikil em-ku-m gaṛi*
 that-GEN after-LOC they-ACC bicycle give-OBJ.3PL-2SG car
em-ku-m.
 give-OBJ.3PL-2SG
 ‘After that give them bicycles (and) cars.’
- (xxii) *ini nawa dulhai hej^ʔ-kan-a-e kano.*
 this.one new bridegroom come-CONT-IND-3SG no
 ‘The new bridegroom is visiting, isn’t he?’
- (xxiii) *ini-ke am job^hi k^husi hen-a job^hi em-ku-m.*
 he-ACC you whatever happiness be-IND that.much give-OBJ.3PL-2SG
 ‘To him (lit. them) you give as much as you are happy [having a new son-in-law].’
- (xxiv) *ar du ro:j^ʔ-ra? ba:d-re bida-ku-a.*
 and two day-GEN after-LOC see.off-OBJ.3PL-IND
 ‘And after two days they see them off.’

The following is a well-known story from Aesop’s Fables, *kac^hua ar k^hargos kahani* ‘the story of the tortoise and the hare’, told by Ms Pushpa Kacchap of Doranda in a flowing narrative style.

- (i) *e:k t^ho: kac^hua oḍo k^hargos dahin-kan-a-kin.*
 one CLSSFR tortoise and hare be-CONT-IND-3DL
 ‘There was a tortoise and a hare.’
- (ii) *in-kin a:pas-re bahut laṭ^hao-kan-a-kin.*
 they-DL self-LOC much fight-CONT-IND-3DL
 ‘Between themselves they were fighting a lot.’
- (iii) *ke: aṅ bese am-a? se bese nir daṛi-ṅ kahi-ke-le.*
 COMP? I well you-GEN from well run can-1SG say-PST-3SG
 ‘Saying “I can run faster (lit. better) than you.”’

- (iv) *ke: to k^hargos kaji-re-a, dol, tab alaj.*
 COMP? then hare say-NPST-IND²⁶ INTERJ then we.INCL
re:s lagao-e-a-laj
 race conduct-OBJ-IND-1DL.INCL.
 ‘Then the hare says “Come on, then let us do a race.”’
- (v) *to kac^hua kaji-ra?-e, t^hik hen-a?*
 then tortoise say-PROG-3SG good be-3SG.INAN
 ‘Then the tortoise says “Okay”.’
- (vi) *dol, tab ek jagah inkin jagah-kin nijcit-le-a-kin,*
 INTERJ then one place they-DL place-3DL fix-ANT-IND-3DL
 ‘Okay, then, they fixed one place’,
- (vii) *oɔo dauʔ lagao-e-a-kin.*
 and run set.to-OBJ-IND-3DL
 ‘And they set themselves to race.’
- (viii) *in-kin dauʔ suru-e-a-kin ek jaga se.*
 3.DL run begin-OBJ-IND-3DL one place from
 ‘They start running from one place.’
- (ix) *to k^hargos, ea, jo:r se nir-a-e.*
 then hare INTERJ power with run-IND-3SG
 ‘Then the hare, well, he runs powerfully.’
- (x) *hã:, oɔo kac^hua, na, d^hire: d^hire: reyɔao-loŋ se-?en-a-e.*
 INTERJ and tortoise INTERJ slowly walk-with go-PST-IND-3SG
 ‘Well, and the tortoise went walking slowly.’
- (xi) *to k^hargos socao-ra?-e.*
 then hare think-PROG-3SG
 ‘Then the hare is thinking.’
- (xii) *dola kac^hua to q^he:r doya-re hin?-ia.*
 INTERJ tortoise FOC very far.back-LOC be-3SG.IND
 ‘Look, tortoise is very far back.’
- (xiii) *se qn tanisun aram-le-ŋ kahi-ke-le.*
 so I a.little rest-ANT-1SG say-PST-3SG
 ‘He said “So, I first take a little rest.”’
- (xiv) *ek t^ho: daru-ra? lata(r)-re gitic⁷-ka-e.*
 one CLSSFR tree-GEN below-LOC lie.down-PF-3SG
 ‘He has lain down under a tree.’
- (xv) *to ini gitic⁷-ka-e ki ini duʔum-ka-e.*
 then he lie.down-PF-3SG COMP he sleep-PF-3SG
 ‘Then he has lain down and then fell asleep.’
- (xvi) *itna duʔum-ka-e, ki k^hargos kac^hua ayur-ka-e.*
 so sleep-PF-3SG COMP hare tortoise pass-PF-3SG
 ‘He has slept so soundly that the tortoise has passed the hare.’

- (xvii) *oɔo?* *k^hargos* *acanak* *ini-ra?* *duɽum* *k^hulao-a* *to*
 and hare suddenly he-GEN sleep cause.to.open-IND then
lel-e-a.
 see-OBJ-IND
 ‘And the hare suddenly wakes from the sleep and then sees him.’
- (xviii) *are* *kac^hua* *do* *ɖ^here:* *duɽ* *sen-?en-a-e.*
 INTERJ tortoise FOC very far go-PST-IND-3SG
 ‘Oh! the tortoise has gone very far.’
- (xix) *aŋ* *do* *neta?* *ge* *hin-a-ŋ.*
 I FOC here EMPH be-IND-1SG
 ‘I am only here.’
- (xx) *to* *acanak* *birid-a,*
 then suddenly stand-IND
 ‘Then suddenly he gets up.’
- (xxi) *ekdam* *nir-loy* *sen-o?*,
 at once run-with go-REFLV
 ‘At once he goes running.’
- (xxii) *to* *lel-e-a* *to* *kac^hua* *enta?* *sida-re* *gen.*
 then see-OBJ-IND then tortoise there first-LOC EMPH
teba-kan-ge-a-e
 arrive-CONT-EMPH-IND-3SG
 ‘Then he sees the tortoise has arrived there first.’

NOTES

* We wish to thank our Kera? Mundari consultants, Ms Birsi Minj, Ms Pushpa Kacchap, Ms Vina Kacchap, Ms Asha Tirkey and Ms Vinita Kacchap, Naguri, Tamaria, and Hasada? Mundari consultants Ms Louni Kandulna, Mr Bisheshwar Munda, and Mr Birendra Soy. We thank Dr Ram Dayal Munda and Dr Toshiki Osada for their advice on Mundari in general, Dr Gregory D.S. Anderson for encouraging us to work on Kera? Mundari and for giving us illuminating suggestions on the drafts of this paper, and Dr John Peterson for reading the final draft of this paper and giving us many invaluable comments. This research was conducted under the auspices of the Grant-in-aid (No. 14710379) of the Japanese government. The recordings of Kera? Mundari we made in our fieldwork are available at <http://www.hakuoh.ac.jp/~masatok>, accessed on 29 October 2007.

1 Roy (1912[1995:77]).

2 As is often the case with Oraon folklore, the Mundas have a similar legend. See Roy (1912[1995:64]). See Peterson’s article in the present volume for a Kharia version of this story.

3 A few non-Oraons living in Oraon villages may speak Kera? Mundari, but it is primarily an ethnic dialect and there are no Munda villages we know of where Kera? Mundari is used.

4 If Kera? Mundari came to share some common features with Santali (section 1.4) through contact with the Santals, it must have been spoken when they still lived in

- their original homeland, the present-day Hazaribagh District, which adjoins Ranchi District to its south.
- 5 Roy (1915:30).
 - 6 These two groups of the Oraons share the same totemic system, and marriage between them is also considered to be acceptable. Cf. the case of Gutob and Konekor Gadabas in Griffiths's article in the present volume.
 - 7 Hoffmann, *EM* 8, p. 2294, s.v. *kera?-Munḍa*.
 - 8 The reason why we compare Kera? with Hasada? is that the latter is the best described dialect. Hasada? is also a dialect least influenced by neighboring Indo-Aryan languages, although not necessarily the most conservative, and makes a good contrast with Kera? Mundari.
 - 9 In Kherwarian languages, dialects often share common words and phonological features across language boundaries. See Kobayashi *et al.* (2003:341ff.)
 - 10 For possible Proto-Munda origin of the minimal word length of two morae, see Anderson and Zide (2002).
 - 11 According to Anderson (personal communication), fronting before /ɲ/ is also found in Mayurbhanj Ho.
 - 12 This verb is pronounced /sag/ and not /sab/ in careful speech (e.g. in Text A, 4). See Osada (1992:29) for the complementary distribution of /g/ and /ʎ/.
 - 13 Actually, one of our consultants said *-ki?a* in careful speech, but as she inserted /ʎ/ in other morpheme boundaries as well, we interpret that the /ʎ/ here is a non-etymological morpheme separator.
 - 14 Cf. the temporal suffix *-to* in Santali, for example, *nito* 'now'.
 - 15 See section 3.2.3 for a possible case of a direct object marker coming before a tense suffix.
 - 16 Double subject marking is found in a few other Kherwarian languages/dialects such as Mayurbhanj Ho (Anderson, this volume), cf. example 7 in Section 3.1.9.
 - 17 A few other Kherwarian languages such as Hazaribagh Korwa also mark only primary objects in their verbal forms (Anderson 2007).
 - 18 According to Anderson (2007), *-e* is found as a future marker in Bhumij, some Ho varieties, and in some South Munda languages such as Kharia, Juang or Gta?/Didey as well. The *-e-* in example (36) might also be an object marker referring to *dau?* 'running'.
 - 19 In examples (37) to (40), 'perfect' might better describe the meaning of *-kan/-akan-*.
 - 20 In the pronunciation of one of our consultants, this form sounds more like *sa?kukoe*. Singular and plural object forms are freely used to denote an object in the dual number (section 3.2.2). Another possibility John Peterson suggested (personal communication) is that *-ka* comes from past suffix *-ke* and indicative marker *-a*, so *sa?kukae* is analyzed *sa?-ku-k-a-e* catch-OBJ.3PL-PST-IND-3SG.
 - 21 There is at least one case of free-form reduplication, *jom jom* 'meal' from the verb *jom* 'eat'.
 - 22 Probably a calque of Hindi *sone ja: raha: hai*.
 - 23 One consultant also said *kami hena-ɲ-a* work be-1SG-IND 'I have work (to do)' in spontaneous speech. This sentence is yet to be cross-checked with other speakers.
 - 24 In other dialects of Mundari, a speaker of the Tamaria dialect judged these sentences to be understandable, and made the following sentence, in which a demonstrative pronoun is used but the relative clause comes after the main clause: Tamaria Mundari *en kuri ke lil-i-pe ini? duraŋ-tan-a* this girl ACC see-OBJ.3SG.AN-2PL she sing-PROG-IND 'Look at the girl who is singing.'
 - 25 Or *saga-e-a-ko*.
 - 26 *-re-a* is probably the same as the non-past suffix *-e-a* (section 4.3).

REFERENCES

- Anderson, Gregory, D.S. (2007) *The Munda Verb. Typological Perspectives*, Trends in Linguistics Studies and Monographs 174, Berlin: Mouton de Gruyter.
- Anderson, Gregory D.S. and Norman H. Zide. (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, 55–74.
- Cook, Walter Anthony (1965) *A Descriptive Analysis of Mundari*, PhD Dissertation, Georgetown University.
- Droese, Ernest (1884) *Introduction to the Malto Language*, Agra: Secundra Orphanage Press.
- Grignard, A. (1924) *A Grammar of the Oraon Language*, Calcutta: Catholic Orphan Press.
- Hoffmann, John B. (1903) *Mundari Grammar*, Calcutta: Bengal Secretariat Press.
- (1950–1979) *Encyclopaedia Mundarica*, 16 vols, Patna: Superintendent, Secretariat Branch Press, (Reprinted by Gyan Publishing House, New Delhi, in 1998).
- Jordan-Horstmann, Monika (1969) *Sadani: A Bhojpuri Dialect Spoken in Chotanagpur*, Wiesbaden: Harrassowitz.
- Kobayashi, Masato, Ganesh Murmu and Toshiki Osada (2003) 'Report on a preliminary survey of the dialects of Kherwarian languages', *Journal of Asian and African Studies*, 66: 331–364.
- Osada, Toshiki (1992) *A Reference Grammar of Mundari*, Tokyo: ILCAA, Tokyo University of Foreign Studies.
- (2000) *Munda-Go Kyōhon*, Tokyo: ILCAA, Tokyo University of Foreign Studies.
- Roy, Sarat Chandra (1912) *The Mundas and Their Country*, Calcutta: Kuntaline Press, (Second reprint by Catholic Press, Ranchi, in 1995).
- (1915) *The Oraons of Chota Nagpur: Their History, Economic Life, and Social Organization*, Ranchi: Published by the author at the Bar Library.

CHAPTER FIVE

HO AND THE OTHER KHERWARIAN LANGUAGES

Gregory D.S. Anderson, Toshiki Osada, and K. David Harrison

1 INTRODUCTION

Strictly speaking, we have no consensus which ‘languages’ are to be considered Kherwarian. The Kherwarian language continuum consists of, according to Norman Zide and his mid-1960s Munda languages project group at the University of Chicago, Asuri, Bhumij, Birhor, Ho, Korwa, Mundari, Santali, and Turi. According to the most recent edition of the *Ethnologue* (Gordon 2005), Kherwarian has 12 languages; that is, Asuri, Birhor, Koda, Ho, Korwa, Mundari, Mahli, Santali, Turi, Agariya, Bijori (Binjhia, Birjia), and Koraku. *The International Encyclopedia of Linguistics* (Bright 1991) follows their classification, in which, Bhumij is not considered a separate language, but rather a dialect of Mundari, whereas Agariya, Mahli, Koda, and Koraku are considered to be separate languages.

The language/ethnolect distinction in India in general is complex and in the case of the Kherwarian varieties, particularly so. Many factors contribute to this: Some language designations are questionable because of the confusion of ethnic or *jati*¹ names with language names. For example, ‘Mahli’ denotes a socially distinct and endogamous group among the Santal people, whose language appears to be the same as the Santali language. Further, the Asur, Agariya, and Birjia are three subtribal distinctions of one tribe, the Asuri. Binjhia or Birjia in its normal understanding (there may be others) does not appear to belong to the Munda language family at all, but rather to the Indo-Aryan languages according to Prasad (1961:314). Moreover, the description of the Koda language has been done only by Konow (1904) in *Linguistic Survey of India* (henceforth *LSI*). Munda (1968:10) demonstrated that Kodas ‘speak the dialect of the neighboring tribe, thus the Kodas of Burdwan and Bankura speak Santali and Bengali and those of Madhya Pradesh speak Kurukh’. As far as Koraku is concerned, Barker (n.d.:3) points out: ‘The earlier writers did not draw a distinction between Koraku and Korowa, lumping them together as “Korwa”. Koraku is spoken in southern Mirzapur, southwestern Bihar, and northern Surguja, and is more closely related to Santali. Korowa is spoken in southern Surguja, Jashpur, parts of Raigarh, and so on. These two groups must be carefully separated on the basis of mutual unintelligibility, for, though the phonemic systems are almost identical, the morphology and vocabulary vary.’ Although this first-hand information should be respected, to our dismay we do not have a proper description of Koraku so far. At the moment we exclude the Koraku language from the Kherwarian languages here but further work on Koraku is urgently required.² As a whole, it can be said that an understanding of the actual number of

distinct Kherwarian speech varieties that exist, whether judging on formal linguistic or ethnolinguistic/sociolinguistic criteria, remains in its infancy.

The designation Kherwari or Kharwari is used to denote the common stock from which the Santals, the Mundas, the Hos, and so on have sprung, as the Santal myths tell us. Since Sten Konow adopted this term in the *LSI*, we use it for this language group. Descriptions of Santali (Ghosh), Mundari (Osada), and a special variety of Mundari used by a (former) Kurukh population (Kobayashi and Murmu) have been given elsewhere in this book. In this chapter we describe the lesser known ‘major’ Kherwarian language, Ho, and several other ‘minor’ Kherwarian languages (Bhumij, Asuri, Turi, Korwa, and Birhor).

1.1 Ho

Ho speakers are mainly located in the East Singhbhum district of Jharkhand and the Mayurbhanj and Keonjhar districts of Orissa. According to the Census of India, the number of speakers of Ho is as follows: 949,216 in 1991; 802,434 in 1981; 648,066 in 1961; 599,876 in 1951; and 383,126 reported in *LSI*. The most recent figure is from the *Ethnologue*.

The designation of Ho is derived from *hor* ‘human’ in the Ho language. From a linguistic point of view, Ho is quite close to Mundari, although the Mayurbhanj variety that forms the basis of this description is less similar both lexically and structurally to Mundari than the Ho of Jharkhand. While linguistically quite close, the ethnic identity of Ho speakers is, however, quite different from that of Mundari speakers. Similar arguments can be made for many of the other languages, for example, Bhumij or Birhor, Turi, or Asuri.

The Ho call their language *hor: basa* and *ho haçam*, but *ho kaji* is more general. It is in the Mundari-type set of Kherwarian speech varieties. Ho has been written in Romanized transcription, Hindi-based Devanagari, Oriya script and the indigenous Warang Chiti script, devised by the Ho pandit Lako Bodra in the early 1950s (Pinnow 1972, Zide 1996, 2000). Though a standard orthography for Ho has not been fully institutionalized, there are considerable efforts among the Ho intelligentsia of Mayurbhanj, Orissa to promote the Warang Chiti script in educational materials and publications. The authors have been working closely with both the Ho community in Orissa and the Unicode consortium over the past couple of years to help codify and encode the Ho script for computational and Internet purposes (see section 8 and Harrison and Anderson 2007).

1.2 Bhumij

Bhumij speakers are spread across the state of Jharkhand and Mayurbhanj district, Orissa. According to the Census of India, the number of speakers is as follows: 46,680 in 1981; 131,258 in 1961; 101,938 in 1951; 108,230 in 1941 and 79,078 recorded in *LSI*. According to the Census reports, the number of speakers is in decline. As Indian Census statistics reflect ethnolinguistic identity not actual linguistic competence, so as with most other Munda languages, in particular the minor ones, it is impossible to know how accurately these Census numbers reflect reality. One Bhumij speaker from Mayurbhanj estimated that there were 50,000 Bhumij speakers in 2005.

As with all Kherwarian languages, whether Bhumij or any of the other minor Kherwarian languages discussed here (or even Santali and Mundari for that matter) constitute separate languages in a linguistic sense is debatable. In terms of ethnolinguistic identity they are all distinct groups and that is the only criteria that can be meaningfully applied. Relatively minor lexical, phonological, or structural/morphosyntactic differences are attested between any two groups or between 'dialects' (or even individual speakers) of a given (set of) Kherwarian language[s] (e.g. between Mayubhanj and Chaibasa varieties of Ho). We cite Munda's (1968:33) comment here:

One can conclude by saying that Bhumij is a name of a Munda tribe and not of a Munda language. The Bhumij speakers are speakers of a more or less Latar (=Talaria) dialect of Mundari with a heavy borrowing from the local dialects of Indo-Aryan. Culturally, too, as has been pointed out earlier, they are the most acculturated people among the Munda-speaking peoples. As their geography extends to the Santali speaking area, there can be borrowings of certain formal characteristics from the former language, but its overall pattern seems to be Mundari.³

The acculturation process alluded to includes both socioreligious consequences as well as processes of language shift. Bhumij like many other minor Munda languages is rapidly losing speakers to local Indo-Aryan variants, for example, Sadani/Sadri or Oriya.

1.3 Korwa

According to the Census of India, the number of speakers is in the following: 28,286 in 1981; 17,720 in 1961; 26,447 in 1951; 13,021 in 1941; and 20,227 in *LSI*. This number includes Kodaku/Koraku speakers. Korwa speakers have also been erroneously classified as Koraku. Apart from the *LSI*, Philip Barker carried out field work among Korwa people in 1953 and K.C. Bahl of the University of Chicago did field work in 1962. No real linguistic publications have resulted from these and only minimal unpublished source materials exist.

1.4 Asuri

According to Census of India, the speaker population of Asuri was 7,703 in 1981; 4,540 in 1961; 1,510 in 1951; 4,564 in 1941; and 19,641 in *LSI*. The recent figure is 16,596 in 2001 from *Ethnologue* (Gordon 2005). It is impossible to say how many actual speakers of this fairly acculturated Munda group exist.

1.5 Birhor

Birhor is one of the most endangered languages in India. We have the following number for Birhor speakers in the Census of India: 5,950 in 1981; 590 in 1961; 37 in 1951; 2,550 in 1941; and 1,234 in *LSI*. *The Ethnologue* (Gordon 2005) improbably reports 10,000 in 1998.

Osada has done field work among Birhor speakers in 1989 and published the short paper on Birhor (Osada 1993). Birhors are nomadic hunter/foragers. When

settled, they take on the ethnolinguistic identity of the locally dominant group. Such practices further complicate the issues in assessing numbers of speakers, levels of endangerment, or actual discrete (if any) distinctions in communalects or Kherwarian speech varieties.

1.6 Turi

According to the Census of India 1981, the number of Turi people is as follows: 133,137 in Bihar (now Jharkhand); 26,443 in West Bengal; 7,374 in Orissa. This number is not that of actual speakers of Turi but of ethnic or caste members of Turi. Turi is usually considered to belong to the Santali-type end of the Kherwarian continuum, as is shown in the *International Encyclopedia of Linguistics* (Bright 1991: (3)16). Ram Dayal Munda has, however, pointed out the following:

The place of Turi was left undefined in Grierson's *LSI* but we feel that it – along with Asuri, Birhor and Korwa – is now more like Mundari than Santali. In certain respects (e.g. in sharing the same vowels in a few items and in dropping morpheme final vowels in certain forms), however, they look more like Santali than Mundari but they can be derived for the most part as simply from Pre-Mundari. (Munda 1968:i–ii)

However, until better systematic field data are collected from a wide range of Kherwarian varieties and these are coherently and globally analyzed within the context of the ongoing comparative Munda lexical and typological databases, such statements need to be taken with considerable caution. Osada (1991a) has published Father Ponette's field notes on Turi. Anderson and Harrison are planning a field expedition to locate remaining Turi speakers across Jharkhand, northern Orissa, and western West Bengal in the near future.

Generally the discussion in the following sections focuses on Ho and then adds data where relevant and available from the various other Kherwarian speech varieties after the presentation of Ho facts. Sections lacking data from these other varieties should be understood to mean that either there is no significant deviation across these varieties with respect to the particular feature at hand from the perspective of Ho, or that data of this type is simply lacking. Where possible we comment on such.

2 PHONOLOGY

2.1 Vowel inventory

There is a five-vowel system for Mundari-like Kherwarian such as Ho, Bhumij, Asuri, Turi, Birhor, Korwa, (+Koḍa/Kora, Koraku/Koḍaku); some varieties of Santali have more vowels.

(1)		Front	Back
	High	i	u
	Mid	e	o
	Low	a	

Some examples demonstrating the contrasts in these basic vowels are offered in (2).

- (2) Ho
- | | | |
|-------------------------------------|---|-------------------------|
| <i>ko?</i> 'egret, crane' | <i>ku?</i> 'cough' | <i>ka?</i> 'crow' |
| <i>bal</i> 'burn hole into' | <i>bul</i> 'make drunk' | <i>bil</i> 'spread mat' |
| <i>oɾa?</i> 'throw out, get rid of' | <i>oɾa</i> 'bathe' | |
| <i>lin</i> 'press between fingers' | <i>len</i> 'press oil' | |
| <i>gel</i> '10' | <i>gil</i> 'strike with swinging fist (not straight)' | |
- (Deeney 1978a)

Bahl has distinguished between long and short vowels within morphemes for Korwa. But this description seems dubious as Norman Zide (1991:2538) has correctly pointed out. Rather, it appears that syllable nuclei may be short (oral or nasalized) long or glottalized as in Ho; in the last mentioned case, a short echo release vowel is common following the glottal catch. A few minimal pairs exist but the oppositions (length and glottalization of syllable nuclei) are too few to speak of a system, rather they appear to be idiosyncratic features of particular lexemes (and gramemes) in Ho. Some examples of such contrasts include the following:

- (3)
- | | |
|--|--|
| <i>nare</i> 'after a little while' | <i>na?e</i> 'surely(interjection)' |
| <i>ni:</i> 'this one' | <i>nü?</i> 'to open' |
| <i>ne</i> 'this' | <i>ne?</i> 'here (interjection)' |
| <i>oe</i> 'a bird' | <i>o:e</i> 'to tear' |
| <i>jo?</i> 'to sweep' | <i>jo:</i> 'fruit' |
| <i>tɪn</i> 'set fire to firewood' | <i>tɪn</i> 'lift, hold, carry in hand' |
| <i>mui?</i> 'small ant' | <i>mui?</i> 'sink down into non-liquid' |
| <i>gā:</i> 'side shoots of paddy' | <i>ga?</i> 'sew' |
| <i>eto</i> 'teach, know' | <i>eto:[?]</i> 'be feasible' |
| <i>-re</i> 'in' | <i>re?</i> 'to rob' (cf. <i>rē?:</i> 'joy of companionship') |
| <i>kōe</i> 'beg[gar]' | <i>koe?</i> 'gulp' <i>ko?</i> 'egret, crane' |
| <i>he?</i> ~ <i>he:d</i> 'pluck whole leaf' | <i>hē</i> 'chaff' <i>hē?</i> 'a swallow' |
| <i>ram</i> 'brushwood/thorny branches used to protect tree or shut a path' | |
| <i>ram</i> 'having salty taste' (Deeney 1978a) | |

Diphthongs are found in a number of common stems in Ho:

- (4)
- | | |
|------------------------|---|
| <i>hau</i> 'a red ant' | <i>pai?i</i> 'work' |
| <i>goe</i> 'wilt' | <i>goe?</i> 'kill' (cf. <i>gojo?</i> 'die', <i>goe?en</i> 'commit suicide') |
- (Deeney 1978a)

What are considered 'poetic' forms or variants in Ho often reflect older pronunciations when compared with other Kherwarian languages (e.g. Mundari). Thus, some long vowels in Ho are produced by the loss of Proto-Kherwarian **ɪ* (including also the autoethnoym *ho:* < **hoɾo:*).

- (5) *raɾa* = *ra:* 'untie' *raɾe* = *ra:e* 'cool something off' (Deeney 1978a)

There is conditioned but only semi-regular vocalic variation in certain words in different Ho-speaking regions, such that it may be possible to speak of a Northern (Jharkand) and Southern (Mayurbhanj, Orissa [=M(ayurbhanj) H(o)]) dialect. This variation includes harmonic and assimilative variants. For example, some original *i*-forms become *u* when *u* follows in the next syllable (i.e. a kind of leftward spread of roundness) while an *a* following certain high vowels (any *i* sounds and some *u* sounds) fronts to *e* in Mayurbhanj (Orissa) Ho.

- (6) MH
cilike = *cilika*
miyu ~ *muyu* ‘calf’
muṭe[ʔ] = *muṭa* ‘nose’ (Deeney 1978a, Field Notes, KCN)

2.2 Suprasegmental phenomena (tone, register)

Vowels can be nasalized in Ho and Bhumij in the same circumstance as is shown in the chapter on Mundari (Chapter 3). Some morphemes are inherently nasalized, for example, the first vowel in *rāsa* ‘happy’ (KCN, Ho of Orissa). As for allophones, the vowel /a/ is realized as [ə] and [ɛ] after a syllable with a high vowel /i/ or /u/. The stop consonants /b/, /d/, and /g/ are realized as checked consonants; that is, unreleased and preglottalized [b̚] and [d̚], and as the glottal stop [ʔ] respectively in morpheme-final position.

As already briefly alluded to above within the domain of single lexemes, harmonic alternation of vowels in affixes (and in at least one case, possibly stems as well) is also found in Mayurbhanj Ho. Here stems with *i* and some with *u* (i.e. most high vowel stems) require front-vocalism affixes for certain suffixes (for certain speakers at certain times or when using certain registers or communalects). One such affix is the progressive/present in *-tan-* ~ *-ten-*. Compare the two sentences in (7i, 7ii) which were uttered in sequence.

- (7) (a) *cimiŋ ho:ko kaji-ten-e*
 how.many HO-PL speak-PROG-FIN
 ‘how many Ho speak (their language)?’ [KCN]
 (b) *cimiŋ ho:ko jagar-tan-a*
 how.many HO-PL speak-PROG-FIN
 ‘How many Ho speak (their language)?’ [KCN]

This ‘harmony’ can take place with intervening consonants, so there need not be a vowel directly preceding the affix to trigger the alternation.

- (8) (a) *aca okoŋde=m tayi-ten-e*
 so where=2 stay-PROG-FIN
 ‘so where do you stay?’ [KCN]
 (b) *ijŋ jayapura-a? tayi-n-ten-ijŋ*
 I Jaypur-GEN stay-RFLXV-PROG-FIN:1
 ‘I live in Jaypur.’ [DH]

As alluded to above, not only do *-i*-stems trigger this alternation but certain stems with *u*-vocalism as well. Why this is the case remains a subject for future research.

- (9) (a) *jilike=pe huju?-ye-n-e*
 how=2PL come-T/A-ITR-FIN
 ‘How did you come (here)?’ [KCN] 31:30
 (b) *ale-do bas-te=le huju?-ye-n-e*
 WE.PL.EX-EMPH bus-ABL=1PL.EX come-T/A-ITR-FIN
 ‘We came by bus.’ [KCN] 0:03

Another affix that shows this alternation is the first dual inclusive subject marker, =*laŋ/leŋ*. If it attaches to the word preceding the verb in the following spontaneous

utterances with *a*-vocalism, then the vowel is *-a-* but when it attaches to the ‘front’ stem *n/nu[e]* it is realized as *-e-*

- (10) (a) *ca ka=laj juy-e*
 tea NEG=1DL drink:3-IMP tea
 ‘Let’s not drink the tea.’ [DH]
- (b) *ca jue=lej*
 tea drink:3:IMP=1DL
 ‘Let’s drink the tea.’ [KCN]
- (c) *ca=laj juy-e*
 tea=1DL drink:3-IMP
 ‘Let’s drink the tea.’ [KCN]

2.3 Consonant inventory

The consonant system of Ho is straightforward and typical of Kherwarian Munda languages. Perhaps 22 phonemes are to be distinguished. The phonemic status of *ŋ* is dubious as is the opposition of *n* and *ɲ*.

(11)		Labial	Dental	Retroflex	Palatal	Velar	Glottal
	Obstruents:						
	Voiceless	<i>p</i>	<i>t</i>	<i>ʈ</i>	<i>c</i>	<i>k</i>	<i>ʔ</i>
	Voiced	<i>b</i>	<i>d</i>	<i>ɖ</i>	<i>j</i>	<i>g</i>	
	Fricatives		<i>s</i>				<i>h</i>
	Nasals	<i>m</i>	<i>n</i>	<i>[ŋ]</i>	<i><n></i>	<i>ŋ></i>	
	Flaps			<i>r</i>	<i>ɽ</i>		
	Lateral		<i>l</i>				
	Glides	<i>w</i>			<i>y</i>		

Note: [] marginal, allophonic, < > perhaps one segment underlyingly

Examples of words/minimal pairs showing some of these phonemic oppositions in Ho include

- | | | |
|------|-----------------------------------|--|
| (12) | <i>ba-re</i> ‘outside’ | <i>bale</i> ‘young’ |
| | <i>munu</i> ‘beginning’ | <i>mulu</i> ‘new moon’ |
| | <i>muka</i> ‘settlement’ | <i>muga</i> ‘coral’ |
| | <i>bo</i> ‘head’ | <i>mo</i> ‘to swell’ |
| | <i>bir</i> ‘forest, jungle’ | <i>biŋ</i> ‘a snake’ |
| | <i>biɖ</i> ‘to put in the ground’ | <i>bil</i> ‘to spread’ (e.g. a bed or table) |
| | <i>bo:r</i> ‘smooth to touch’ | <i>bo:r̥</i> ‘straw rope’ |
| | <i>cara</i> ‘bait’ | <i>caɽa</i> ‘bald’ |
| | <i>ota</i> ‘press down’ | <i>oɽa</i> ‘open up, expose’ |

Among the characteristic differences seen between Chaibasa Ho which is the variety that has been described generally in the literature and that of Mayurbhanj district, Orissa, where most of the field data for this chapter were collected, includes individual speaker variation in the latter Ho variety between *d* : *j* and *t* : *c*. For example, ‘how do you say’ in Mayurbhanj Ho may be heard as either *tilekepe kadiye* or *cilekepe kajiye*, ‘how many’ can be heard as *timiŋ* or *cimiŋ*, and so on.

- (13) (a) *tileke=pe kadi-ye* vs. (b) *cileke=pe kaji-ye*
 how=2PL say-T/A:FIN how=2PL say-T/A:FIN
 ‘How do you say?’ ‘How do you say?’ [KCN]
- (c) *cimiŋ sirmi=m hoba-ya-n* vs. (d) *timiŋ sirme?*
 how.many year=2 happen-T/A-ITR how.many year
 ‘How old are you?’ [KCN] ‘How old.’ [KCN]

The status of *ʈ* in Ho is unclear. There are sufficient numbers of contrasts to consider *ʈ* phonemically distinct from *t*, *ʈ*, and *ɖ* (e.g. *caʈu* ‘earthen pot’ vs. *caʈu* ‘throw/fall into deep water’), but nevertheless in Ho dialects, there is variation between *ʈ* ~ *t* and *ʈ* ~ *ʈ* in various words, as in the following examples:

- (14) *ʈai* ~ *tai* ‘remain’
peʈelpeʈe ‘pluck twig or small branch with one or both hands’
keʈelkeʈe ‘hard’

Further, *n* and *ŋ* are in near complementary distribution in Ho, with *n* appearing after *i* and a small number of words with *e* (except in Mayurbhanj Ho where *e* arises ‘harmonically’ from *a*, for example, *diyən* for *diyən* ‘country rice liquor’) and appearing as *ŋ* after *o* and *a*, *u* being found in a small number of words with both (and in fact perhaps the only minimal pairs), again speaking to a *phonologically* front vs. back *u*. Neither sound is ever found in onset position. There are a small number of exceptions to this general pattern including the first singular pronoun which for some Ho speakers is *aŋ* [also *aɪn* and *iŋ*] and *germoŋ* ‘fleeting smile’. There are almost no minimal pairs except in some dialects between vocative and first singular possessed forms of certain kinship terms, for example, *apuŋ* : *apuŋ* ‘father’ and the sets *aŋ* ‘I’ : *aŋ* ‘dawn’ and *ruŋ* ‘sensation of having limb asleep’, *ruŋ* ‘to husk’ (cf. *ruŋ* ‘poke, prod’) among a small group of others. There is also plenty of evidence of individual pronunciations, even in spontaneous repetitions of the same word by a single speaker, of *n* alternating with *ŋ*.

- (15) (a) *nui=n* (b) *nui=ŋ*
 drink:3/T:A/IMP=1 drink:3T/A:IMP=1
 ‘I will drink mine.’ [KCN]

Some words in standard Jharkhand (Chaibasa) Ho with *iŋ* have *iŋ* in Mayurbhanj Ho: Chaibasa Ho *barsiŋ* = M. H. > *barsiŋ* ‘two days’ and all ‘day’ combinations < *singi*, for example, *tisiŋ* (not *tisiŋ* ‘today’). As mentioned above, there are almost no words in Ho that show the configurations **aŋ* and **oŋ*. The former is found only in certain pronunciations of the first singular pronoun ‘I’ *a[i]ŋ*, and before phonetic *n* in a small number of words with *-nj-* medial clusters, and the latter in the words *germoŋ* ‘fleeting smile’ and *loŋ*.

Note also the following alternation between *n* and *ŋ*:

- (16) (a) *alay nui-ten-e* (b) *alay nu-tan-a*
 we:2 drink:3-PROG-FIN we.2 drink-PROG-FIN
 ‘We 2 are drinking it.’ [KCN] ‘We 2 are drinking.’ [KCN]

ŋ appears to be non-phonemic as it occurs only before *ɖ* medially and in the word *duna* ‘resin of sal tree used to make incense’ (albeit in a near minimal contrast with *-n-* in *duna[:]* ‘times’). It seems better to treat it as an allophone of */n/* before *ɖ* (and *ʈ*?) and a form that occurs in a small number of mainly loanwords.

Individual speaker variation can be seen between *y* and *j* in intervocalic position in Ho. Indeed, the following two variants were uttered by a single speaker within seconds of each other:

- (17) (a) *hobo-ya-n-a* (b) *hobo-ja-n-a*
 happen-T/A-ITR-FIN happen-T/A-ITR-FIN
 ‘happened’ ‘happened’ [KCN]

(cf. Chaibasa Ho *hoba-*, a variant form also appearing in Mayurbhanj Ho. Mayurbhanj Ho *hobo* likely is from **hoba-o[ʔ]* a form known from other Kherwarian varieties).

2.4 Syllable structure and phonotactics

Words and morphemes may have the following syllabic structural or CV-skeletal shape in Ho.

- (18) V only morphemes like *-a* FIN OR *-a-* BEN
 Vʔ *aʔ* ‘bow’
 V:ʔ *a:ʔ* ‘herbs’ *i:ʔ* ‘excrement’
 CV only morphemes *-ka-* MODAL OR *ka* NEG
 CV: *ba:* ‘flower’ *ti:* ‘hand’
 CV:C *ju:r* ‘smooth’ (cf. *ju:r* ‘crowd around’)
 CVC *bijŋ* ‘snake’
 VC *am* ‘you’
 VCV *ape* ‘you (PL)’
 CVCV *bulu* ‘thigh’
 VCCVC *enɖel* ‘meal left overs’
 CVCVC *daʔob* ‘press compactly’, *maray* ‘big’
 CVCCVC *halmaɖ* ‘salt lick’, *hambud* ‘embrace’
 CVCCCV[C] *tumbrub* ‘short’
 CVCCV *dumbu* ‘any kind of grass or weed’, *dursu* ‘aim (a bow)’,
 ganta ‘four-cornered basket’, *gun̄gu* ‘great grand parent/
 child’, *ca[ɖ]ɽa* ‘bald’
 CVʔ [V] *ɖaʔa* ‘water’
 Cṽ *hē* ‘chaff’
 Cṽ: *gā:* ‘side shoots of paddy’
 Cṽʔ *hēʔ* ‘a swallow’

There are significant phonotactic restrictions on which of the Ho consonantal phonemes may appear in which syllabic structural position, for example, in onset or coda position. For example, in onset position, the following consonants are not permitted in Ho: **#ŋ-*, **#ɽ-*, **#[ŋ]* (which is non-phonemic anyway), **#w*, **#j*, and **#y* except in recent loans). As for coda restrictions, there is no **h#* permitted, *w#* is found only minimally [possibly loans?] *buɽaw* ‘wipe out completely’) *y#* is rare, **#ɽ* never occurs, (nor does **#ŋ#* which is non-phonemic anyway). Further, there are no final voiceless stops except in the word *map* ‘forgive[ness]’, and no *-s#* except in loanwords like *biswas* ‘belief’ and *asis* ‘bless[ing]’ or *bogsis* ‘reward’.

3.1 Nominal morphology

3.1.1 Number

Nouns are divided into two classes; that is, animate and inanimate, the distinction being covert, realized only in terms of the concord system between subject/object and verb. Animate nouns refer to human beings and animals. Besides these, the following nouns are classified into animate nouns in Ho: (1) Celestial bodies, for example, *siŋgi* ‘sun’, *ipil* ‘star’, and so on. (2) Supernatural beings, for example, *boŋga* ‘spirit’, *siŋ boŋga* ‘the supreme God’, and so on. (3) Natural phenomena, for example, *gama* ‘rain’, *aril* ‘hail’, and so on.

There are three numbers; singular, dual, and plural. Singular is unmarked while dual and plural are marked by *-kijn* and *-ko* respectively. Dual and plural marking is compulsory for animate nouns but optional for inanimate nouns. For example, animate nouns: *hon* ‘a child’, *hon-kijn* ‘two children’, and *hon-ko* ‘children’; inanimate nouns: *daru* ‘tree’, *daru(-kijn)* ‘two trees’, *daru(-ko)* ‘trees’.

- (22) (a) *esu lali ho-ko*
 very enough people-PL
 ‘quite a few people, more than enough’ [KCN]
- (b) *cimiŋ ho-ko Bhubaneswar-re mena?-ko*
 how.many person-PL Bhubaneswar-LOC COP-PL
 ‘How many people are there in Bhubaneswar?’ [CMH]

Sometimes the agreement rules are lax and plural agreement is found where dual might be expected in Mayurbhanj Ho. This may be an idiolectal feature or a speech error, as forms such as these are infrequently attested.

- (23) *am-a? barea hon-ko mena?-a=ko*
 YOU-GEN two child-PL COP-FIN=PL
 ‘You have two children.’ [CMH]

In Mayurbhanj Bhumij variation may be seen in the same speaker with respect to plural or dual number assignment with the numeral ‘two’. Note also the N Num order (see section 4.1), which may reflect local Indo-Aryan influence.

- (24) Bhumij
- (a) *kuŋi hon barie meŋa?=ko* (b) *kuŋi hon-kijn barie*
 girl child two COP=3PL girl child-DL two
 ‘I have two daughters.’ [CMS] ‘Two daughters.’ [CMS]

3.1.2 Case

Grammatical case marking, that is, that which indicates grammatical relations, for example, nominative/accusative or ergative, and so on, is not found in Ho. The unmarked form of the noun thus appears in both subject and object functions.

- (25) (a) *ale maŋdi=le jom-i-a* (b) *ale jilu jom-e-a=le*
 we food=1PL eat-3:T/A-FIN we meat eat-3:T/A-FIN=1PL.EX
 ‘We (would/used to) eat food.’ [KCN] ‘We (do) eat meat.’ [KCN]

The case roles of benefactive, instrumental, locative, comitative, goal, on the other hand, can be expressed by a case or clitic postposition following a noun or pronoun. For instance, there is locative *-re*, benefactive *-nangen* ‘for’, source *-ete* ‘from’, instrumental *-te* ‘by’ (or general oblique) comitative *-lo?* ‘with’.

As for the locative *-re*, it marks location at, in, or near a place.

- (26) (a) *cimij hor-ko Bhubaneswar-re mena?-ko*
 how.many person-PL Bhubaneswar-LOC COP-PL
 ‘how many people are there in Bhubaneswar?’ [CMH]
 (b) *owa?-re i?-a? mā mepi?y-a*
 family-LOC I-GEN mother COP:3-FIN
 ‘I have my mother in my family.’ [DH]

In rapid speech with inanimate possessors of inanimate nouns, the possessive/genitive (see below) *-rea?* may be reduced to *-re*.

- (27) *ape hatu-re numu cie?*
 you.PL village-LOC name what
 ‘What is the name of your village?’ [KCN]
 NB: Chaibasa Ho *nutum* ‘name’ with infix *-t-* is also used in
 Mayurbhanj Ho

In Bhumij it may appear on a nominal complement of *pai?i* ‘work’ to mark one’s profession.

- (28) Bhumij
bilqar-re-ŋ pai?-[t]en-e
 builder-LOC-1 work-PROG-FIN
 ‘I work as a builder.’ [CMS]

The multi-purpose oblique marker is *-te*. It may mark source, goal, instrument, and so on. With verbs it marks various kinds of dependent clauses (see below).

- (29) (a) *abu hotel-te=bu seno?-tan-a=bu*
 we hotel-ABL=1PL.INC go-PROG-FIN=1PL.INC
 ‘We are going from the hotel.’ [CMH]
 (b) *ale-do bas-te=le huju?-ye-n-e*
 we.PL.EXCL-EMPH bus-ABL=1PL.EXCL come-T/A-ITR-FIN
 ‘We came by bus.’ [KCN]
 (c) *ti-te=le jom-i-a*
 hand-ABL=1PL.EX eat-3:t/a-FIN
 ‘We (would/used to) eat with our hands.’ [KCN]
 (d) *jilike-te=m huju?-ye-n-e*
 how-ABL=2 come-T/A-ITR-FIN
 ‘How did you come (here)?’ [KCN]
 (e) *uqao ga?-te=ŋ huju?-ye-n-e*
 airplane-ABL=1 come-T/A-ITR-FIN
 ‘I came by airplane.’ [KCN]

In the instrumental function, the use of the oblique marker appears to be optional, at least in interrogative forms.

- (30) (a) *jilike=m huju?-ye-n-e*
 how=2 come-T/A-ITR-FIN
 ‘How did you come (here)?’
 [KCN]
- (b) *jilike-te=m huju?-ye-n-e*
 how-ABL=2 come-T/A-ITR-FIN
 ‘How did you come (here)?’
 [KCN]

In addition to these case roles, the possessive/genitive can be made by adding the nominal suffixes *-a?*, *-rea?l-ra?*, and *-ren* to a noun or pronoun. The possessive suffix *-a?* demonstrates possession by an animate noun while *-rea?l-ra?*, and *-ren* denote possession by an inanimate noun of an inanimate and animate one, respectively. Thus the elements encode animacy of both possessor and possessum. For instance, *hon-a? hatu* ‘the village of the child’, *owa?-rea?l-ra? ba?* ‘the flower of the house’; *owa?-ren hon* ‘a child of the house’.

- (31) (a) *am-a? nutum do cikana*
 YOU-GEN name EMPH what:COP:FIN
 ‘Your name, what is it.’ [KCN]
- (b) *ay-a? desum mglad*
 3:GEN country England
 ‘His country is England, he is English, from England?’ [KCN]
- (c) *side munu-ri-a? kaji*
 past times-LOC-GEN story
 ‘The story of past times.’ [KCN]
- (d) *ij? jaypur-rea? tayi-n-ten-ij?*
 I Jaypur-GEN.INAN stay-RFLXV-PROG-FIN:1
 ‘I’m Jaypurian.’ [DH]

3.1.3 Person

A bound form or short form may be used with first and second person’s possession of kinship terms; for example, *apu-m* ‘your father’ but *apu-te* ‘his/her father’.

3.1.4 Definiteness

Definiteness as a category of syntactically nominal elements is not formally defined *per se* in Ho. Within the verbal system, object agreement is generally restricted (but not for all speakers) to definite referents.

- (32) *kadal daru mara? mara? sakam-ko mena?-a*
 banana tree big big leaf-PL COP-FIN
 ‘Banana tree; (it) has really big leaves.’ [KCN]

3.1.5 Class/gender

As noted above, the possessive can be made by adding the nominal suffixes *-a?*, *-rea?l-ra?*, and *-ren* to a noun or pronoun, with *-a?* denoting possession of an animate noun, *-rea?l-ra?* and *-ren* denoting possession of or by an inanimate noun. For instance, *hon-a? hatu* ‘the village of the child’, *owa?-rea?l-ra? baa* ‘the flower of the house’; *owa?-ren hon* ‘a child of the house’. The demonstrative pronoun, which

is derived from a demonstrative base, is divided into two classes, that is, animate and inanimate and three forms of numbers, that is, singular, dual, and plural.

(33)		Animate			Inanimate		
		Proximate	Inter- mediate	Remote	Proximate	Inter- mediate	Remote
	Singular	<i>nii?</i>	<i>ini?</i>	<i>hani?</i>	<i>neyal/na</i>	<i>ena</i>	<i>hana</i>
	Dual	<i>ni-kij</i>	<i>in-kij</i>	<i>han-kij</i>	<i>neyal/na-kij</i>	<i>ena-kij</i>	<i>hana-kij</i>
	Plural	<i>ne-ko</i>	<i>en-ko</i>	<i>han-ko</i>	<i>neyal/na-ko</i>	<i>ena-ko</i>	<i>hana-ko</i>

3.1.6 Pronouns

The personal pronouns of Ho of Chaibasa are as follows:

(34)	Personal Pronouns	Full/free form			Short/bound form		
		Singular	Dual	Plural	Singular	Dual	Plural
	1 (inclusive)	<i>aŋ</i>	<i>alaŋ</i>	<i>abu</i>	<i>-ŋl-eŋ l-iŋ</i>	<i>-laŋ</i>	<i>-bu</i>
	(exclusive)		<i>aliŋ</i>	<i>ale</i>		<i>-liŋ</i>	<i>-le</i>
	2	<i>am</i>	<i>aben</i>	<i>ape</i>	<i>-ml-mel-em</i>	<i>-ben</i>	<i>-pe</i>
	3	<i>ae?</i>	<i>akij</i>	<i>ako</i>	<i>-e?l-i</i>	<i>-kij</i>	<i>-ko</i>

The personal pronoun has two forms, that is, a free form and a bound form. The free or full form appears as subject or object of sentence and head of a postpositional phrase. Full forms are used as NPs filling argument functions and may be omitted. The short forms are used for possession with a restricted set of inalienably possessed noun forms (e.g. *apu-m* 'your father' but *apu-te* 'his/her father') and as subject or object markers in the verbs, the difference being determined by their structural position within the verbal template: before the finite marker the short form pronominal encodes various kinds of object functions, while following the finite marker (or preferentially enclitic to the word immediately preceding the verb), it functions as a subject argument agreement marker (see sections 3.2.1 and 3.2.2). Under certain yet-to-be determined discourse conditions, agreement morphology with short form pronouns may be lacking on the verb as well (for more see section 3.2.1).

Possession is marked either syntactically via combination with a preposed possessive form of the pronoun or by the *t*-series of inflection added to the noun (or in a Santali-esque manner in verbs as well, see section 3.2.2 for more). Examples of the *t*-series of possessives include *hon-taben* 'your (dual) child' or *disum-tabu* 'our (inclusive) land'. The 'postpositional' or morphological form is to be considered archaic. However the preposed syntactic variant is dominant in recent usage. Thus *aben-a?* 'hon' your (dual) child' and *abu-a?* 'disum' our (inclusive) land' are frequently heard nowadays.

(35)		Preposed form			Postposed form		
		Singular	Dual	Plural	Singular	Dual	Plural
	1 (inclusive)	<i>aŋ-a?</i>	<i>alaŋ-a?</i>	<i>abu-a?</i>	<i>-taŋ</i>	<i>-talaŋ</i>	<i>-tabu</i>
	(exclusive)		<i>aliŋ-a?</i>	<i>ale-a?</i>		<i>-talij</i>	<i>-tale</i>
	2	<i>am-a?</i>	<i>aben-a?</i>	<i>ape-a?</i>	<i>-tam</i>	<i>-taben</i>	<i>-tape</i>
	3	<i>ay-a?</i>	<i>akij-a?</i>	<i>ako-a?</i>	<i>-te</i>	<i>-takij</i>	<i>-tako</i>

The following set of first and second person pronouns is from Mayurbhanj Ho and shows little difference from Chaibasa/Jharkhand Ho.

- (36) *ajn aliŋ alaŋ ape abu ale aben*
 I we.2 we.2.EX you.PL we.PL we.PL.EX you.2 [CMH]

Note that *ajn* ~ *ijn* all may be used for the first person pronoun among self-identifying Ho speakers in multilingual and multiethnic Kherwarian communities in Mayurbhanj where ‘Santali’, ‘Mundari’, ‘Ho’, ‘Bhumij’, and ‘Munda’ might be spoken.

- (37) (a) *ijn jayapura-a? tayi-n-ten-ijn*
 I Jaypur-GEN stay-RFLXV-PROG-FIN:1
 ‘I live in Jaipur.’ [DH]
 (b) *ajn-a? nutum kej si naik*
 I-GEN name K C Naik
 ‘My name is K C Naik.’ [KCN] 3:00
 (c) *ijn-a? nutum Damodar Hembram*
 I-GEN name D. H.
 ‘My name is Damodar Hembram.’ [DH]

In terms of degree of respect, there are three forms for second person singular in New Indo-Aryan languages; for example, *tū*, *tum*, are and *āp* in Hindi. These kinds of various forms of personal pronoun are rare in Munda. In Ho, however, dual forms are used for respect; for example, *aliŋ* first person exclusive dual form for first person singular (respect to addressee); *aben* second person dual form for second person singular (respect); *akijn* third person dual form for third person singular (respect to the third person).

- (38) Bhubaneswar Ho
aliŋ baro ca ju-te-liŋ sanaŋ-tan-a
 we.DL two tea drink-ABL-1DL desire-PROG-FIN
 ‘I/We two wish to drink tea.’ [CMH]

3.1.7 Demonstratives

The demonstrative system has a set of three terms for location (proximate, intermediate, and remote); that is, *nelnen* ‘this’ (dialectal variants), *en* ‘that’ and *han* ‘yonder’. These demonstrative bases are used as adjective; that is, *ne bir* ‘this forest’, *en baa* ‘that flower’, and *han owa?* ‘yonder house’. The demonstrative pronoun, which is derived from demonstrative base, is divided into two classes, that is, animate and inanimate, and three forms of numbers.

- | | | | | | | |
|----------|----------------|-------------------|-----------------|-----------------------|-------------------|------------------|
| (39) | Animate | | | Inanimate | | |
| | Proximate | Inter-
mediate | Remote | Proximate | Inter-
mediate | Remote |
| Singular | <i>nii?</i> | <i>ini?</i> | <i>hani?</i> | <i>neyalnena</i> | <i>ena</i> | <i>hana</i> |
| Dual | <i>ni-kijn</i> | <i>in-kijn</i> | <i>han-kijn</i> | <i>neyalnena-kijn</i> | <i>ena-kijn</i> | <i>hana-kijn</i> |
| Plural | <i>ne-ko</i> | <i>en-ko</i> | <i>han-ko</i> | <i>neyalnena-ko</i> | <i>ena-ko</i> | <i>hana-ko</i> |

The demonstrative adverbials are made by the combination of demonstrative bases and postpositions. For instance, in the case of *nelnen* ‘this’ there are 11 derivations:

- (40) (a) *-ta?* ‘exact place’ (b) *-paa* ‘vicinity’
nen- ta? ‘this exact place’ *nen-paa* ‘near here’

- (c) *-re* ‘in’
ne(n)-re ‘here’ *nen-taʔ-re* ‘in this exact place’ *nen-paa-re* ‘near here’
- (d) *-te* ‘to’
ne(n)-te ‘hither’ *nen-taʔ-te* ‘to this exact place’ *nen-paa-te* ‘to near here’
- (e) *-ete* ‘from’
nen-ete ‘from here’ *nen-taʔ-ete* ‘from this exact place’
nen-paa-ete ‘from near here’
- Note: *ne(n)* can be replaced by *en* ‘that’ and *han* ‘yonder’.

In addition to demonstrative adverbials, definite demonstrative derivations can be found. Definite demonstrative adjectivals are *nimin* ‘this much’ and *imin* ‘that much’, definite demonstrative nominals are *niminan* ‘only this much’ and *iminan* ‘only that much’.

3.1.8 Numerals

Numerals in Ho occur in two forms, a long form and short form. For example,

(41)	1	<i>miyaq</i>	<i>miq</i> , <i>mi</i>	6	<i>turuiya</i>	<i>turui</i>
	2	<i>bariya</i>	<i>bar</i>	7	<i>aiya</i>	<i>ai</i>
	3	<i>apiya</i>	<i>apee</i>	8	<i>irleya</i>	<i>iril</i>
	4	<i>upuniya</i>	<i>upun</i>	9	<i>areya</i>	<i>aree</i>
	5	<i>mooya</i>	<i>moe</i>	10	<i>geleya</i>	<i>gel</i>

The short forms of numerals are used in counting money, for any units of measuring while the long forms of numerals are used in counting other things. For instance, *apiya kanci-ko* ‘three baskets’ but *apee canḍuʔ* ‘three months’; *mooya oe-ko* ‘five birds’ but *moe ṭaka* ‘five rupees’; *apiya daru-ko* ‘three trees’ but *apee maa* ‘three days’, and so on. The numeral classifiers *ho[:]* for counting humans, *owaʔ* for counting houses and *bo:ʔ* for counting cattle are used in the following; for example, *gel ho hon-ko* ‘ten children’. This classifier requirement has weakened in Mayurbhanj Ho and Bhumij.

The following set was collected from speakers of Mayurbhanj Ho and Bhumij. Note the spontaneous use (and incidentally following the Ho speaker) of Indo-Aryan stems for numbers ‘7’ and higher. This reflects the endangered and tenuous status of Bhumij with respect to Ho, which is a vigorous and flourishing language.

(42)	Bhumij	Ho	gloss
	<i>mujəʔ</i>	<i>moʔy miyeʔ[d]</i>	‘1’
	<i>barieʔ</i>	<i>báriě</i>	‘2’
	<i>apieʔ</i>	<i>apié</i>	‘3’
	<i>upunieʔ</i>	<i>upunieʔ</i>	‘4’
	<i>monie/mõrĩã</i>	<i>mõne</i>	‘5’
	<i>turieʔ</i>	<i>turie[ʔ]</i>	‘6’
	<i>satta</i>	<i>ayeʔ</i>	‘7’
	<i>atta</i>	<i>eriliaʔ</i>	‘8’
	<i>nõ</i>	<i>areyəʔ...areliaʔ</i>	‘9’
	<i>dosta</i>	<i>gele ... geli</i>	‘10’
		<i>gelimiyəḍ</i>	‘11’
		<i>hisi</i>	‘20’
		<i>hisibarie</i>	‘22’
		<i>hisigel</i>	‘30’

<i>barhisi</i>	‘40’
<i>barhisigel</i>	‘50’
<i>apehisi</i>	‘60’
<i>apehisigel</i>	‘70’
<i>upunhisi</i>	‘80’
<i>upunhisigel</i>	‘90’
<i>mōyhisi</i> [CMS]	<i>mōn[e]hisi</i> [KCN]
	‘100’

As is clear from the higher numerals, Ho like many Munda languages makes use of a vigesimal counting system. Note the difference in placement of the numeral and associated noun in the following Ho of Bhubaneswar and rural Mayurbhanj Bhumij forms. Bhumij is usually N Num and this urbanized Ho is Num N like the interrogative expression (*cimiṅ hon-ko*) and variable in varieties used by educated Mayurbhanj Ho speakers and in response to a preposed interrogative quantifier in Bhumij. Nouns following numerals that are animate take plural (or dual) marking, but inanimate ones remain unmarked.

(43) Bhubaneswar Ho

- | | |
|--------------------------------------|---------------------------------------|
| (a) <i>am-a?</i> <i>barea hon-ko</i> | (b) <i>am-a?</i> <i>upunia hon-ko</i> |
| you-GEN two child-PL | you-GEN four child-PL |
| <i>mena?-a=ko</i> | <i>ban-ko-a</i> |
| COP-FIN=3PL | NEG.COP-PL-FIN |
| ‘You have two children.’ [CMH] | ‘You don’t have four kids.’ [CMH] |

(44) Mayurbhanj Ho

- | |
|---|
| (a) <i>cimeniṅ hon-ko mena?-ko-a</i> |
| how.many child-PL COP-PL-FIN |
| ‘How many children do you have?’ [KCN] |
| (b) <i>barhisi sirmi=ṅ hobo-ya-n-a</i> |
| forty year=1 happen-T/A-ITR-FIN |
| ‘I am 40’ [KCN] |
| (c) <i>hisigeliril sirme=ṅ hobo-ja-n-a</i> |
| 38 year=1 happen-T/A-ITR-FIN |
| ‘I am 38 years old.’ [KCN] |
| (d) <i>am-a?</i> <i>cimiṅ sirme hobo-ja-n-a</i> |
| you-GEN how.many year happen-T/A-ITR-FIN |
| ‘How old are you?’ [KCN] |

(45) Mayurbhanj Bhumij

- | | |
|-------------------------------|---------------------------------|
| (a) <i>barhisituri sirme?</i> | (b) <i>mōrĩā hon-ko</i> |
| forty-six year | five child-PL |
| ‘[I am] 46 years old.’ [CMS] | ‘[I have] five children.’ [CMS] |

(46) Mayurbhanj Ho

- | |
|---------------------------------------|
| <i>kuṛi hon-ko cimiṅ mena?-ko-a</i> |
| daughter child-PL how.many COP-PL-FIN |
| ‘How many daughters?’ [KCN] |

(47) Mayurbhanj Bhumij

- | | |
|------------------------------------|-------------------------------|
| (a) <i>kuṛi hon barie meṅa?=ko</i> | (b) <i>kuṛi hon-kij barie</i> |
| girl child two COP=3PL | girl child-DL two |
| ‘I have two daughters.’ [CMS] | ‘Two daughters.’ [CMS] |

Distributive numerals are very common in Kherwarian. The following distributive numerals are found in Ho, with both long and short forms:

- (48)
- | | | |
|----|--------------------|------------------|
| 1 | <i>miipiya</i> | <i>mimi</i> |
| 2 | <i>baa-bariya</i> | <i>baa-bar</i> |
| 3 | <i>aa-apiya</i> | <i>aa-apee</i> |
| 4 | <i>uu-upuniya</i> | <i>uu-upun</i> |
| 5 | <i>moo-mooya</i> | <i>moo-moe</i> |
| 6 | <i>tuu-turuiya</i> | <i>tuu-turui</i> |
| 7 | <i>aa-aiya</i> | <i>aa-ai</i> |
| 8 | <i>ii-irleya</i> | <i>ii-iril</i> |
| 9 | <i>aa-areya</i> | <i>aa-aree</i> |
| 10 | <i>gee-geleya</i> | <i>gee-gel</i> |

The distributive numeral forms are basically made by repeating the first syllable and doubling the first vowel of cardinal numeral forms except 'one' which is formed by inserting the infix *-pV-*. These constructions are slightly different from Mundari, where *-pV-* infixation predominates, for example, *apapia* 'three each', *upunia* 'four each' in Mundari.

3.1.9 Adpositions

The distinction between adpositions (postpositions) and case markers is not clear synchronically, and the connection diachronically in Ho is obvious. Postpositions follow their nominal complements in Ho and show varying degrees of phonological dependency on the noun, for example, *mente* is still free-standing while *=/leka* is well on its way to affixal 'case' status. Combined with the demonstrative stem, *leka* is a common clause-initial adverbial used for stringing clauses together in certain discourse registers and conversational genres.

- (49) (a) *wā mente ma mente ka?-e? kaji-e wā mente kaji-e[?]*
wā like ma like NEG-3 speak-FUT wā like say-FUT-?
 'like *wā* s/he speaks not *ma*, s/he speaks like *wā*' [KCN]
- (b) *enleka ne-ya=ko men-iy-a a? cilike kaji-re*
like.that this-EMPH=3PL say-FUT-FIN a? how say-LOC
 'it's like that, how they will say it, like *a?*' [KCN]
- (c) *enleka uy?-ke-redo yuw-e*
like.that jump.up-T/A-COND fall.down-FUT:FIN
 '(and) if (s/he) jumps up like that, (then) she will fall down' [KCN]

Like many Eurasian SOV languages, Ho uses case-marked relational nouns, like *biter/bitar* or *danaŋ* in the following examples:

- (50) (a) *akij-e? mon biter-re jetan pap uŋu?-ko ka*
they.two-GEN mind inside⁴-LOC any sin thought-PL NEG
taiken-a
 PST.COP-FIN
 'They two had no sinful thoughts in their minds.' [KCN]
- (b) *akij-e? mon biter-re eta?lekan rāsa uŋu?-ko*
they.2-GEN mind inside-LOC other.type joy thought-PL

huju?-*ye-n-a*

come-T/A-ITR-FIN

'In their minds they felt another type of joyful thoughts.' [KCN]

(c) *barsij danar-re diyen upude-ya-n-a*

two.day later-LOC liquor liquor.become.ready-T/A-ITR-FIN

'After two days the liquor was ready.' [KCN]

3.1.10 Derivation

There is no productive nominal derivation in Ho and the minor Kherwarian languages except for verbal nouns often, but not necessarily with instrumental semantics, marked by infixation of *-nV-*, with V representing a short copy of whatever the vowel of the initial syllable of the stem is, whether it is monosyllabic (as is typical) or bisyllabic. That is, the shape of infix is seen by the lack of vowel length, nasalization, or glottalization of the vowel before the *-n-* which is always otherwise identical to the one following it.

- (51) *sumutu?* 'follow after' < *sutu?* 'next one in birth'
tanagoe? 'jaw' < *tagoe?* 'chew'
tonol 'bond, strap' < *tol* 'tie'
enem 'gift' < *em* 'give'
anasi 'petition' < *asi* 'ask for'
anacu 'a command to do some work' < *acu* 'command to do some work'
ana:du < *adu* 'descent'
danalob 'covering' < *dalob* 'cover'
dunub 'meeting' < *dub* 'sit'
gonoe? 'death' < *goe?* 'kill'
giniyu? 'private body parts' < *giyu?* '(be) [a]shame[d], bashful, shy'
gononj 'cost, brideprice' < *gonj* 'give girl in marriage' cf. *gononjan* 'valuable'
hanatijn 'share' < *hatijn* 'divide'
jiniq 'life, living' < *jiq* 'to live'
hunuju? < *huju?* 'a coming, an advent'
jono? 'broom' < *jo?* 'sweep'
kanaji 'something told related' < *kaji* 'tell'
onol 'writing, inscription' < *ol* 'write'
manaraj 'greatness, glory' < *maraj*
ranapiq 'eyelid' < *rapid* 'wink, blink'
pana: 'fork in road, part in hair' < ? *pa:?* 'chop (wood)'
ranakab 'steep ascent' < *rakab* 'climb'
sanari 'truth' < *sari* 'true' (note in passive means 'believe')

Some forms are variably formed with an *-n-* infix in Ho.

- (52) *racam* ~ *ranacam* 'scissors'
uka ~ *unuka* 'elbow'

There is also evidence of nouns derived via a *-p-* infix, found in both its verbal reciprocal meaning and in certain kin terms.

- (53) *sopola* 'reconciliation'
sepe:q 'young man' (cf. *honse:q* 'older sister's nephew')
hapanum 'young woman'

epeser ‘counter claims of possession’
gopoe? ‘a fight, battle’
kupusur ‘mutual enmity’

One element that appears in a number of Ho lexemes with identifiable or probable roots is *-ob*.

- (54) *kaŋkom* ~ *kaŋob* ‘crab’
apir ‘fly’ *aparob* ‘wing’

In addition, another common sequence found in a range of nouns is *-kom*; there are a number of phonologically similar elements that occur less frequently, for example, *-com* (~*-lom*) and *-qom*.

- (55) *kama:kom* ‘the stem of a leaf’ *damkom* ‘young bull’
kaŋkom ‘crab’ *klparkom* ‘bed’
baqcom ‘sabai grass’ *caqcom* ~ *caqlom* ‘tail’
tonqom ‘knot’ *kandom* ‘rim of vessel or basket’

There is a semi-productive means of deriving adjectival forms via either full reduplication (CVCVC-CVCVC) or stem reduplication (CV[C]-CVC), often (though far from obligatorily) verbal in nature. Note that C₂ can be *r* in the reduplication pattern and contain a homo-organic nasal/stop cluster in the duplication pattern. In addition to reduplication and duplication, an ‘expressive’ (derivational) use of lengthening of a verb stem vowel can show the same function: *jilad* ‘slip’ > *ji:lad* ‘slippery’.

- (56) *taŋga* ‘separate (vb)’ > *taŋga-taŋga* ‘separate (adj)’
ji:ʔl ‘slip’ > *ji:ji:ʔl* ‘slippery’
juwa ‘to stick’ > *jujuwa* ‘sticky’
je:ʔr ‘to weld’ > *je:ʔr-je:ʔr* ‘to be jammed’
jerem ‘wet or damp and sticky’ > *jerem-jerem* ‘wet or damp and sticky’
ipil ‘star’ > *ipiʔ-ipiʔ* ‘twinkle (of stars)’
- (57) *tepe?-tepe?* ‘shallow’ *kurkur* ‘angry’⁵
lolo ‘hot’ *sasaŋ* ‘yellow’
mamaran ‘proud’ *sasa:* ‘cold’

Reduplication may also be found in words such as *sasan* ‘burial place’. Syntactic compounding is also found in certain noun phrase-cum-lexical expressions.

- (58) (a) *kani aŋguʎi* (b) *bo?* *bale* [*kan-a*]
 ?pinkie finger head hair [COP-FIN]
 ‘pinkie finger’ ‘(it is) head hair’ [KCN]
 (c) *muteʔ-re sunui[ʔ]*
 nose-LOC ‘snot’
 mucous (cf. *sului?* Deeney 1975; Mundari: *sului?*, *sunui?*, *suli?*, *sumi?*) [KCN]
 (d) *sim jilu* vs. *merom jilu*
 ‘chicken meat’ ‘goat meat’ [KCN]

3.1.11 Adjectives

Adjectives precede the nouns they modify in Ho. On the distinction of parts of speech in Mundari, including adjectives, most arguments of which appear valid for

Ho and the other minor ‘Mundari’-like Kherwarian speech varieties, see Evans and Osada (2005) for a discussion about Mundari. Adjectives may be modified by degree adverbs like *esu* ‘very’.

- (59) (a) *ciliken lipi-ko mena?-a*
 what.kind letter-PL COP-FIN
 ‘What kinds of letters there are.’ [KCN]
- (b) *tisiŋ aliŋ kaji tey-a?-a-do nen citi-do-ko*
 today we say AUX-ITR/PASS-FIN-EMPH this writing-EMPH-PL
meta?-a waraŋ citi lipi
 say-ITR/PASS-FIN waraŋ citi letter
 ‘Today we will speak about these ways of writing, that are called the Warang Chiti letters.’ [KCN]
- (c) *ne-ya-do waraŋ citi ne-ya-do huriŋ*
 this-FOC-EMPH big case this-FOC-EMPH small
 ‘This one is the upper case and this one is the lower.’ [KCN]
- (d) *esu pure?e sitme parom-e-ya-n-a*
 very many year pass-EPEN-T/A-ITR-FIN
 ‘very many years have passed since then.’ [KCN]
- (e) *esu pure? kili bai-ye-n-a*
 very many clan form-T/A-ITR-FIN
 ‘Very many clans have been formed.’ [KCN]
- (f) *nimiŋ cera ʃarna-ko taiken-a*
 very beautiful spring-PL PST.COP-FIN
 ‘There were very beautiful springs.’ [KCN]
- (g) *akiŋ-e? mon biter-re eta?lekan rāsa uŋu?-ko*
 they.2-GEN mind inside-LOC other.type joy thought-PL
huju?-ye-n-a
 come-T/A-ITR-FIN
 ‘In their minds they felt another type of joyful thoughts.’ [KCN]

Doubling of adjective stems is also seen in cases of emphasis.

- (60) *kadal maraŋ maraŋ sakam-ko mena?-a*
 banana big big leaf-PL COP-FIN
 ‘Banana; it has really big leaves.’ [KCN]

Augmentatives of certain stems may be formed by copying a lengthened form of the stem vowel and infixation of *-p-* after the first consonant.

- (61) *ma:paraŋ* ‘very big’ < *maraŋ*

3.1.12 Adverb(ial)s

Ho realizes a range of different adverbial forms. Formally speaking, adverbials are a heterogeneous class of elements, and include case forms of nouns, opaque derived, compound or root forms. Some elements are preferentially clause-initial, but adverbials generally appear in a position preceding the verb or adjectival form they modify.

- (62) (a) *ente akij=do=kij juʔi-ye-n-a*
 then they.2=EMPH=3DL copulate-T/A-ITR-FIN
 ‘Then they started copulating.’ [KCN]
- (b) *ena-ge saben etoi-e-ka bugi-leka-te ondo?o*
 that-FOC all learn-3-OPT do.well-ADV=ABL and
desum-re nitir-e=pe
 country-LOC spread-FUT:FIN=2PL
 ‘(So that) it will all have been learned well and you will spread it out in the country.’ [KCN]
- (c) *tisiŋ aliŋ kaji tey-aʔ-a-do nen citi-do-ko*
 today we say AUX-ITR/PASS-FIN-EMPH this writing-EMPH-PL
meta-ʔ-a waraŋ citi lipi
 say-ITR/PASS-FIN waraŋ citi letter
 ‘Today we will speak about these ways of writing, that are called the Warang Chiti letters.’ [KCN]
- (d) *nimiŋ cera ʃ^harna-ko taiken-a*
 very beautiful spring-PL PST.COP-FIN
 ‘There were very beautiful springs.’ [KCN]
- (e) *barsiŋ danaŋ-re diyey upud-e-ya-n-a*
 two.day later-LOC liquor ferment-EPEN-T/A-ITR-FIN
 ‘After two days the liquor was ready.’ [KCN]

Both locative (-*re*) and oblique/ablative (-*te*) forms of nouns may function as adverbial forms in Ho, sometimes from the same stem. Usually ‘auxiliary’ or relational nouns appear in locative case forms and adjectival-type stems in the oblique/ablative.

- (63) *samanayre ~ sanamayre* ‘in front of’
sanaŋte ‘intentionally’
sari sarite ‘really’
rāsate ‘willingly’
bar ‘outside’
bar-re ‘outside’
bar-te ‘(from/to) outside’

Adverbs often appear to be (partially) reduplicated words, though the base is not always attested as an independent lexeme.

- (64) *mir mir* ‘just before daylight’
pari pari ‘in turn, by turns’
amna samna ‘face to face’
gene gene ~ gete gete ‘in a line’

3.2 Verbal morphology

The verbal morphology in Ho is very complex. It involves the following grammatical categories: TAM(=Tense, Aspect, Mood), Voice, Transitivity, Finiteness, as well as subject, object, and possessor markers. The verbal base is formed by the verbal root and (additionally, if motivated) a derivational affix. Serial verb and auxiliary verb

constructions frequently occur. A verbal intensifier is also found. The Ho verbal template (Box 5.1) is as follows:

BOX 5.1: HO VERB TEMPLATE

VB + (Asp) + (Trans/Intrans) + (Obj) + (IND/FIN) + (Subj)
 VB = VR + (VR) + (INTEN) + (Affix)

Tense is not morphologically marked and is divided into future (usually unmarked) and non-future. There are three moods; that is, indicative (unmarked), imperative, and optative.

3.2.1 *Subject*

As in Mundari and Santali, the expression of subject takes one of two basic patterns in the Ho verb. The most common pattern is for the short form of the pronominal to appear enclitic to the word immediately preceding the (finite) verb, regardless of what kind of word that preverbal word is. It can be a negative particle, a verbal complement in a complex predicate or complement structure, an interrogative pronoun, case-marked oblique or complement phrases, argument noun phrases, or even subject pronoun elements themselves (as in example (65g)).

- (65) (a) *ka=m samao-tan-a*
 NEG=2 understand:ITR-PROG-FIN
 ‘You don’t understand.’ [KCN]
- (b) *aliŋ baro ca ju-te=liŋ sanay-tan-a*
 we.DL two tea drink-ABL=1DL desire-PROG-FIN
 ‘We two wish to drink tea.’ [CMH]
- (c) *okonde=m jonom-le-n-a*
 where=2 born-T/A-ITR-FIN
 ‘Where were you born?’ [KCN]
- (d) *ijŋ jaypura hatu-r=ijŋ jonom-le-n-a*
 I Jaypur village-LOC=1 born-T/A-ITR-FIN
 ‘I was born in Jaypur village.’ [DH]
- (e) *aiŋ ho:ko=jŋ nel-ko-a*
 I person-PL=1 see-3PL-FIN
 ‘I will see them.’ [KCN]
- (f) *ale-do bas-te=le huju-ye-n-e*
 we.PL.EX-EMPH bus-ABL=1PL.EXCL come-T/A-ITR-FIN
 ‘We came by bus.’ [KCN]
- (g) *ne?ene-ya abu-do=bu mene-tan-a cimitay*
 this.one-FOC we-EMPH=1PL.INC say-PROG-FIN when
abu=bu boŋga-e-tan-a
 we=1PL.INC worship-3-PROG-FIN
 ‘This one here we say it when we are worshipping.’ [KCN]

The other pattern is to mark subject at the very end of the verbal complex, after the finitizer *-a*.

- (66) *maŋd̪i jom-ke-q-a=m am andi-ka-n-a=m*
 rice/food eat-T/A-TR-FIN=2 you marry-T/A-ITR-FIN=2
 'You ate rice/food.' [CMH] 'Are you married?' [KCN]

Under yet to be determined discourse or conversational conditions, in particular, but certainly not obligatorily, in the presence of an overt subject pronoun, subject marking in the verb can be omitted in Mayurbhanj Ho.

- (67) *aj̪n nel-ki-e? aj̪n ho: nel-ki-e?*
 I see-T/A-3:FIN I person see-T/A-3:FIN
 'I saw him.' [KCN] 'I saw the person.' [KCN]

In conjoined phrases, one subject clitic may suffice for both conjuncts in a same subject structure.

- (68) *daʔa ka=laŋ ju-e ca ju-e*
 water NEG=1DL drink-IMP tea drink-IMP
 'Let's not drink water, let's drink tea.' [KCN]

Given that plural number marking can be optional, it is sometimes not clear whether a given form has subject marking or not, as in the following sentence. It is not clear whether the *-ko* is a nominal plural marker, a third animate plural subject marker, or a haplologized form of both.

- (69) *ho-l=ko nel-ij̪n-e*
 person-/[=3]PL see-1-FIN
 'The people will see me.' [KCN]

The preverbal and postverbal subject marking strategies alternate freely, with no apparent difference in meaning as in the following three nearly sequential spontaneous utterances. What exactly triggers the use of one variant over the other in given contexts remains a subject for future research. Note that there is a difference between *jom-i-a* and *jom-e-a* in some Kherwarian dialects, meaning 'eat Animate being' and 'eat Inanimate thing'. It is possible that this particular difference in the forms below signals a differential gender class between food and meat in these examples. It could just be simple phonetic variation, however.

- (70) (a) *ale maŋd̪i=le jom-i-a* (b) *ti-te=le jom-i-a*
 we food=1PL eat-3:T/A-FIN hand-ABL=1PL.EX eat-3:T/A-FIN
 'We (would/used to) eat food.' 'We (would) eat with our hands.'
 [KCN] [KCN]
 (c) *ale jilu jom-e-a-le*
 we meat eat-3:T/A-FIN=1PL.EX
 'We eat meat.' [KCN]

Other minor Kherwarian languages show similar patterning, with a statistical dominance of the preverbal pattern, but both well attested.

- (71) Mayurbhanj Bhumij
 (a) *am cie-ko=m paiʃi-ten-e* (b) *cie=ʃn paiʃi-ten-e*
 you what-PL-2 work-PROG-FIN what work-PROG-FIN
 'What jobs do you do?' [CMS] 'What do I do?' [CMS]

- (72) Asuri
duɾup-geɾ-tan-a-iŋ
 sit-CAUS-PROG-FIN-1
 ‘I cause to sit.’ (Grierson 1906:139)

Double marking of subject is also possible in Ho and the minor Kherwarian languages, primarily in contrastive or expressive discourse. For example, in Koḍa, double marking of subject (with both the preverbal and post-finitizer positions of subject marking filled) is found in negative forms (under certain discourse conditions), as well as in Mayurbhanj and Bhubaneswar Ho varieties. Note that this is also characteristic of Kera? Mundari (Kobayashi and Murmu, this volume).

- (73) (a) Koḍa
ka:m äm-ta-t-iñ-a:m
 NEG-2 give-ASP-TR-1-FIN-2
 ‘You didn’t give me (it).’
 (Grierson 1906:112)
- (b) Koḍa
ka:n taɾam-paɾã:m-ta-t-iñ
 NEG-1 sin-ECHO-ASP-TR-1
 ‘I didn’t sin.’
 (Grierson 1906:112)

- (74) Bhubaneswar Ho
abu hotel-te=bu seno?-tan-a=bu
 we hotel-ABL=1PL.INC go-PROG-FIN=1PL.INC
 ‘We are going to/from the hotel.’ [CMH]

3.2.2 Object types

A range of different object types may be encoded in the Ho verb. This includes patient and recipient arguments as well as possessor of object. The last two are preceded by the benefactive and possessive affixes, respectively. Usually only animate beings trigger object agreement, but some culturally significant objects (optionally) trigger this argument encoding pattern as well. The object position in the verb template, encoded by short form pronominal elements, is following the transitivity marker and before the finitizer. Note that nouns that fill patient or recipient argument roles themselves appear unmarked even if encoded within the object agreement morphology of the verb, although recipients may optionally take *-te* as well.

- (75) (a) *merom jilu bugine bugi=le jom-i-a*
 goat meat well well=1PL eat-T/A:3-FIN
 ‘We like to eat goat meat.’ [KCN]
- (b) *aij hor-ko=p nel-ko-a*
 I person-PL=1 see-3PL-FIN
 ‘I will see them.’ [KCN]

Subcategorized ‘object’ arguments of verbs may, depending on the inflectional paradigm/category and a variety of still poorly understood discourse factors, appear as object morphology in the verb, as a transitivity marker, or both.

- (76) (a) *ale diyen=le ju-i-e*
 we.PL country.liquor=1PL drink-T/A:3-FIN
 ‘We (like to) drink country liquor.’ [KCN]
- (b) *diyen diyen nele-ke-q-a*
 country.liquor country.liquor see-T/A-TR-FIN
 ‘Country liquor. Have (you) tried country liquor?’ [KCN]

Minor Kherwarian languages form two subgroups with respect to the formal encoding in the verb of non-patient arguments, in particular recipients. One group, which includes Ho, treats recipients as morphosyntactically distinct from patient-type object arguments, and this is encoded by the use of the benefactive ‘aspect’ marker *-a-*. Thus, recipient marking in Ho consists of the object markers preceded by *-a-*.

- (77) *ca=jn em-a-m-tan-a*
 tea=1 give-BEN-2-PROG-FIN
 ‘I’m giving you tea.’ [KCN]

Many minor Kherwarian languages not surprisingly show a pattern similar to that of Ho. The benefactive marker itself may appear in a variety of phonetic realizations across these Kherwarian varieties, for example, *-a-*, *-wa-*, *-ova-*, or even *-ga-*.

- (78) (a) Singbhum Bhumij (b) Asuri
ka:yi-a-d-i-ya: *ha:ʃiŋ-ova-t-kiŋ-a*
 say-BEN-TR-3-FIN divide-BEN-TR-3DL-FIN
 ‘said to him’ ‘divided to them 2’
 (Grierson 1906:100) (Grierson 1906:138)
- (c) Palamau Brijia Asuri (d) Palamau Korwa
ha:ʃi-wa-d-i-a: *ha:ʃiŋ-wa-iŋ-me*
 divide-BEN-TR-3-FIN divide-BEN-1-2
 ‘divided to him’ ‘divide to me!’
 (Grierson 1906:143) (Grierson 1906:151)
- (e) Hazaribagh Korwa
mene-m em-ga-d-iñ-a
 not-2 given-ASP-TR-1-FIN
 ‘you haven’t given to me’
 (Grierson 1906:161) *w > *g ??
- (79) (a) Erṅa Korwa (b) Erṅa Korwa
ñaw-ā-iŋ *kiyā-wā-iŋ-a*
 seek-BEN-1 buy-BEN-1-FIN
 ‘seek for me’ ‘buys for me’
 (Grierson 1906:163) (Grierson 1906:166)

Less commonly, other minor Kherwarian languages treat recipients in common ditransitive verbs as formally similar to patient objects, in a primary object like configuration (Dryer 1986). Such languages include Birhor.

- (80) Birhor
kañi-ki:ch-a-e
 tell-ASP:TR:3-FIN-3
 ‘He told him.’ (Grierson 1906:103)

In a restricted set of instances, it is also possible in Ho to encode the possessor of a logical argument (object), similar to the system seen in Santali (Ghosh this volume, Neukom 2000). Similar forms are reported for Koḍa as well (Grierson 1906).

- (81) Ho
bo:ʔ=e? heḍ-ki-ḍ-iŋ-a *bo:ʔ=e? heḍ-ki-ḍ-iŋ-a*
 head=3 pull-T/A-TR-1-FIN head=3 pull-T/A-TR-POSS-1-FIN
 ‘She/he pulled my hair.’ ‘She/he pulled my hair.’ [Burrows 1915:40]
 confirmed by C.M. Haibru

- (82) Koḍa
tusiŋ-taxe-pe:
 put.on- POSS:3SG -2PL
 'put on his (y'all)!' (Grierson 1906:110)

On rare occasion for emphasis an object can be doubled in Ho, as in the following utterance:

- (83) *ca=jn em-a-jn-me*
 tea=1 give-BEN-1-2
 'Give me tea.' [KCN] 18:40

3.2.3–3.2.4 Tense/aspect

Kherwarian tense/aspect systems (which are generally conflated) are decidedly complex. Perhaps in a strict sense, it is possible to say that tense proper is not reckoned within the Ho and minor Kherwarian verbal complexes, and that aspectual distinctions of some sort are superordinate to any functional temporal categories, one possible exception to this is a future element in *-e-* (\sim *-i-*) found in Mayurbhanj Ho and some varieties of Bhumij. Anderson (2007) reckons a two-way set of oppositions in the Kherwarian verbal complex, labeled Series-A and Series-B, respectively (although the primary aspectual marker in Series-A is of Series-B origin, all diachronically semi-fused 'auxiliary' or complex predicate structures). Inimately connected with, but logically and formally independent of, (tense)/aspect marking in Ho verb morphology are categories of valence/transitivity, possibly derived themselves from original tense markers from Proto-Munda. Specifically, some τ/A markers in Series-B come in transitive and intransitive sets, while in other instances, one variant or another has been grammaticalized in a particular function, a system which may in fact be older, dating from Proto-North Munda. Thus in 'aorist' or 'perfect', that is, default 'past' contexts, the preferred affix combination is *ke-d* for transitives or class-I verbs and *ya-n* for intransitives (detransitive) or class-II verbs; *-ke-n* occurs now mainly in a frozen form in the 'past' copula form *taiken* and *ya-d* is an ostensible transitive form corresponding to intransitive *tan* in Series-I/progressive (Ramaswami 1992) but this does not occur in our field corpus. The set of these τ/A -cum-transitivity elements in Ho and Bhumij is offered below.

- | | | | | | | |
|------|--------|---------|-----------------------------|-------------|---------------------------|--------------------------|
| (84) | Ho | TRANS | <i>ke-d</i> | <i>le-d</i> | <i>aka-d</i> | <i>ke-d</i> |
| | | INTRANS | <i>(e)ya-n</i> ⁶ | <i>le-n</i> | <i>aka-n</i> ⁷ | <i>ke-n</i> ⁸ |
| | Bhumij | TRANS | <i>-taʔt</i> , | <i>ke-d</i> | <i>le-d</i> | <i>ke-dlki</i> |
| | | INTRANS | <i>-ya-n</i> | <i>le-n</i> | <i>aka-n</i> | <i>ke-n/ya-n</i> |
- (85) (a) *maṅḍi jom-ke-d-a=m* (b) *seno-ya-n-a=m*
 rice eat- τ/A -TR-FIN=2 go- τ/A -ITR-FIN=2
 'You ate the rice.' [CMH] 'You went.' [KCN]

A somewhat large set of these are used across the minor Kherwarian languages, at least in terms of realizations of the relatively restricted functional (and formal/templatic) set of affixes. Different languages prefer different combinations in particular lexical/functional configurations, but most show the same *-ke-d* : *-ya-n* opposition attested in Ho. Note that the τ/A + transitivity combinations minus the finitizer (but

optionally including object agreement morphology) may function in non-sentence final position as participial verb forms in different complex clause configuration subtypes.

- (86) (a) Palamau Brijia Asuri
kul-tad-i-a:
send-ASP-TR-3-FIN
'sent him'
(Grierson 1906:144)
- (b) Palamau Brijia Asuri
raje-tha:n-a:
poor-ASP-ITR-FIN
'became poor'
(Grierson 1906:144)
- (87) (a) Jashpur Korwa
bol-ta:
enter-ASP:FIN
'he entered'
(Grierson 1906:150)
- (b) Palamau Korwa
sen-te-n-a:
go-ASP-ITR-FIN
'he went'
(Grierson 1906:151)
- (88) (a) Koḍa
pap-ta-t-iñ
sin-ASP-TR-1
'I sinned'
(Grierson 1906:111)
- (b) Asuri
sen-ta-d-a:
go-ASP-TR!-FIN
'he went'
(Grierson 1906:139)
- (89) 'Raigarh Mañjhi'
torya:kul-ta-i-ya:
away-send-ASP-3-FIN
'sent him away' (Grierson 1906:146)
- (90) (a) Turi
idi-ta-n-a-ku:
take.away-ASP-ITR-FIN-PL
'They took away.'
(Grierson 1906:131)
- (b) Turi
go:t-cha:ba:ta-n-a-ku:
pluck-COMPLT-ASP-ITR-FIN-PL
'They finished plucking.'
(Grierson 1906:131)
- [NB: original serial verb construction]
- (91) (a) Koḍa
da:l-ek-et-iñ
strike-INAN-ASP-1
'I strike (it).'
(Grierson 1906:114)
- (b) Koḍa
da:l-i:ch-et-iñ
strike-3-ASP-1
'I strike him.'
(Grierson 1906:114)
- (92) (a) Asuri
sen-ya-n-a:
go-ASP-ITR-FIN
'She went.'
(Grierson 1906:139)
- (b) 'Raigarh Mañjhi'
para-ya-n-a:
fall-ASP-ITR-FIN
'He fell.'
(Grierson 1906:146)
- (93) (a) Asuri
sen-e-n-a:
go-ASP-ITR-FIN
'He went.'
(Grierson 1906:139)
- (b) Birhor
nam-e-d-e-a-e
find-ASP-TR-3-FIN-3
'He found him.'
(Grierson 1906:103)

- (94) Jashpur Korwa⁹
sen-e-q-a'
 go-ASP-TR!-FIN
 '(he) went' (Grierson 1906:151)
- (95) (a) Koḍa
da:l-e't-a-ñ
 strike-ASP-TR-FIN-1
 'I struck'
 (Grierson 1906:110)
- (b) Palamau Brijia (Asuri)
seno:-a-n-a'
 go.ITR/PASS-ASP-ITR-FIN
 'went'
 (Grierson 1906:140)

Note that in the sets that have both transitive and intransitive members in Ho, a transitive stem can be semantically marked as detransitive (passive, reflexive) by using the *-n-* for the expected *-q-l-d- quḍi tisiḅ rapud-ea-n-a* 'today the bridge was broken'; cf. *quḍi-e rapud-ke-d-a* 'he broke the bridge'. (Burrows 1915:33)

- (96) *ne-ya-do=ko* *huriḅ huriḅ-ti-a* *maray-ti-a-do*
 this-FOC-EMPH=3PL small small-ABL//T/A-FIN big-ABL//T/A-FIN-EMPH
maray-leka-ge *ol-aka-n-a*
 big-PP-FOC write-PLUP-ITR-FIN
 'these ones here, small and large, like, have been written' [KCN]

In terms of historical origin, the Series-B 'progressive' or present marker *tan* (~ harmonic variant in Mayurbhanj Ho *ten*) is **-ta-n* of the same formal shape as the Series-B *t/A+*transitivity markers. However, in Ho, it is used with intransitive and transitive verbs alike. Although historically a part of Series-B, this is no longer the case in Ho, and *-tan* is now a part of a formally different paradigmatic class, distinguished by the use of the inflectional detransitive marker *-o[?]* on verbs (since the *-n-* is found with both transitive and intransitive stems, and the functional flexibility exhibited by *-n-* seen in Series-B is not available in Series-A inflections) and the pre-aspectual position of the object agreement markers (see example (97e)).

- (97) (a) *cimiḅ* *ho:-ko* *kaji-ten-e*
 how.many Ho-PL speak-PROG-FIN
 'How many Ho speak (their language)?' [KCN]
- (b) *cimiḅ* *ho:-ko* *jagar-tan-a*
 how.many Ho-PL speak-PROG-FIN
 'How many Ho speak (their language)?' [KCN]
- (c) *aliḅ* *baro* *ca* *ju-te=liḅ* *sanay-tan-a*
 we.DL two.people tea drink-ABL=1DL desire-PROG-FIN
 'We two wish to drink tea.' [CMH]
- (d) *ka=m* *samao-tan-a*
 NEG=2 understand:ITR-PROG-FIN
 'You don't understand.' [KCN]
- (e) *ca=n* *em-a-m-tan-a*
 tea=1 give-BEN-2-PROG-FIN
 'I'm giving you tea.' [KCN]

Asuri and other 'minor' Kherwarian languages make similar use of *-tan-*

- (98) Asuri
duɽup-ge:-tan-a:-ij
 sit-CAUS-PROG-FIN-1
 'I cause to sit.'
 (Grierson 1906:139)
- (99) (a) Turi
idi-ta:-n-a:=kuw
 steal-ASP-ITR-FIN=PL
 'They stole.'
 (Grierson 1906:131)
- (b) Turi
go:t-ca:ba:-ta:-n-a:=kuw
 gather-COMPL-ASP-ITR-FIN=PL
 'They finished gathering.'
 (Grierson 1906:131)
- (100) (a) Koḍa
da:l-ek-et-ij
 strike-INAN-ASP-1
 'I strike (it).'
 (Grierson 1906:114)
- (b) Koḍa
da:l-i:ch-et-ij
 strike-3-ASP-1
 'I strike him.'
 (Grierson 1906:114)
- cf. (c) Koḍa
da:l-ed-i:ch'-tan-a-ṅ
 strike-TR-3-PROG-FIN-1
 'I'm striking him.'
 (Grierson 1906:110)

Some varieties of Jharkhand Bhumij may have preserved an older state with the *tan* intransitive progressive, and with antipassivized or detransitivized semantics with transitive stems, in opposition to the *-ja²t* (<*-*ya-d*) transitive progressive.

- (101) (a) Bhumij
Soma sen-ta-n-a-?e¹⁰
 Soma GO-ITR-FIN-3
 'Soma is going.'
 (Ramaswami 1992:99)
- (b) Bhumij
ape nu?u-ya?-t-a-pe
 you drink-ASP-TR-FIN-2PL
 'You are drinking it.'
 (Ramaswami 1992:99)
- (c) Bhumij
hɔɽ:ta-ke le:(ja²t)-ji-a-ij
 man:CLSSFR-OBJ see-ASP.(TR)-3-FIN-1
 'I am looking at the man.'¹¹
 (Ramaswami 1992:99; 93)
- (d) Bhumij
daruṭa le:(ja²t)-a-ij
 tree:CLSSFR see-ASP-FIN-1
 'I am looking at the tree.'
 (Ramaswami 1992:99)
- (e) Bhumij
jom-ta-n-a-ij
 eat-PROG-ITR-FIN-1
 'I'm eating.'
 (Ramaswami 1992:98)
- (f) Bhumij
ir-ta-n-a-ij
 reap-PROG-ITR-FIN-1
 'I am reaping.'
 (Ramaswami 1992:98)

This is further suggested by the range of elements that may function as an imperfect or past progressive, utilizing many of the T/A elements found in Series-B.

- (102) (a) Bhumij¹²
sen-lentailen-a-e?
 GO-ASP.TR.AUX.ASP.ITR-FIN-3
 'She/he was going.'
 (Ramaswami 1992:99)
- (b) Bhumij
jom-tantailen-a-ij
 eat-PROG. ITR.AUX.ASP.ITR-FIN-1
 'I was eating.'
 (Ramaswami 1992:99)

- | | |
|---|---|
| (c) Bhumij
<i>jom-tantaiken-a-ij</i>
eat-PROG.ITR.AUX.ASP.ITR-FIN-1
'I was eating.'
(Ramaswami 1992:99) | (d) Bhumij
<i>jom-ke-n-a-ij</i>
eat-ASP-ITR-FIN-1
'I was eating.'
(Ramaswami 1992:99) |
|---|---|

As alluded to above, in certain verb forms in Mayurbhanj Ho (and in Bhumij as described by Ramaswami 1992) one finds the use of a seeming future tense marker in *-e/-i-*.

- (103) *en gapa paṛa-o?* *hoba-i-redo kitep-ko*
that tomorrow read-PASS/ITR AUX-FUT-COND book-PL
cilike udub-kete en na? saben
how make.understand-DS that now all
kaji=lij udub-e-pi-a
say=1DL make.understand-FUT-2PL-FIN
'Tomorrow reading will continue and we will teach you all to understand how (to read) these books and how to say (the sounds).' [KCN]

3.2.5 Mood

Modal categories in the Ho verb include imperative, conditional, and capabilitive and optative/subjunctive (marked by *-ka*). Imperatives are formed by a stem and person marker (second singular is optionally zero or overt in Ho imperatives), and with some stems, an imperative marker as well.

- | | |
|---|--|
| (104) (a) <i>ca eya jui-ya=bu</i>
tea yes drink:3-IMP=1PL.INC
'Tea, yes, let's drink it.' [KCN] | (b) <i>nu-ya=bu</i>
drink-IMP=1PL.INC
'Let's drink.' [KCN] |
| (c) <i>am ca nuy-a</i>
you tea drink-IMP
'Drink your tea.' [KCN] | (d) <i>ace ca jui=me</i>
so tea drink:3=2
'So drink your tea.' [KCN] |
| (e) <i>jui-pe</i>
drink:3-2PL
'Drink <it>' (you-PL)! | (f) <i>jome-pe</i>
eat:3-2PL
'Eat <it>' (you.PL) |
| (g) <i>ka=laṅ juy-e</i>
NEG=1DL drink:3-T/A:FIN/IMP
'Let's not drink it.' [KCN] | (h) <i>ca ka=laṅ juy-e</i>
tea NEG=1DL drink:3-IMP
'Let's not drink the tea.' [DH] |
| (i) <i>jue=leṅ</i>
drink:3:T/A=1DL
'Let's drink (it).' [KCN] | (j) <i>ca jue=leṅ</i>
tea drink:3:IMP=1DL
'Let's drink the tea.' [KCN] |
| (k) <i>ca=laṅ juy-e</i>
tea=1DL drink:3-IMP
'Let's drink the tea.' [KCN] | |

Bhumij prefers the use of the second singular subject marker in the imperative as well, à la Santali.

- (105) Bhumij
jo jome-me
fruit eat[:3]-2
'Eat the fruit!' [CMS]

In Ho, there is a conditional formation that consists historically of a locative case marker *-re-* attached to a participial form of the verb, followed by an emphatic clitic *-do*, that is, *-redo*:

- (106) *águ-le-d-redo-ij*
bring-ANT-TR-COND-1
'if I will bring it' (Burrows 1915)

The capabillitive in *-dail-diy* is from a fused auxiliary structure (see section 3.2.12 below for more on such structures in Ho and other Kherwarian languages).

- (107) *aṭkar-diy-e-bu*
know-CAP-FUT:FIN-1PL
'We will come to know.' [CMH]

The optative/subjunctive *-ka* appears to occupy the same slot in the verb template as the finitizer *-a*, appearing between object marking and subject marking in Ho.

- (108) *enage saben etoi-e-ka bugi-leka-te ondo?*
that-FOC all learn-3-OPT do.well-ADV=ABL and
desum-re nitir-e=pe
country-LOC spread-3:FIN=2PL
'(so that) it should all have been learned well and you will spread it out in the country' [KCN]
- águ-e-ká-ij* *ema-m-ká-e*
bring-3-OPT-1 give:BEN-2-OPT-3
'I may bring it.' (Burrows 1915:55) 'He may give you.' (Burrows 1915:56)

3.2.6 Orientation/directionality

Not investigated for Ho or the other so-called minor Kherwarian languages for this study.

3.2.7 Voicelversion

Like other Munda languages, categories of grammatical voice form a part of the Ho and other Kherwarian verbal systems. Causative is realized by a number of formal strategies, primarily suffixes as in Asuri *-ge-* or through auxiliary (explicator/compound/light) verb constructions.

- (109) Asuri
duṛup-ge-tan-a-ij
sit-CAUS-PROG-FIN-1
'I cause to sit.' (Grierson 1906:139)

Only a very small number of lexical items preserve the old prefixal marker of causative (Anderson 2004, Anderson and Zide 2001) in these North Munda languages.

- (110) Bhumij
ajom- 'feed' *anu?u-* 'give to drink' (Ramaswami 1992:86)

A passive is formed by the inflectional detransitivizer (or passive/intransitive) marker of series-A and by the alternation of *-q-* with *-n-* in Series-B in Ho (as seen earlier).

- (111) *ondo?o cilike=ko paʔa-oʔ-iy-a*
 and how=PL read-PASS/ITR-FUT-FIN
 'And how they are to be read?' [KCN]

Reciprocals marked by an infix *-p-* occurs in a number of stems, but its productivity has not been tested.

- (112) *dol=laŋ kapaji-a aloka=laŋ eperaŋ*
 come=1DL talk/RECIP/-FIN PROHIB/NEG=1DL fight/RECIP
 'Come let's talk together.' 'Let us not quarrel.' (Burrows 1915:61)

An infix reciprocal *-p-* seems to be optional in the following verbal pair *darom ~ daparom* 'meet'.

3.2.8 Finiteness

The finitizer or declarative marker in *-a* appears in all finite declarative utterances. It is often but not always or obligatorily dropped in interrogative questions. Compare the following forms in this regard.

- (113) *am-a? ɖarie mena? am-a? ɖarie? mena?-a*
 YOU-GEN beard COP YOU-GEN beard COP-FIN
 'Do you have a beard?' [KCN] 'You have a beard.' [KCN]

Some elements appear to occupy the same slot as the finitizer in the verbal template and are mutually exclusive with it, for example, modal clause operators like the subjunctive, conditional, and so on.

3.2.9 Negation

Negatives are formed first and foremost by the negative particle *ka* which appears in preverbal position (and thus often with the subject marker) in finite declarative, interrogative and imperative forms.

- (114) (a) *ka=m samao-tan-a*
 NEG=2 understand:ITR-PROG-FIN
 'You don't understand.' [KCN]
 (b) *ka=laŋ juv-e* (c) *ca ka=laŋ juv-e*
 NEG=1DL drink:3-T/A:FIN/IMP tea NEG=1DL drink:3-IMP
 'Let's not drink it.' [KCN] 'Let's not drink the tea.' [DH]

Negative copular forms are made off of the stem *ban[o]*.

- (115) *ama? upunia hon-ko ban-ko-a*
 YOU-GEN four child-PL NEG.COP-PL-FIN
 'You don't have four children.' [CMH]

Other minor Kherwarian languages have *ka* as well, for example, Turi. Chaibasa Ho on the other hand prefers the *ban* element that functions as a negative copula in Mayurbhanj Ho.

- (116) Turi
ini-ke ka=ko em-a-i-ke-n-a
 he-DAT/ACC NEG=3PL give-BEN-3-ASP-ITR-FIN
 ‘They didn’t give him.’ (Grierson 1906:130)

The *alolalu* prohibitive form on the other hand seems to be lacking or little used in Mayurbhanj Ho.

- (117) (a) Turi (b) Jashpur Korwa
alu do-ij-me alo-i bol
 PHB keep-1-2 NEG-3 enter
 ‘Don’t keep me!’ ‘He did not enter.’
 (Grierson 1906:130) (Grierson 1906:151)

Both *alo* and *ka* may be used together as well in Chaibasa Ho.

- (118) *aloka=laj eperaj*
 PROHIB/NEG=1DL fight/RECIP
 ‘Let us not quarrel.’ (Burrows 1915:61)

3.2.10 Derivation

Derivation *per se* in Ho verb stems has not been examined. Some verbs seem to have a (lexicalized?) *-Vn* suffix to mark reflexive action (e.g. *goe?* ‘kill’ *goe?en* ‘kill self’). How productive this is is currently unknown.

3.2.11 Noun incorporation and combining forms

Unlike South Munda languages where incorporation is widely attested or even productive, there is little evidence of noun incorporation in North Munda languages as a whole. However, much as ‘poetic’ language preserves certain phonological archaisms (e.g. intervocalic *-r-), so too does it hold (morpho)syntactic archaisms as well, here in the form of VN order in some fixed frozen expressions in an archaic form used in prayers, just as in older forms of incorporation attested across South Munda languages (see Anderson 2004, 2007 for more).

- (119) *jom-mandī nu:-da?-ka-e?*
 eat-food drink-water-OPT-3
 ‘May she/he take a dinner’ (Deeney 1978a:248)

3.2.12 Auxiliary verb constructions

Fused or synchronically (semi) bi-partite auxiliary verb constructions or complex predicate structures form an integral part of Ho sentence and verbal structure. As alluded to above, the tense/aspect forms discussed above all seem themselves to derive from a stage in [(pre)Proto-]North Munda when complex predicate structures marked a range of aspectual categories and tense marking was split according to two mainly semantically determined inflectional classes, roughly transitive and intransitive (or active/neutral). Such a system has been preserved more or less to this day in South Munda Juang (see Anderson 2001).

While Munda languages as a whole show considerable formal and functional variation with respect to the structure of auxiliary verb constructions [AVC] (as the whole set of complex predicate types is known), Kherwarian languages in their current state prefer AUX-headed and split inflectional configurations, often the former with zero-marked lexical verbs, probably resulting from a nuclear serialized structure (Anderson 2006). The number of such elements that appear as Verb₂ or the ‘auxiliary’ in such AVCs or complex predicate subtypes in Ho and the minor Kherwarian languages is quite large, and the system seems to allow new members rather easily as needed by speaker or speech context. Some examples are given below.

- (120) (a) *uʔi=ben-ten-e*
 jump-AUX-PROG-FIN
 ‘jumping’ [CMH]
 (c) *jom hoba-n-a*
 eat AUX-ITR-FIN
 ‘Eating will continue/happen.’
 [CMH]
 (e) *aʔkar-ɖiy-e-bu*
 know-CAP-FUT:FIN-1PL
 ‘we will come to know’ [CMH]
- (b) *nir=ban-tan-a*
 run-AUX-PROG-FIN
 ‘running’ [CMH]
 (d) *inuʔ hoba-n-a*
 play AUX-ITR-FIN
 ‘Playing will continue/happen.’
 [CMH]
- (121) (a) Turi
go:t-cha:ba:ta:n-a-ku
 gather-COMPL-ASP-ITR-FIN-PL
 ‘They have gathered.’
 (Grierson 1906:131)
- (b) Asuri
goj-doho-le-n-a
 die-AUX-ASP-ITR-FIN
 ‘had been dead’
 (Grierson 1906:141)

Note that the progressive marker, which, probably derives from Series-B intransitive T/A inflections – which themselves likely reflect grammaticalizations of (now opaque) tense-marked auxiliary verbs – is seen in the Turi form above (*ta-n*) in a non-present, non-progressive function.

The progressive and imperfect forms of the verb in Kherwarian Munda are either synchronic AVCs or undergoing phonological fusion in the modern states of the various individual languages.

- (122) ‘Raigarh Mañjhi’
para:-ya-n-a:
 fall-ASP-ITR-FIN
 ‘He fell.’ (Grierson 1906:146)
- (123) Jashpur Asuri
holate ij huʔu ir-ij sen-tehin-en-a-ij
 yesterday I paddy cut-1 go-T/A-ITR-FIN-1
 ‘Yesterday I went and cut rice.’ (Grierson 1906:142)

Note the double subject marking in this form.

- (124) Ho
aliʔ baro ca ju-te-liʔ sanay-tan-a
 we.DL two tea drink-ABL-1DL desire-PROG-FIN
 ‘We two wish to drink tea.’ [CMH]

3.3 Expressives

Reduplicated expressive elements may serve as verb stems, and bear a range of (at times unexpected) inflectional suffixes. Note the following paradigm of *mēmē* ‘bleat (of a goat)’.

- (125) (a) *merom mēmē-tan-a=e?*
 goat bleat-PROG-FIN=3
 ‘The goat is bleating.’ [CMH]
 (b) *merom mēmē-e-a=e?*
 goat bleat-FUT-FIN=3
 ‘The goat will bleat.’ [CMH]
 (c) *merom mēmē-ke-ɖ-a=e?*
 goat bleat-T/A-TR!-FIN=3
 ‘The goat bleated.’ [CMH]

It appears with the transitive past form which is quite unexpected. This remains to be checked with other speakers before an explanation is offered. It is true that some monovalent verbs unexpectedly belong to the inflectional class that takes *-ke-ɖ-* in the ‘aorist/perfect/past’. Other expressive reduplicated elements may function as verb stems in a similar manner in Ho.

- (126) *seta bu?bu?-tan-a=e?*
 dog bark-PROG-FIN=3
 ‘The dog is barking.’ [CMH]

Reduplication is common in a number of lexical items, often ones that have distributed or collective semantics. As mentioned above, it also appears to have limited productivity in producing adjectives with vaguely distributive, collective, and so on, semantics derived from verbs or unattested simplex stems.

- | | | |
|---|--|--------------------------|
| (127) <i>sun ~ susun</i> ‘dance’ | <i>nu:</i> ‘drink’ | <i>munu</i> ‘nurse (tr)’ |
| <i>ipi?ipi?</i> ‘star twinkle < <i>ipil</i> ‘star’ | <i>sasan</i> ‘burial place’ | |
| <i>balbal</i> ‘perspire/sweat’ | <i>sisir</i> ‘dew’ | |
| <i>dudu:</i> ‘steam’ | <i>sosoe?</i> ‘tree sap’ | |
| <i>ɖeɖem</i> ‘sparrow’ | <i>didi</i> ‘vulture’ | |
| <i>baba</i> ‘husked rice, paddy’ | <i>sarsar</i> ‘fingernail’ | |
| <i>tepe?tepe?</i> ‘shallow’ | <i>lolo</i> ‘hot’ | |
| <i>sasaj</i> ‘yellow’ | <i>mamaraj</i> ‘proud’ | |
| <i>sasa:</i> ‘cold’ | <i>tanga-tanga</i> ‘separate (adj)’ < <i>tanga</i> | |
| | ‘separate (v)’ | |
| <i>jü:l ~ jü:jü:l</i> ‘slip’ > slipper | <i>juwa</i> ‘to stick’ > <i>jujuwa</i> ‘sticky’ | |
| <i>jerem-jerem</i> ‘sticky’ < <i>je:ʔr</i> ‘to stick’ | | |

A small number of lexical items share sufficient lexical and phonological content to believe them to be in some kind of ‘affective’/expressive relationship to one another. One such pair involves a kind of initial [de-]voicing (the ‘base’ form is unclear and therefore the process as well): *jur* ‘surround’ vs. *cur* ‘surround so as to prevent escape’; *jeŋga ~ jeŋged ~ juŋguɖ* ‘degrees of red’. Some seem to involve the same kind of vowel alternation (plus stem augmentation) typical of the highly developed and rich expressive [re]duplication and ablaut system described below.

A preliminary survey of the Ho lexical materials (e.g. Deeney 1978b) reveals rich use of expressive, mimetic or echo reduplication patterns of a type found across the

Munda family, and more broadly across the Eurasian continent as an areal feature. These include (128) full reduplication, (129) full reduplication with overwriting of one or more vowels, and (130) full reduplication with overwriting (or prothesis) of an initial consonant, typically [s], [m] or [b]. Semantically these encode a range of meanings: intensity, iterativity, continuity (esp. of sounds), rapidity, distributivity, diversity, and so on.

- (128) *mir-mir* ‘twilight’
hiti-hiti ‘sensation which precedes fainting’ (= English ‘seeing stars’)
biur-biur ‘tortuous’ (< *biur* ‘to turn’)
banka-banka ‘roundabout’
joroy-joroy ‘everlasting’
ungud-ungud ‘doubled up’
rusu-rusu ‘ague’
saq-saq ‘the sound of reaping and grazing’
sae-sae ‘to look repeatedly at from the corner of one’s eyes’
a: col-col ‘a long beak of a stork’ (*a:* ‘mouth, beak’)
- (129) *sida-sada* ‘frank’, ‘open’, ‘simple’
dampa-dumpu ‘to stagger’
ɖaŋ-ɖuŋ ‘to swing’
balu-balu ~ *bala-balu* ‘mad (of animals only)’
saq-sod ‘to do something very quickly or in a very short time’
- (130) *amna-samna* ‘face to face’
kili-mili ‘various, different’ (< *kili* ‘a tribe, sect’)
sago-bago ‘a low rumbling noise, for example, of the breathing of a weak or dying person ... or the sound of fermenting rice beer, or the sound made by a crowd of flying ants moving around in their hole before coming out’
sagor-bagor ‘the sound made by rice being cooked when the water is very little, the sound made by a cow eating straw’

4 SYNTAX

4.1 Syntax of the simple sentence

One distinction that can be made for many Kherwarian languages syntactically is between ‘verbal’ and ‘copular’ sentences. The former is the majority of sentences while the latter contains one of a number of elements filling the role of clause-final, finitizable copula. Verbal clauses can be declarative, interrogative, or imperative.

- (131) Bhumij
jo jome-me
 fruit eat:3-2
 ‘eat the fruit!’ [CMS]
- (132) Ho
 (a) *mandi jom-ke-ɖ-a-m*
 rice eat-T/A-TR-FIN-2
 ‘you ate rice’ [CMH]

- (b) *cenaʔ=ko menaʔ-a oŋ*
 what=3PL say-FIN oŋ
 ‘What do they call it?’ [KCN]

The basic forms of the copula are *menaʔ* (~ *meniʔ* optionally if animate singular) in the positive and *ban[o]* in the negative. The copula can be used in both existential and possessive functions.

- (133) (a) *cimiŋ ho:ko Bhubaneswar-re menaʔ-ko*
 how.many person-PL Bhubaneswar-LOC COP-PL
 ‘How many people are there in Bhubaneswar?’ [CMH]
 (b) *ay-eʔ era menaʔ*
 3:GEN wife COP
 ‘Is he married?’ [DH]
 (c) *kadal daru maraŋ maraŋ sakam-ko menaʔ-a*
 banana tree big big leaf-PL COP-FIN
 ‘The banana tree; it has really big leaves.’ [KCN]
 (d) *am-aʔ barea hon-ko menaʔ-a=ko*
 you-GEN two child-PL COP-FIN-PL
 ‘you have two children’ [CMH]
 (e) *am-aʔ ɖarie menaʔ* (f) *am-aʔ ɖarie menaʔ-a*
 you-GEN beard COP you-GEN beard COP-FIN
 ‘Do you have a beard?’ ‘You have a beard.’
 [KCN] [KCN]

Note also

- (134) Bhumij
ne aŋn-a[ʔ] era menaʔ-ye
 yes I-GEN wife COP-FIN
 ‘Yes, I have a wife.’ [CMS]

Examples with negative copula forms include the following:

- (135) (a) Ho
am-aʔ upunia hon-ko ban-ko-a
 you-GEN four child-PL NEG.COP-PL-FIN
 ‘you don’t have four kids’ [CMH]
 (b) *ɖariye aŋn-aʔ bano[ʔ]-a*
 beard I-GEN NEG.COP-FIN
 ‘I don’t have a beard.’ [KCN] 6:05

Identificational copula constructions seem to involve the element *kan-a*, similar to the progressive/present form in Santali, in this idiolect from Gungupari village, Mayurbhanj district, Orissa.

- (136) (a) *ne-ya-do aŋn-aʔ boʔ kan-a*
 this-FOC-EMPH I-GEN head:COP-FIN
 ‘This here is my head.’ [KCN] 3:50
 (b) *ne-ya-do aŋn-aʔ muʔte kan-a*
 this-FOC-EMPH I-GEN nose:COP-FIN
 ‘This here is my nose.’ [KCN] after 3:50

- (c) *ne-ya-do me²q kan-a*
 this-FOC-EMPH eye:COP-FIN
 ‘This here is (my) eye.’ [KCN]
- (d) *ne-ya-do ti: kan-a*
 this-FOC-EMPH hand:COP-FIN
 ‘This is (my) hand, arm.’ [KCN]

This element is optional however:

- (137) (a) *ne-ya-do qãta*
 this-FOC-EMPH tooth
 ‘This here (is my) tooth.’ [KCN]
- (b) *ne-ya-do alay*
 this-FOC-EMPH tongue
 ‘This here (is my) tongue.’ [KCN]
- (c) *ne-ya jua*
 this-FOC cheek
 ‘This (is my) cheek.’ [KCN]
- (d) *ne-ya-do lutur*
 this-FOC-EMPH ear
 ‘This here (is my) ear.’ [KCN]
- (e) *ne-ya-do hofo?*
 this-FOC-EMPH throat
 ‘This here (is my) throat.’ [KCN]
- (f) *ne-ya-do kaða*
 this-FOC-EMPH foot/leg
 ‘This here (is my) foot.’ [KCN]

Occupying the same functional space as this *kana* copula element may also be found the Ho ‘progressive’ or Series-A marker in a verbal construction. Perhaps the *kana* forms are best understood as a special use of a Series-B form in a copular (present) construction.

- (138) *ne-ya-do sarsar-tan-a*
 this-FOC-EMPH nail-PROG-FIN
 ‘This here is (my) nail.’ [KCN]

That no copula is obligatory is seen above with the identificational forms and in the following Wh-question with the interrogative element in clause-final position.

- (139) *ay-a? nutum ci-ka-n-a*
 3-GEN name what-T/A-ITR-FIN [or what]
 ‘What’s his name?’ [KCN]

4.1.1 Typological features

SOV is the basic clausal constituent order found in Ho. Other orders (SVO, OSV) are permitted in various circumstances, for example, focus, topicalization, and so on, but the details of this remain a subject for future research.

- (140) SOV
aliy baro ca ju-te=liy sanay-tan-a
 we.DL two tea drink-ABL=1DL desire-PROG-FIN
 ‘We two wish to drink tea.’ [CMH]
- (141) SVO
cimiy hor-!=ko jagar-tan-a ho kaji
 how.many people-[/=3]PL speak-PROG-FIN Ho language
 ‘How many people speak Ho language?’ [KCN] 20:40
- (142) OSV
Bhumij kaji cimiy hor-ko jagar-tan-a
 Bhumij language how.many person-PL speak-PROG-FIN
 ‘How many people speak the Bhumij language?’ [KCN]

Within the noun phrase, adjectives, demonstratives, and genitives precede the noun, while numerals may precede or follow the governed noun. Adverbs precede both adjectives and verbs, while auxiliaries follow the lexical verb as is typical in Eurasian SOV languages (Anderson 2006). Within possessive sentences the order can be Gen N or N Gen (either order between possessor and possessum is permitted).

(143) *am-a? seta mena?a-koa*
 you-GEN dog COP-PL:FIN
 ‘Do you have any dogs?’ [KCN]

(144) *saxdom aŋ-a? ban-kue*
 horse I-GEN NEG.COP-PL:FIN
 ‘I have no horses.’ [KCN]

4.2 Complex sentence structure

4.2.1 Relative-type clauses

There are no spontaneous forms of this type in our corpus. Non-finite verbs marked by *-te* (see below) may appear in clauses that have relative-type functions.

(145) *nen-te saben=ko ajaraj-te nen-do-ko meta-?a omm*
 this-FOC/EMPH all=PL come.first-ABL that-EMPH-PL say-PASS/ITR-FIN *Om*
 ‘That one there, the ones that always come first, those are said “Om.”’
 [KCN]

4.2.2 Other subordinate clauses (time, manner, cause, purpose, and so on)

One particularly common means of marking subordinate clauses in Ho is through a system of case-marked clausal subordination (cf. Anderson 2002 for an elaborate such system in Burushaski). The general oblique or ablative marker marks complements of certain head verbs in various complex predicate structures in Ho.

(146) *aliŋ baro ca ju-te=liŋ sanaŋ-tan-a*
 we.DL two tea drink-ABL=1DL desire-PROG-FIN
 ‘We two wish to drink tea.’ [CMH]

Note that as is commonly the situation in the history of complex predicates, the case-marked (same subject) verb complement structure has the same formal characteristics to the (same subject) serialized formation.

(147) *goe?-te jonom-le-n*
 die-ABL born-ANT-ITR
 ‘stillborn’

Attached to a participle- (or T/A+transitivity)-marked stem, temporally subordinate clauses may be formed.

(148) (a) *seno-ja-n-te* (b) *paŋa-ja-n-te* (c) *jome-ja-n-te*
 go-T/A-ITR-ABL read-T/A-ITR-ABL eat-T/A-ITR-ABL
 ‘after going’ ‘after reading’ ‘after eating’ [CMH]

- (d) *hoba-o?-tan-a* *cilike* *lele-do* *ne?enen talu-re*
 happen-PASS/ITR-PROG-FIN how tongue-EMPH this.one jaw-LOC
joga-o-ka-n-te *ra?a-ja* *cilike* *wã wã wã*
 touch-PASS/ITR-T/A-ITR-ABL cry-[TA:]FIN how wã wã wã
 ‘is like this when the tongue touches the upper jaw and it cries like
wã wã wã’ [DB]

Another formation involves a finitizer or a genitive marker added to the case and makes a nominal sentence of the following type.

- (149) *ol-e-ti-a*
 write-T/A-ABL-FIN/GEN??
 ‘this is to write with’ [CMH]

Attaching to a T/A form (with or without the transitivity marker), the locative case marker in *-re* may form temporally subordinate clauses as well:

- (150) *ena-do* *ne-ya-do* *mis-ete* *ule-tan-re*
 that-EMPH this-FOC-EMPH throat-ABL vomit-PROG-LOC
jantan-ko-re *ka=bu* *men-iy-a*
 different.places-PL-LOC NEG=1PL.INC say-FUT-FIN
 ‘That’s not what we call that one, this is when vomit comes out all over.’ [KCN]

Mixed structures are also found with case-marked verbs in combination with (co-) relative pronominal-type elements. Such mixed varieties are found in high-contact varieties of Siberian languages, where such subordination strategies abound but are lacking in the socially dominant Russian (for more see Anderson 2003, 2004).

- (151) *ne-ya-do* *hoba-tan-a* *cimitaj* *citi=e?* *jonom-o?-a*
 this-FOC-EMPH happen-PROG-FIN when baby=3 birth-PASS/ITR-FIN
imitej *si?i=e?* *jonom-ta-re* *ra?a-ya* *ina-do*
 then small.child=3 birth-T/A-LOC cry-[T/A:]FIN that-EMPH
 ‘So it is that when a baby is born, at that time it starts to cry that one.’ [KCN]

On the other hand, a possibly innovative feature of complex sentence structure in the speech of younger Ho speakers is the presence of a clause-initial subordinator and two finite verbs rather than a case-marked subordinate verb.

- (152) *ne?ene-ya* *abu-do=bu* *mene-tan-a* *cimitaj*
 this.one-FOC we-EMPH=1PL.INC say-PROG-FIN when
abu=bu *bojga-e-tan-a*
 we=1PL.INC worship-3-PROG-FIN
 ‘This one here we say it when we are worshipping.’ [DB]

4.2.3 Coordination and switch reference

Clauses can be coordinated by the conjunction *ondo[?o]* or by simple concatenation. In the latter structure, the last verb is sometimes the only fully finite one.

- (153) *enleka* *uru?-i-pe* *ondo* *kaji-pe*
 like.that think-FUT-2PL and speak-FUT-2PL
 ‘Think about it like that and say it.’ [KCN]

ean-ge *saben* *etoi-e-ka* *bugi-leka-te* *ondo?o*
 that-FOC all learn-3-OPT do.well-ADV=ABL and

desum-re *nitir-e=pe*
 country-LOC spread-FUT:FIN=2PL

‘(So that) it will all have been learned well and you will spread it out in the country.’ [KCN]

(154) Jashpur Asuri

holate *ij* *hu?u* *ir-ij* *sen-tehin-en-a-ij*
 yesterday I paddy cut-1 go-T/A-ITR-FIN-1

‘Yesterday I went and cut rice.’ (Grierson 1906:142)

Note the double subject marking in this form but tense/aspect marking only on the final conjunct. This is an example of a semi-finite conjunctive structure.

5 SEMANTICS/DISCOURSE

5.1 Semantics

Semantics in Ho have not been extensively investigated. Some things were touched on briefly in section 3.3 above in the discussion of expressive formation. Another feature of Ho that would be apparent to anyone casually looking through the Ho dictionary of Deeney (1978a) is the fine-grained semantic nature of many Ho verbs. Thus, there are numerous verbs translatable as English ‘cut’, depending on motion used, instrument, and so on. A small selection of these are offered below.

(155) *hār* ‘cut with small instrument’

oq ‘cut down, by self’

ma?a ‘cut with axe, heavy strokes’

ir ‘cut paddy’

ged ‘cut (meat) or w/saw’

sa ma?a ‘cut into small pieces’ [CMH]

5.2 Discourse

The study of the discourse structure of Ho and the minor Kherwarian languages is in its very infancy. Only a brief overview and some impressionistic comments are offered below on Ho conversation.

In emphatic speech, multiple references to a single referent is possible in Ho.

(156) *abu* *hotel-te=bu* *seno?-tan-a=bu*

we hotel-ABL=1PL.INC go-PROG-FIN=1PL.INC

‘we are going to/from the hotel’ [CMH]

Dropping of noun phrases recoverable from the discourse is commonly attested in spoken varieties of Mayurbhanj Ho. Copular forms, like verbs, can themselves alone serve as answers to questions.

(157) (a) *am-a?* *mena?-k^wa*

YOU-GEN COP-PL:FIN

‘Do you have any?’ [KCN]

- (b) *ban-ku-e*
 NEG.COP-PL-FIN
 ‘(I’ve) got none.’ [KCN]

6 LEXICON

6.1 Austroasiatic/Munda components

Ho like any other language reflects numerous layers of its history in the make up of its lexicon. Austroasiatic and Munda components are found alongside Indo-Aryan, English, and other loan strata.

The Kherwarian languages are closely related to each other. The language/dialect/communalect/ethnolect continua are complex with respect to the Kherwarian Munda languages. There are groups that are linguistically very similar but whose ethnocommunity identities are quite distinct. One such group consists of the Mah[aj]li, Karmali, and Santali. Also, variation of a not-insignificant sort (i.e. cross-cutting ostensible language-defining structural boundaries) is found, with features of individual idiolects showing one or more Kherwarian language features, for example, in multi-lingual Kherwarian villages in Mayurbhanj, Orissa, where ethnolinguistic identity is largely Ho but linguistic features of individual community members may show more Mundari or even Santali-like features (see below for more and examples). Suffice it to say that how many Kherwarian ‘languages’ or ‘speech varieties’ there are remains an open question, and how these can be meaningfully distinguished using linguistic criteria, rather than ethnolinguistic or sociohistorical identity is a subject for future research. All of this should be borne in mind when considering the discussion which follows.

The correspondence among Kherwarian languages including Santali and Mundari is remarkable in Table 5.1. The following phonological correspondences and discrepancies among the Kherwarian languages should be noted:

- (i) The identical correspondences like (1) and (2) are copious. We can find easily word correspondences of these words in Austroasiatic. Thus (1) ‘forest’ *bri* in Sre, *pri?* in Riang, *brei* in Palaung, *brai* in Khmer and *brai?* in Wa; (2) ‘thigh’ *bhlau* (*phlou*) in Khmer, *blaul bleu* in Palaung and *pelau* in Tareng.¹³
- (ii) Intervocalic **ɾ* is lost in Ho but retained in other Kherwarian in (3) and (4).
- (iii) Intervocalic **d̪* is changed to *ɾ* in Ho, Mundari, Birhor, and Turi but retained in Santali, some dialects of Mundari, Korwa, and Asuri in (5).
- (iv) Intervocalic **h* is lost in Ho and Mundari but retained in the other Kherwarian languages in (6) and (7).
- (v) The position of articulation of checked consonant in word-final position is a retroflex *d̪* in Ho and Korwa but a dental *d* in the other Kherwarian languages in (8) and (9).
- (vi) The mid-vowel in Santali and Korwa corresponds to the low vowel in other Kherwarian in (10) and (11). The vowel correspondences of mid-vowel *e*, *o* in Santali and *i*, *u* in other Kherwarian, especially Mundari and Ho are already discussed in Pinnow (1959), Zide (1966), Zide and Munda (1966), and Osada (1996). According to Minegishi (1990) and a recent report, that is, Kobayashi *et al.* (2003), the high vowel correspondences *i*, *u* are found even in some dialects of Santali. Thus we have given these cases in { } of (10) and (11).

TABLE 5.1: A COMPARATIVE VOCABULARY IN THE KHERWARIAN LANGUAGES

	Santali	Mundari (Hasada/Naguri)	Ho	Bhumij	Korwa	Asuri	Birhor	Turi	References
1 'forest'	<i>bir</i>	<i>bir</i>	<i>bir</i>	<i>bir</i>	<i>bhir</i>	<i>bir</i>	<i>bir</i>	<i>bir</i>	P:V321; M:52
2 'thigh'	<i>bulu</i>	<i>bulu</i>	<i>bulu</i>	<i>bulu</i>	<i>bul</i>	<i>bulu</i>	<i>bulu</i>	<i>bulu</i>	P:V145; M:65
3 'human'	<i>hɔɽ</i>	<i>hoɽo</i>	<i>ho:</i>	<i>hɔɽɔ</i>	<i>hoɽ</i>	<i>hɔɽ</i>	<i>hɔɽ</i>	<i>hoɽ</i>	P:V311; M:160
4 'house'	<i>oɽak'</i>	<i>oɽa?</i>	<i>owa?</i>	<i>oɽa?</i>	<i>oɽa?</i>	<i>oɽa?</i>	<i>oɽa?</i>	<i>oɽa?</i>	P:V200; M:90
5 'duck'	<i>geɽɛ</i>	<i>geɽe/geɽe</i>	<i>geɽe</i>	—	<i>geɽe</i>	<i>geɽe</i>	<i>geɽe</i>	<i>geɽe</i>	P:V160
6 'flower'	<i>baha</i>	<i>baɽ/baha</i>	<i>ba:</i>	<i>baha</i>	<i>baha</i>	<i>baha</i>	<i>baha</i>	<i>baha</i>	P:V21; M:123
7 'to place'	<i>dɔhɔ</i>	<i>doɽ/doho</i>	<i>do:</i>	<i>doo</i>	<i>doho</i>	<i>doho</i>	<i>doho</i>	<i>doho</i>	P:V262; M:95
8 'to plant'	<i>bit'</i>	<i>bid</i>	<i>biɽ</i>	<i>bid</i>	<i>biɽ</i>	<i>bid</i>	<i>bid</i>	<i>bid</i>	P:V285; M:51
9 'mushroom'	<i>ot'</i>	<i>ud</i>	<i>uɽ</i>	<i>ud</i>	<i>uɽ</i>	<i>ud</i>	<i>ud</i>	<i>ud</i>	P:V237; M:74
10 'arm'	<i>sopo</i> { <i>supu</i> }	<i>supu</i>	<i>supu</i>	<i>supu</i>	<i>sop</i>	—	<i>supu</i>	<i>supu</i>	P:V134; Mu:76
11 'ebony'	<i>terel</i> { <i>tiril</i> }	<i>tiril</i>	<i>tiril</i>	<i>tiril</i>	<i>terel</i>	<i>tiril</i>	<i>tiril</i>	<i>tiril</i>	P:V227; M:74
12 'mango'	<i>ul</i>	<i>uli</i>	<i>uli</i>	<i>uli</i>	<i>ul</i>	<i>ul</i>	<i>ul</i>	<i>ul</i>	P:V144; Mu:61
13 'tiger'	<i>kul</i>	<i>kula</i>	<i>kula</i>	<i>kula</i>	<i>kul</i>	<i>kul</i>	<i>kul</i>	<i>kul</i>	P:V281; M:61

- (vii) The vowel in word-final position in some words is dropped in Santali, Korwa, Asuri, Birhor, and Turi but retained in Mundari and Ho in (12) and (13). A notable example is *ot* 'earth' in Santali, Korwa, Asuri, Birhor, and Turi but *ote* in Mundari and Ho.
- (viii) The intervocalic *-m-* in Santali and some dialects of Mundari corresponds to a consonant cluster; that is, of homorganic nasal and stop *-mb-* in other Kherwarian shown in (14) and (15). But the distribution of the following case is slightly different; for example, *rimil* 'cloud' in Ho, Santali, some dialects of Mundari while *rimbil* in Mundari.
- (ix) Two different forms are distributed among Kherwarian in (16) and (17). These are usually divided into two groups; that is, Santali, some dialects of Mundari, Korwa, Asuri, Birhor, Turi on the one hand, Ho and Mundari on the other hand. Such instances are as follows: *rama* 'nail' in Santali and so on while *sar-sar* in Mundari and Ho; *hoṛo* or *huṛu* 'paddy' in Santali while *baba* in Mundari and Ho, except Mayurbhanj Ho which often uses *huṛu*.
- (x) The sporadic vowel correspondences are found in (18) and (19). Additional cases include the following instances: *taren* 'shoulder' in Santali while *taran* in other Kherwarian; *laser* 'blade' in Santali while *leser* in other Kherwarian.
- (xi) The initial *ɟ-* in Santali, Korwa, and Asuri is correspondent to the initial *n-* in other Kherwarian languages in (20) and (21).
- (xii) The CaC(C)e(C)/CeC(C)a(C) and CaC(C)o(C)/CoC(C)a(C) of the disyllabic words in Santali corresponds to the CaC(C)i(C)/CiC(C)a(C) and CaC(C)u(C)/CuC(C)a(C) in other Kherwarian languages in (22), (23), (24) and (25). In these cases Osada (1996) suggested that the high vowel *i/u* became lowered due to the assimilation of the low vowel *a* in Santali under the influence of an adjoining Indo-Aryan language; that is, Bengali in which the same assimilation can be found according to Dimock (1957).

6.2 Loan strata

Ho is covered quite excellently in the lexical materials of Father Deeney in which numerous loanwords are identified. As mentioned in section 3.1.8 above, numeral stems in Bhumij even down to 'five' reflect Indo-Aryan sources.

7 BRIEF ANALYZED TEXT

The following text was recorded by Anderson and Harrison in Bhubaneswar, India, September 2005. It is a lesson given to a group of Ho children by a well-known Ho orator and performance artist, Mr. K.C. Naik Biruli. In this lesson, he emphasizes the spiritual and cultural significance of Ho writing. The syllabary begins with 'om', a grapheme of religious significance that serves no orthographic purpose. Each successive grapheme has a mythology that relates its shape, sound, or appearance to the unique Ho cultural context and worldview. This mythology serves as a mnemonic used in teaching and memorizing the Ho syllabary.

- (i) *joarge* *hon-ko*
greetings child-PL
'Greetings children!'

- (ii) *enredo abu abu-e? ho: hayam-te ho: bhasa-te*
 ok we we-GEN Ho language-ABL Ho language-ABL
ho: language-te tisij
 Ho E.language-ABL today
 ‘OK today we are going to (talk) about our Ho language.’
- (iii) *ciliken lipi-ko mena? -a*
 what.kind letter-PL COP-FIN
 ‘What kinds of letters there are?’
- (iv) *ondo?o cilike=ko paʔa-oʔ-iy-a*
 and how=PL read-PASS/ITR-FUT-FIN
 ‘and how they are read’
- (v) *ena-ko ape-ko samaj-re=lij jagar-iy-a*
 that-PL you.PL-PL front-LOC=1DL.EXCL speak-FUT-FIN
 ‘I (POL) will speak about these in front of you.’
- (vi) *ena-ge saben etoi-e-ka bugi-leka-te ondo?o*
 that-FOC all learn-3-OPT do.well-ADV=ABL and
desum-re nitir-e=pe
 country-LOC spread-FUT:FIN=2PL
 ‘(so that) it will all have been learned well and you will spread it out in the country.’
- (vii) *ayum-ke-q-a=pe ondo?o may-ta-ko babu-ta-ko*
 listen-T/A-TR-FIN=2PL and girl-EMPH-PL boy-EMPH-PL
 ‘Did you listen girls and boys?’
- (viii) *tisij alij kaji tey-aʔ-a-do nen citi-do-ko*
 today 1DL.EX say AUX-ITR/PASS-FIN-EMPH this writing-EMPH-PL
meta-ʔ-a waraj citi lipi
 say-ITR/PASS-FIN waraj citi letter
 ‘Today I will speak about these ways of writing, that are called the Warang Chiti letters.’
- (ix) *ne-ya-do Pandit Lako Bod[o]ra tikiʔ nen*
 this-FOC-EMPH Pandit Lako Bodra himself this
lipi-do=kij olen=taiken-a
 letter-EMPH=3DL write=PROG.PST-FIN
 ‘This Pandit Lako Bodra himself produced in writing these letters.’
 [NB: use of dual in polite discourse for singular referent]
- (x) *nen citi-re-a? nutum-do cikana-ʔ-a waraj citi*
 this writing-GEN:INAN name-EMPH what-FIN waraj citi
 ‘What is this way of writing(? it is) the Warang Chiti.’
- (xi) *cinakana-ʔ-a waraj citi*
 call-RFLXV-T/A-ITR-FIN waraj citi
 ‘was called Warang Chiti’

- (xii) *nen-te* *saben=ko* *ajaraj-te* *nen-do-ko*
 this-FOC/EMPH all=PL come.first-ABL that-EMPH-PL
meta-ʔ-a *omm*
 say-PASS/ITR-FIN Om
 ‘That one there, the ones that always come first, those are said “Om.”’
- (xiii) *cinaka-a* *meta-ʔ-a* *omm*
 call-PASS/ITR-FIN say-PASS/ITR-FIN Om
 ‘How is it called? It is said Om.’
- (xiv) *ena* *canaʔp-te* *ne-ya-do* *siʔie-ko* *raʔay-e=ko-leka*
 that finish-ABL that-FOC-EMPH small.child-PL cry-FUT=PL-PP
wā *cenaʔa=ko* *kaji-e* *wā*
wā what=3PL say-FUT:FIN *wā*
 ‘That next one there, that one is like as babies cry, *wā*, what do they say, *wā*.’
- (xv) *saben-ko* *ajte* *ayar-re* *siʔi-eʔ* *jonom-eya-n-[d]re wā*
 all-PL first come.first-LOC small.child-3 be.born-T/S-ITR-LOC wā
mente *raʔay* *aci* *bano*
 like cry:FUT:FIN or not
 ‘When all of them first come out, when a small child is born, (he) cries *wā*, isn’t it so?’
- (xvi) *manoa* *jonom-re* *ratom* *kaji* *eʔeʔ-ti-a-do*
 human birth-LOC just.at? speak AUX-T/A-FIN-EMPH
 ‘Just when a human is born, [s]he starts to speak.’
- (xvii) *wa* *mente* *ma* *mente* *kaʔ=eʔ* *kaji-e* *wa* *mente* *kaji-e[ʔ]*
wa like *ma* like NEG=3 speak-FUT *wa* like say-FUT-?
 ‘like *wā* s/he speaks not *ma*, s/he speaks like *wā*’
- (xviii) *ne-ya-do* *maraj* *citi* *ne-ya-do* *huriʔ*
 this-FOC-EMPH big case this-FOC-EMPH small
 ‘This one is the upper case and this one is the lower.’
- (xix) *enete* *ena* *canaʔp-te* *siʔie-do-ko=eʔ* *maraj-ya-n-[d]relte*
 then that finish-ABL little.child-EMPH-PL=3 grow-T/A-ITR-LOC/ABL
cinaʔ *men-iy-a*
 what say-FUT-FIN
 ‘Then this next one, little children, when they grow up, what do (they) say?’
- (xx) *eʔga-do* *aʔay-ta-ye* *ka=ko* *men-iy-a*
 mother-EMPH set.free-T/A-3:FIN NEG=3PL say-FUT-FIN
 ‘The mother sets them free, isn’t it so?’
- (xxi) *enleka* *ne-ya=ko* *men-iy-a* *aʔ* *cilike* *kaji-re*
 like.that this-EMPH=3PL say-FUT-FIN set.free how say-LOC
 ‘It’s like that, how they will say it, like *aʔ*’
- (xxii) *siʔi=eʔ* *aʔa-k-i-redo* *iniʔi-do* *sen-iy-a*
 child=3 set.free-PASS/ITR-FUT-COND those-EMPH go-FUT-FIN
 ‘When the child is set free, those ones will go.’

- (xxiii) *ban-ta-redo iyu?-e yu?e iyu?-e ban-ta-re yuw-e*
 not-EMPH-COND jump.up(3x) not-EMPH-LOC jump.up-FUT:FIN
parkom-ete iyu?-e
 bed-ABL jump.up-FUT:FIN
 ‘If not, (s/he will) jump up, if not, (s/he) will jump up from the bed.’
- (xxiv) *ente ni-ya-do=ko men-iy-a uy?*
 then this-FOC-EMPH=3PL say-FUT-FIN *uy?*
 ‘Then this is how they say this one *uy?*’
- (xxv) *cina?=ko men-iy-a uy?*
 what=3PL say-FUT-FIN *uy?*
 ‘What do they say? *uy?*’
- (xxvi) *enleka uy?-ke-redo yuw-e*
 like.that jump.up-T/A-COND fall.down-FUT:FIN
 ‘(and) if (s/he) jumps up like that, (then) she will fall down’
- (xxvii) *ena cina? ne-ya cina? =ko meta yu?*
 that what this-FOC what=3PL say *yu?*
 ‘What’s that? This one they pronounce *yu?*’
- (xxviii) *enayte ne?e ne-ya-do a?e*
 then this this-FOC-EMPH *a?e*
 ‘Then this one here, this one is *a?e*’
- (xxix) *cina=ko mena-ye a?e*
 what=3PL say-FUT:FIN *a?e*
 ‘What they call this is *a?e*’
- (xxx) *ondo ne-ya-do o?*
 and this-FOC-EMPH *o?*
 ‘and this one is *o?*’
- (xxxi) *cina=bu meta o?*
 what=1PL.INC say *o?*
 ‘What we call this one is *o?*’
- (xxxii) *ena-do ne-ya-do mis-ete ule-tan-re*
 that-EMPH this-FOC-EMPH throat-ABL vomit-PROG-LOC
jantan-ko-re ka=bu men-iy-a
 different.places-PL-LOC NEG=1PL.INC say-FUT-FIN
 ‘That’s not what we call that one, this is when vomit comes out all over.’
- (xxxiii) *iya? o? ka=bu men-iy-a*
iya? o? NEG=1PL.INC say-FUT-FIN
 ‘*iya?* (we say) we don’t say *o?*’
- (xxxiv) *heya? cina?-ko-ta*
heya? what-PL-T/A:FIN
 ‘*heya?* is what they (say)’

- (xxxv) *ente ne-ya-do=ko meta?-a heyo?*
 then this-FOC-EMPH=3PL say-FIN heyo?
 ‘Then this one here they pronounce heyo?’
- (xxxvi) *cina?=ko meta?-[a] heyo?*
 what=3PL say-FIN heyo?
 ‘What do they say, heyo?’
- (xxxvii) *ena cana?p-te ne-ya-do i*
 that finish-ABL this-FOC-EMPH i
 ‘That next one, this is i.’
- (xxxviii) *cenaka=ko men-iy-a i*
 what=3PL say-3:FIN i
 ‘What they call this is i.’
- (xxxix) *ondo?o ne-ya-do cina=bu men-iy-a u*
 and this-FOC-EMPH what=1PL.INC say-FUT-FIN u
 ‘and this one here what we call it is u’
- (xl) *ne-te enayte nen ayer-te ko nen-ta-re*
 this-ABL then this come.first-ABL they? this-EMPH-LOC
 ‘after this then, starting from the beginning here’
- (xli) *ena ete nen wa a[?] uy? yu? a?e o? i u heyo? heya?*
 that begin:FUT:FIN this wā a[?] uy? yu? a?e o? i u heyo? heya?
 ‘that begins like this wā a uy yu a?e o? i u heyo? heya?’
- (xlii) *ne?en ne-ya-do-ko meta?-a b^howels b^howels*
 this:EMPH this-FOC-EMPH=3PL say-FIN vowels vowels
 ‘These ones here they call vowels.’
- (xliii) *b^howel english-te vowel citi=ko meta?-a hor-te-do...*
 bowel English-ABL vowel word=3PL say-FIN ho-ABL-EMPH
 ‘Bowel (we say) in Ho, from the English word vowel.’
- (xliv) *iya enlekan doburi citi-te=ko meta?-a*
 DISC like.that double/diphthongs letter-ABL=3PL say-FIN
 ‘Those ones like that they call doburi (diphthongs).’
- (xlv) *enayte abu na?-do nen-ta-ete=bu e?e? -ey-a*
 then we now-EMPH this-EMPH-ABL=1PL.NC start-FUT-FIN
 ‘Now then we will start with these here.’
- (xlvi) *ne-ya cina=ko mena?-a oŋ*
 this-FOC what=3PL say-FIN oŋ
 ‘This one is what they call oŋ.’
- (xlvii) *cena?=ko mena?-a oŋ*
 what=3PL say-FIN oŋ
 ‘What do they call it? oŋ’
- (xlviii) *cena?=ko men-ey-a goy?*
 what=PL say-FUT-FIN goy?
 ‘What they call (this) is goy?’

- (xlix) *cina?*=*ko* *men-iy-a* *ko*
 what=3PL say-FUT-FIN *ko*
 ‘What they call (this) is *ko*’
- (l) *ena-ya-ko* *meta* *wij*
 that-FOC=3PL say:FIN *wij*
 ‘That one they call *wij*.’
- (li) *kaji-pe* *sitje-ko* *bjn* *wij*
 say-2PL child-PL *bjn* *wij*
 ‘Say it children, *bjn*, *wij*.’
- (lii) *enete* *na?*-*ya-do*=*ko* *meta* *wij*
 then this!-FOC-EMPH=3PL say *wij*
 ‘This one here they call *wij*.’
- (liii) *cena*=*ko* *meta* *wij*
 what=3PL say *wij*
 ‘What they call it is *wij*.’
- (liv) *enate* *cana?**p-te* *cena*=*ko* *meta?*-*a* *wic*
 then finish-ABL what=3PL say-FIN *wic*
 ‘Then the next one is what they call *wic*.’
- (lv) *ena* *cana?**p-te* *nen-ta*=*bu* *nele-tan-a*
 that finish-ABL this-EMPH=1PL.INC see:3?-PROG-FIN
 ‘the next one we see’
- (lvi) *ne-ya-do*=*ko* *meta* *en*
 this-FOC-EMPH=3PL say:FIN *en*
 ‘This one they call *en*.’
- (lvii) *en* *kaji-pe*
en say-2PL
 ‘Say *en*.’
- (lviii) *enayte* *ne-ya-do* *daru* *od-ia-n-a*
 then this-FOC-EMPH tree cut.down-T/A-ITR-FIN
ka=*ko* *meta?*-*a*
 NEG=3PL say-FIN
 ‘Then this next one here, trees are cut down, don’t they say?’
- (lix) *daru* *nen-ta* *od-ia-n-a*
 tree this-FOC cut.down.by.person-T/A-ITR-FIN
- (lx) *rapud-iy-a-n-[a]*
 cut.down.by.nature-T/A-ITR-FIN
 ‘a tree is (/may be) cut down’
- (lxi) *nen* *lipi*=*ko* *meta?*-*a* *od*
 this letter=3PL say-FIN *od*
 ‘This letter they call *od*.’
- (lxii) *cina*=*ko* *meta?*-*a* *od*
 what=3PL say-FIN *od*
 ‘What do they call this, *od*?’

- (lxiii) *ena cana²p-te ne-ya-do jana iyu?-ye-n-dre*
 that finish-ABL this-FOC-EMPH side fall-T/A-ITR-LOC
 ‘The next one, this is when (something) falls on its side.’
- (lxiv) *citana-ete latar otere yu-yen-re te?e mente kasar-iyē*
 up-ABL down ground-LOC fall-T/A-ITR-LOC *te?e* like sound-FUT:FIN
 ‘When from up above it falls and hits the ground, it sounds like *te?e*.’
- (lxv) *enleka uru?-i-pe ondo kaji-pe*
 like.that think-FUT-2PL and speak-FUT-2PL
 ‘Think about it like that and say it.’
- (lxvi) *en cina=ko kaji-ye*
 that what=3PL say-3:FIN
 ‘That is what they call it.’
- (lxvii) *enayte ho: ungur-e-n-re cilike*
 then person bend.forward-T/A-ITR-LOC when
naṅnuṅ ba:n-tan-a ka=ko men-iy-a
 back.and.forth AUX-PROG-FIN NEG=3PL say-3-FIN
 ‘When a person bends forward, they rock back and forth so (*naṅnuṅ*)
 don’t they say it?’
- (lxviii) *ne-ya-do cina=ko kaji-ye nuṅ*
 this-FOC-EMPH what=3PL speak-3:FIN *nuṅ*
 ‘This one is what they call *nuṅ*.’
- (lxix) *ayum-ke-ḍ-a-pe sabēn*
 hear-T/A-TR-FIN-2PL all
 ‘Did you all hear (that)?’
- (lxx) *enayte nen-ta-do cilike nel-o?o-tan-a*
 then this-EMPH-EMPH how see-PASS/ITR-PROG-FIN
 ‘Then this one here how does it look?’
- (lxxi) *nen-ta nen-do cikana em-aka-n-a*
 this-EMPH this-EMPH what give-T/A-ITR-FIN
 ‘This one here what’s it give (the appearance of)?’
- (lxxii) *nemenēṅ leka da?a pere-yeka-n-a leka*
 up.to.this like water be.filled-T/A-ITR-FIN like
nel-o?o-tan-a
 see-PASS/ITR-PROG-FIN
 ‘It looks like water could be filled up to this like that.’
- (lxxiii) *esu rāsa-te=bu urum-ke-ḍ-a*
 very happy-ABL=1PL.INC know/recognize-T/A-TR-FIN
 ‘We very happily recognize it.’
- (lxxiv) *na?a-do cina=bu kaji-e menta-do*
 now-EMPH what=1PL.INC say-3:FIN like-EMPH
 ‘Now what is it that we call this one like.’

- (lxxv) *cina?a=ko kaji-e da?a*
 what=3PL say-FIN:FUT *da?a* (water)
 ‘What they call it is *da?a*.’
- (lxxvi) *abu-e? desum-do go:lge mena?aca gulge*
 WE.PL.INC-GEN country-EMPH round COP.NFIN round[ness]
mena-ke-q-a
 say-T/A-TR-FIN
 ‘Our country is round, (they) have told (us) it is round.’
- (lxxvii) *ne-ya-do enleka abu desum-rea? desum leka*
 this-FOC-EMPH like.that we.PL.INC country-GEN country like
nel-o?-tan leka ot
 see-PASS/ITR-PROG like *ot*
 ‘This one here, we, it looks like our country, like our world.’
- (lxxviii) *cina=bu kaji-e ot*
 what=1PL.INC say-FUT:FIN *ot*
 ‘What we call (this) is *ot*.’ (Note: *desum* = world)
- (lxxix) *ne-ya-do abu cina=bu kaji-e*
 this-FOC-EMPH we.PL.INC what=1PL.INC say-FUT:FIN
ne-ya-do=ko meta homo
 this-FOC-EMPH=3PL say:FIN *homo*
 ‘This one here, what we call this one here, they call it *homo*.’
- (lxxx) *ado enleka-ge ne-ya-do bu cina=ko kaji-e bu*
 it like.that-FOC this-FOC-EMPH *bu* what=3PL say-3:FIN *bu*
 ‘Then that one it is like that, this one is *bu*, we they call this is *bu*.’
- (lxxxix) *ne-ya-do cilike nel-o?-tan-a marda*
 this-FOC-EMPH how see-PASS/ITR-PROG-FIN come.now
miyēḍ sakam-te bai-o?-leka pu leka
 one.NONHUMN leaf-ABL create-PASS/ITR-PP leaf.cup like
nel-o?-tan-a kalgi leka
 see-PASS/ITR-PROG-FIN leaf.plate like
 ‘This one here how does it look, come now, it looks like a leaf cup made from a single leaf, (or) like a leaf plate?’
- (lxxxii) *ne-ya-do=ko meta?-a pu*
 this-FOC-EMPH=3PL say-FIN *pu*
 ‘This one they call *pu*.’
- (lxxxiii) *diyey=ko nui-ti-a? leka kaji-o?-tan-a*
 country.liquor=3PL drink:3-ABL-GEN like say-PASS/ITR-PROG-FIN
enleka nel-o?-tan-ti-a? esu bania-te leka-te
 like.that see-PASS/ITR-PROG-ABL-GEN very beautiful-ABL like-ABL
 ‘It is said like what they drink country liquor from, it looks like that, quite beautiful[ly].’

- (lxxxiv) *esu anya-te yirgel eto-ya marci na?a=bu*
 very quick-ABL be.aware learn-FUT:FIN let's.try now=1PL.INC
kaji-e pu
 say-FUT-FIN *pu*
 'Very quickly (you) will learn it, now let's have a go, let's say *pu*.'
- (lxxxv) *ente ne-ya-do=ko meta?a hoyo*
 then this-FOC-EMPH=3PL say-FIN *hoyo*
 'Then this one here they call *hoyo*.'
- (lxxxvi) *ente=ko cina=ko meta?a hoyo*
 then=3PL what=3PL say-FIN *hoyo*
 'That one, they, they call it *hoyo*'
- (lxxxvii) *ente enleka-ge ne-ya-do holo holo*
 then like.that=FOC this-FOC-EMPH *holo holo*
 'Then this is one like *holo holo*.'
- (lxxxviii) *cina?=ko meta?a holo holo*
 what=3PL say-FIN *holo holo*
 'What they call this one is *holo, holo*.'
- (lxxxix) *ondo?o ne-ya-do-ko meta?a ho?a*
 and this-FOC-EMPH=3PL say-FIN *ho?a*
 'and they call this one *ho?a*'
- (xc) *cina=bu kaji-e ne-ya-do*
 what=1PL.INC say-FIN this-FOC-EMPH
 'what we call this one is'
- (xci) *ente ne-ya-do meta?a=bu kaji-e har*
 then this-FOC-EMPH say-FIN=1PL.INC say-FIN *har*
 'Then this one is what that is called, we call it *har*.'
- (xcii) *cina?=ko meta?a har*
 what=3PL say-FIN *har*
 'What they call this one is *har*.'
- (xciii) *enayte nē?e ne-ya-do su*
 then this:EMPH this-FOC-EMPH *su*
 'Then this one is *su*.'
- (xciv) *cina=bu kaji-e su*
 what=1PL.INC say-FIN *su*
 'What we call this is *su*.'
- (xcv) *ondo?o ne-ya-do esu anya-te toraŋ=bu*
 and this-FOC-EMPH very quick-ABL perhaps=1PL.INC
nel urum-ey-a naye-leka nel-o?o-tan-te-a?
 see+recognize-FUT-FIN plough-PP see-PASS/ITR-PROG-ABL-GEN
cina=bu kaji-e si
 what=1PL say-FIN *si*
 'and this one perhaps very quickly (you) will see and recognize as
 because it looks like a plough, what we call this is *si*.'

- (xcvi) *ne-ya-do=bu si-ti-a*
 this-FOC-EMPH=1PL.INC plough-ABL//T/A-FIN
 ‘This is what we plough with.’
- (xcvii) *enayte ne-ya-do maraŋ kero?o mena?-a kero mena?-a*
 then this-FOC-EMPH big case COP-FIN case COP-FIN
 ‘Then this one here is the upper case, the case.’
- (xcviii) *ne-ya-do=ko hurij hurij-ti-a maraŋ-ti-a-do*
 this-FOC-EMPH=3PL small small-ABL//T/A-FIN big-ABL//T/A-FIN-EMPH
maraj-leka-ge ol-aka-n-a
 big-PP-FOC write-PLUP-ITR-FIN
 ‘These ones here, they, small and large, like this big, have been written’
- (xcix) *enleka kitep-ko cap-aka-n-ti-a? bano-leka-te-a*
 like.that book-PL print-PLUP-ITR-ABL-GEN NEG.COP-PP-ABL-GEN
 ‘because there aren’t any printed books (available to show you).’
- (c) *ka=lij udub daie-ten-e*
 NEG=1DL.EX [make.]understand CAP-PROG-FIN
 ‘I can’t really make you understand.’
- (ci) *uru?u-ye=lij ape esu anyate en*
 think-T/A=1DL you very quickly that
kitep-ko=lij udub-e-pi-a
 book-PL=1DL make.understand-FUT-2PL-FIN
 ‘I think that I could make you understand (how to read) those books very quickly.’
- (cii) *gapa gapa cimite paŋa-o? hob-a-n-a*
 tomorrow tomorrow when read-PASS/ITR AUX-T/A-ITR-FIN
 ‘Tomorrow, tomorrow when the reading continues’
- (ciii) *en gapa paŋa-o? hoba-i-redo kitep-ko*
 that tomorrow read-PASS/ITR AUX-FUT-COND book-PL
cilike udub-kete en na? saben
 how make.understand-DS that now all
kaji=lij udub-e-pi-a
 say=1DL make.understand-FUT-2PL-FIN
 ‘Tomorrow reading will continue and I will teach you all to understand how (to read) these books and how to say.’
- (civ) *tisij=e? paŋ-o?-do nen-ta-re=bu*
 today=EMPH?/3 read-PASS/ITR-EMPH this-EMPH-LOC=1PL.INC
mucire-tan-a
 stop-PROG-FIN
 ‘Today we are going to stop here now.’
- (cv) *eya marena-do tisij nen-ta-re mucire-tan-a*
 yes for.now-EMPH today this-EMPH-LOC stop-PROG-FIN
 ‘Yes, today (let’s) stop here for now.’



FIGURE 5.1 MR K.C. NAIK BIRULI, HO ORATOR AND SCHOLAR, 2005

As told by Mr. K.C Naik Biruli (see Figure 5.1) 12 September 2005, Bhubaneswar. Recorded, translated, and annotated by Gregory D.S. Anderson and K. David Harrison.

8 HO ORTHOGRAPHY

Oral use of the Ho language is vigorous in all speech domains and age groups. The Ho are typically bilingual or trilingual, and literate in languages of wider communication, for example, Oriya, Hindi, English. Use of Ho writing remains limited. It has been written using multiple scripts, including Devanagari and Oriya, but these are considered by the community as nonideal, and the use of an exclusively Ho script is to be preferred. Following the rationale of other writing systems of the Indian subcontinent, the Ho orthography (called Warang Chiti) was designed by the scholar Lako Bodra to appear maximally distinctive from all other writing systems in nearly all its graphemes.

During two field trips to India, Anderson and Harrison recorded the use of the orthography and narratives, including the one presented above, about the meaning and cultural importance of the writing system. We were able to collect samples of Ho handwriting from individuals (Figures 5.2 and 5.4), as well as one printed book, an alphabet primer (see Figure 5.3). The Ho orthography is rapidly gaining a wider user base, despite lack of representation of the language in official media and in state schools. A proposal for inclusion of Ho in the Unicode standard is currently being prepared (cf. Harrison and Anderson 2007).



FIGURE 5.2 THE HO ORTHOGRAPHY, A HANDWRITTEN VERSION FROM 2005, INCLUDING NUMBER SYMBOLS (LOWER ROW), BUT EXCLUDING LOWER-CASE LETTERS



FIGURE 5.3 A SAMPLE PAGE FROM THE HO ALPHABET PRIMER BOOK (CHLEC 2002) SHOWING THE USE OF BOTH UPPER- AND LOWER-CASE FORMS

BIBLIOGRAPHY

- Adhikary, Ashim Kumar (1984) *Society and World View of the Birhor: A Nomadic Hunting and Gathering Community of Orissa*, Calcutta: Anthropological Survey of India.
- Anderson, Gregory D.S. (2001) 'A new classification of Munda: Evidence from comparative verb morphology', *Indian Linguistics*, 62:27–42.
- (2002) 'Case marked clausal subordination in Burushaski complex sentence structure', *Studies in Language*, 26(3):547–571.
- (2003) 'Yeniseic languages from a Siberian areal perspective', in E. Vajda and G.D.S. Anderson (eds) *Studia Yeniseica: Typological Studies on Yeniseic*, special issue of *Sprachtypologie und Universalienforschung*, 56(1/2):12–39.
- (2004) 'The languages of Central Siberia: Introduction and overview', in E. Vajda (ed.) *Languages and Prehistory of Central Siberia*, Amsterdam: John Benjamins, 1–119.
- (2006) *Auxiliary Verb Constructions*, Oxford: Oxford University Press.
- (2007) *The Munda Verb. Typological Perspectives*, Berlin: Mouton de Gruyter.
- Anderson, Gregory D.S. and John P. Boyle (2002) 'Switch-reference in South Munda', in Marlys A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 39–54.
- Anderson, Gregory D.S. and Norman H. Zide (2001) 'Recent advances in the reconstruction of the proto-Munda verb', in Laurel Brinton (ed.) *Historical Linguistics 1999*, Amsterdam: Benjamins, 13–30.
- (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic Constraint', in Marlys A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Bahl, Krishna Chandra (1962) *Korwa Lexicon*, Mimeograph, Chicago: University of Chicago.
- Barker, P.R. (1953(?)) *The Phonemes of Korowa*, Mimeograph, Seattle, WA.
- Bhattacharya, Sudhibhushan (1975) *Studies in Comparative Munda Linguistics*, Simla: Indian Institute of Advanced Study.
- Bhol, Lata and Paramananda Patel (1992) *Ho* (Text in Oriya with roman script for Ho), Bhubaneswar: Academy of Tribal Dialects and Cultures.
- Bodding, P.O. (1929–1936) *A Santali Dictionary*, 5 Vols, Oslo: J. Dybwad.
- Bright, William (ed.) (1992) *International Encyclopedia of Linguistics*, New York: Oxford University Press.
- Burrows, Lionel (1915) *Ho Grammar*, Calcutta: Catholic Orphan Press.
- CHLEC (2002) *A Ho Alphabet Primer* (Text in Ho script). Chaibasa: Ho Language Society.
- Dasgupta, Samir (1994) *Birjia: Society and Culture*, Calcutta: Firma KLM.
- Dasgupta, S.B. (1978) *Birjhia: A Section of the Asurs of Chotanagpur*, Calcutta: K.P. Bagchi.
- Deeney, J. (1975) *Ho Grammar*, Chaibasa: Xavier Ho Publications.
- (1978a) *Ho-English Dictionary*, Chaibasa: Ho Xavier Publication.
- (1978b) 'Comparison of the Munda and Ho languages', in P. Ponette (ed.) *The Munda World: Hoffmann Commemoration Volume*, Ranchi: Catholic Press, 44–52.
- Dimock, Edward Cameron, Jr (1957) 'Notes on semi-vowel alternation in the Bengali verb', *Indian Linguistics*, 17:172–177.
- Donegan, P.J. (1993) 'Rhythm and vocalic drift in Munda and Mon-Khmer', *Linguistics in the Tibeto-Burman Area* 16(1):1–43.
- Dryer, Matthew (1986) 'Primary objects, secondary objects, and antitativity', *Language*, 62:808–845.

- Evans, N. and T. Osada (2005) 'Mundari: The myth of a language without word classes', *Linguistic Typology*, 9(3):351–390.
- Ghosh, Arun (this volume, Chapter 2) 'Santali'.
- Gordon, R. (ed.) (2005) *Ethnologue*, 15th edition, Dallas: SIL International.
- Grierson, G.A. (ed.) (1906) *Linguistic Survey of India, Vol IV: Munda and Dravidian*, Calcutta: Office of the Superintendent of Government Printing.
- Harrison, K. David and Gregory D.S. Anderson (2007) 'Review of proposal for encoding Warang Chiti (Ho orthography) in unicode', unpublished technical report, available online at www.livingtongues.org/moremunda, accessed on 1 November 2007.
- Hoffmann, John (1930–1978) *Encyclopaedia Mundarica*, 16 vols, reprinted in 1990, New Delhi: Gyan Publications.
- Kobayashi, Masato and Ganesh Murmu (this volume, Chapter 4) 'Kera? Mundari'.
- Kobayashi, Masato, Ganesh Murmu and Toshiki Osada (2003) 'Report on a preliminary survey of the dialects of Kherwarian languages', *Journal of Asian and African Studies*, 66:331–364.
- Konow, Sten (1904) 'Munda and Dravidian', *Indian Antiquary* 33:121–125.
- Leuva, K.K. (1963) *The Asur: A Study of Primitive Iron-Smelters*, New Delhi: Bharatiya Adimjati Sevak Sangh.
- Minegishi, M. (1990) 'Santali-English-Japanese wordlist – a preliminary report', *Journal of Asian and African Studies*, 39:69–84.
- Munda, Ram Dayal (1968) 'Proto-Kherwarian phonemic system', MA Thesis (unpublished), Department of Linguistics, University of Chicago.
- Nagaraja, K.S. (1999) *Korku Grammar with Texts*, Tokyo: TUFUS.
- Narayan, Abaninder (1990) *The Korwa Tribe: Their Society and Economics*, Delhi: Amar Prakashan.
- Neukom, L. (2000) 'Argument marking in Santali', *Mon-Khmer Studies*, 30:95–113.
- Osada Toshiki (1991a) 'Father Ponette's field note on Turi', *Journal of Asian and African Studies*, 42:175–189.
- (1991b) 'Notes on linguistic convergence in the Chotanagpur area', in Sanjay Bosu-Mullick (ed.) *Cultural Chotanagpur: Unity and Diversity*, Delhi: Uppal Publications, 99–119.
- (1992) *A Reference Grammar of Mundari*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- (1993) 'Field notes on Birhor', *A Computer-Assisted Study of South-Asian Languages*, 30–40.
- (1996) 'Notes on the Proto-Kherwarian vowel system', *Indo-Iranian Journal*, 39:245–258.
- (2001) 'Personal pronouns and related phenomena in South Asian linguistic area: Convergent features or convergence-resisting features?', *Yearbook of South Asian languages and Linguistics 2001*, New Delhi: Sage, 269–287.
- (this volume, Chapter 3) 'Mundari'.
- Pinnow, Heinz-Jürgen (1959) *Versuch einer historischen Lautlehre der Kharia-sprache*, Wiesbaden: Otto Harrassowitz.
- (1965) 'Personal pronouns in the Austroasiatic languages: A historical study', *Lingua*, 14:3–42.
- (1966) 'A comparative study of the verb in the Munda languages', in Norman Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton, 96–193.
- (1972) 'Schrift und Sprache Lako Bodras', *Anthropos*, 67:822–857.
- Prasad, Narmadeshwar (1961) *Land and People of Tribal Bihar*, Ranchi: Bihar Tribal Research Institute.

- Ramaswami, N. (1992) *Bhumij Grammar*, Mysore: Central Institute of Indian Languages.
- Roy, Sarat Chandra (1925) *The Birhors: A Little-known Jungle Tribe of Chota Nagpur*, Ranchi: Man in India Office.
- Sahu, Chaturbhuj (1995) *Birhor Tribe: Dimensions of Development*, New Delhi: Sarup and Sons.
- Samal, Smrutirekha and Paramananda Patel (1997) *Bhumijia* (in Oriya with roman script for Bhumijia), Bhubaneswar: Academy of Tribal Dialects and Cultures.
- Sandhwar, A.N. (1978) *The Korwa of Palamau: A Study of their Society and Economy*, Wien: Institute für Völkerkunde der Universität.
- Singh, Bageshwar and Ajit K. Danda (1986) *The Kodaku of Surguja*, Calcutta: Anthropological Survey of India.
- Stampe, David (1965–1966) ‘Recent works in Munda linguistics I–IV’, *International Journal of American Linguistics*, 31:332–341, 32:74–88, 164–168, 390–397.
- Zide, Norman H. (1966) ‘Korku low tone and proto-Korku-Kherwarian vowel system’, in Norman Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, The Hague: Mouton, 214–229.
- (1969) ‘Munda and non-Munda Austroasiatic languages’, in Thomas Sebeok (ed.) *Current Trends in Linguistics*, The Hague: Mouton, 5:411–430.
- (1991) ‘Lexicography of other languages of the Indian subcontinent: The Munda languages’, in Franz Josef Hausmann, Oskar Reichmann, Herbert Ernet Wiegard and Ladislav Zgusta (eds) *An International Encyclopedia of Lexicography*, Berlin and New York: Walter de Gruyter, 3:2533–2547.
- (1996) ‘Scripts for Munda languages’ in Peter Daniel and William Bright (eds) *The World’s Writing Systems*, New York: Oxford University Press, 612–618.
- (2000) ‘Three Munda Scripts’, *Linguistics of the Tibeto-Burman Area*, 22:199–232.
- (this volume, Chapter 6) ‘Korku’.
- Zide, Norman and Ram Dayal Munda (1966) ‘Notes on Proto-Kherwarian vocalism’, Abstract appeared in Stampe (1966) No. 732.

KORKU*

Norman H. Zide

1 INTRODUCTION

Korku is a North Munda language. It constitutes the western branch of North Munda, the Kherwarian languages making up the eastern branch. Korku is thus the westernmost language of the Austroasiatic phylum. There are some interesting dialect chain features (e.g. involving the complex demonstrative stems characteristic of Munda) extending across the whole North Munda area. Our knowledge of the minor Kherwarian languages in western Jharkhand and eastern Madhya Pradesh or Chhattisgarh is limited.

This language, also called Kurku in the western regions of its range, is called Muwasi (Mowasi, Mawasi) by the speakers of its eastern dialect, who reject the name ‘Korku’ (see other names in *Ethnologue*). The name comes from *koro-ku*, *koro* ‘person, member of the Korku community’ (Santali *hoꝛo*, Ho *hoꝛ*, etc.), and *-ku*, animate plural. Most of our knowledge of the language comes from the western or Melghat region, and the Melghat dialect (Drake 1903; Girard 1965; Nagaraja 1999; Zide 1960; also Nagaraja 1989). What little we know of Mawasi comes from the old *LSI* (Grierson 1906).

Korku is spoken in northeastern Maharashtra (most of the speakers residing in Amraoti (Amravati) District), and in south central Madhya Pradesh (East Nimar, Betul, and Hoshangabad Districts). The SIL do not mention the area of Chhindwara District just east of Hoshangabad where there must still be Mawasi speakers.

Ethnologue (2000) gives the number of Korkus as 478,000, most of them presumably speaking the (Kurku, Mawasi – *Ethnologue* gives other ‘alternate names’ as well) language. They claim bilingualism in Hindi and Marathi is low (no percentage is given). I suspect ‘low’ is an underestimate, perhaps a considerable underestimate (depending on what degree of competence in Hindi is understood). Literacy in Korku (5% maximum according to *Ethnologue*) probably is fairly low, but literacy in Hindi and/or Marathi (which they do not mention) would be higher, certainly greater than 5%.

The Korku, certainly the men, are almost all bi- or multilingual, speaking the regional ‘varieties of Hindi’ in Madhya Pradesh, and Marathi (and dialectal Hindi) in Maharashtra. Korku has borrowed heavily from neighboring Indo-Aryan languages, particularly in lexicon. There are a few words apparently borrowed (directly?) from some Dravidian language (e.g. *bo* ‘let’s go’). N. Zide observed considerable change in syntax, for example, the new Indo-Aryan-like relative clause structures preferred by some younger speakers.

Inter-dialect intelligibility, they say is high (76 to 97%), but it is not clear that the Mawasi dialect – the most divergent dialect in the extreme east of the Korku area – was included in the inter-intelligibility tests. It is hard, without much data, to judge the kind and scale of difference between Korku now from Melghat, and the Betul–Hoshangabad and Mawasi dialects. Very broad observations of a few Lahi (Hoshangabad) speakers showed that, for instance, they had not only lost the dual, but had no memory of it.

Korku uses the Devanagari script (used for Hindi and the Marathi form of it) and has not, like some of the Kherwarian languages, devised a script of its own. Very little has been written, and I have learned very little about the small amount of earlier-published materials in Korku. Earlier scholars – missionary scholars among them – recorded folk songs and tales in the Melghat region, these preserved more by women than men (Girard, unpublished texts). There was little printed material in Korku: some Bible translations, Christian hymns put into Korku by foreign missionaries, etc. The Madhya Pradesh government prepared simple language-teaching materials. With increasing population and an increased number of literate (mostly in Hindi) speakers, and some greater general recognition of ‘tribal’ languages in India, particularly after the creation of the state of Jharkhand and the promotion of the local Munda languages there, one might expect an increased use of Korku (rather than Hindi or Marathi) in written materials of all sorts. (There are still Christian missionaries in the Melghat area. The earlier missionaries were British, Canadian, and American. The current missionaries are Indian and are from the northeast of India.)

2 PHONOLOGY

2.1 Vowel inventory

Korku possesses a basic five-vowel system.

- (1) *i* *u*
 e *o*
 a

There is considerable intra-dialect (and even intra-idiolect) variation in the phonetic forms of vowel sequences; that is, there are glides, more or less prominent semivowels, etc. There is some phonetic description and detail in Nagaraja (1999), and a few (questionable?) forms in their lexical materials showing contrasting, *Cya* and *Cia* vs. *Ciya*. But note for instance (Zide field notes) *baon*; (Girard), *bavon*; (Nagaraja 1999), and *bawan* ‘wife’s younger sister’; (Zide) *tiälteä*, (Nagaraja) *tyä*, (Girard) *tiyã* ‘husband’s older sister’; (Zide) *miã?* (Nagaraja) *mya*, (Girard) *miya?* ‘one’.

Long vowels, that is, sequences of the same vowel, interpreted by Zide as belonging to separate syllables, occur frequently, almost always in borrowed forms. Final vowels in monosyllabic nouns are automatically lengthened, for example, *jo* {*joo*} ‘fruit’; *ti* {*tii*} ‘hand’; thus Korku appears to conform to the Munda minimal bi-moraic word constraint it inherited from Proto-Munda (and probably Proto-Austroasiatic, see Anderson and Zide [2002] for details).

An *-o* is usually added to consonant-final monosyllabic loan stems in Korku. Examples of monosyllabic nouns in borrowed forms include: Hindi *sath* Korku *saato* ‘with’, Hindi *koṭ* (< English) ‘coat’, Korku *kooṭo*. Zide reports a few examples of syllabic nasal, for example, *jilygod* (Nagaraja: *jilngot*, but Girard *jiriygod* ‘worm species’).

2.2 Suprasegmental phenomena (tone, register)

Aspiration in Hindi gives low tone in Korku: Hindi *pehlaa* (the *a* phonetically a lower *e*), ‘first, before’ Korku *peèla*.

There is a small number of old Korku morphemes of CVV'(C) shape, for example, *doò*, 'to put, place' *koòŋ*, 'to call', but there are many morphemically complex forms of that shape, for example, *qeèn* 'there' (< *qi-* 'that' + *-èn* 'locative') and many reduplicated verbs with initial V, for example, *ira?*, 'to return (intransitive)', *ìra?* 'to return (transitive)'. That is, the reduplicated form can be said to be made by infixing the (necessarily low-toned) verb-stem initial or, by another interpretation, prefixing a non-low initial syllable to a stem with its initial syllable lowered. Nagaraja does not usually distinguish these. Tone was recorded (by Zide, but not by Drake, Girard, or Nagaraja) consisting of a marked low tone contrasting with an unmarked high, for example, *bulù* 'thigh', *lulu* (reduplicated stem) 'to draw water'; non-low before glottal stop slightly rising, *ura?* 'house'; low, before glottal a fall-rise *sasà?* (reduplicated stem) 'to bring'. Low tone in allegro speech has as its domain a whole noun phrase, that is, anything in the phrase after a low tone is low: 'my small black goat' *ijnà? sani kēḍe siṛi* (*ijn* 'I' -à? genitive, *sani* 'small', *kēḍe* 'black', *siṛi* '(male) goat' comes out as *ijn-à sà-nì kē-ḍe si-ṛi*.

Only non-initial syllables exhibit tone distinction. Tone is correlated with aspiration – any non-initial syllable with syllable-initial aspirate is necessarily low-toned, for example, *ukhù* 'to hide', **ukhu* is impossible. But note the exception mentioned above for the reduplicated and automatically lowered final syllable of dissyllabic verb-stems (due to the stem-final glottalized consonant). The reduplicated form of *kab-* is *kakàb*, not **kakhàb*, that is, the preglottalization blocks the aspiration. Thus a more economical transcription would be: *ukHu*, with H used to mark low tone aspiration. Where a lost syllable is contracted (i.e. VV' > V) initially, since initial syllables cannot be 'low', its low tone shifts to the following syllable. For example, the verb-stem *koòŋ-* is reduced to *koŋ-* when preceding a consonant-initial syllable, but the aspiration-tone-lowering affects the following syllable, for example, *koòŋ-kij* 'call them-2' > *koŋ-khij* which, as suggested above, could be morphophonemically written as *koHŋ-kij*. The only aspirated stops that occur medially are voiceless stops. Where reduplication of a voiced initial aspirate occurs, the aspiration in the reduplicated second syllable is absent, but the low tone expected from a non-initial aspirate – that is, **gha-ghal* – remains in, for example, *ghal-* 'to show', reduplicated *ghagàl*. With some speakers the medial *g* is slightly aspirated.

Tone has a limited functional load and has been lost in some speakers as Zide observed, and Nagaraja looked but didn't find it at all. But where a name ends in a vowel, the only morphological indicator of genitivity is the low tone, for example, *dadù ura?* (from *dadu-à? ura?* 'Dadu's house'.

2.3 Consonant inventory

The Korku consonants are those standard for the area along with the typical (more fully North than South) Munda pre-glottalized final stops. Bhattacharya (personal communication c.1956) entertained the idea that Korku had, like Dravidian, three series of coronal stops, roughly: dental, alveolar, and retroflex, but finally decided that there were only two. The early Proto-Munda system must have had two terms for coronal stops – dental *t* and retroflex *ḍ* (this arrangement surviving most clearly in Sora). It looks as if such a two-term system persisted into earlier Korku, before the present system was firmly established, with heavily borrowed lexicon resulting in an Indo-Aryan-like 'pattern-filling' dental *d* and *t*, and retroflex *ḍ* and *ṭ*. The later phonemic alignment of the phonetic data was perhaps different in different dialects

and idiolects, the aligning having gone on – back and forth – over a period of time. The different phonemic transcriptions (in some words) of, for example, initial *t* in Nagaraja and Zide (Nagaraja *t*, Zide *t*) probably reflect different judgments on how to classify allophones of alveolar *t* (in, for example, *ti*, *tɪ* ‘hand’). Girard transcribed almost no forms with *t* and *th*. If there were fuller lexical data more words would be listed where alternative forms are now found – forms with both *t* and *t̪*, and *d* and *d̪*.

The current consonantal inventory for Melghat Korku as described by Zide (1960) is offered in (2).

(2)	<i>b</i>	<i>bh</i>	<i>d</i>	<i>dh</i>	<i>d̪</i>	<i>d̪h</i>	<i>j</i>	<i>jh</i>	<i>g</i>	<i>gh</i>	
	<i>p</i>	<i>ph</i>	<i>t</i>	<i>th</i>	<i>t̪</i>	<i>t̪h</i>	<i>c</i>	<i>[ch]</i>	<i>k</i>	<i>kh</i>	<i>ʔ</i>
	<i>m</i>	<i>mh</i>	<i>n</i>				<i>ɲ</i>		<i>ŋ</i>		
	<i>s</i>									<i>h</i>	
	<i>r</i>	<i>l</i>	<i>ɽ</i>	<i>y</i>							

In most older speakers of the dialects studied, the many loanwords with final stops in the source language are given the final vowel *o*, when borrowed into Korku and their final syllables lengthened, either by doubling their syllable-final consonant or lengthening that syllable vowel, for example, *kuluppo* (but Girard, Nagaraja *kulupo*) ‘lock’, dial. Hindi *kulup[h]*, *cooʔo*, Hindi *coʔ*, ‘wound, injury’, *koʔo* ‘coat, jacket’.

Nagaraja does not find *ʔ* in many forms, stems, and affixes, where Zide and Girard record them, for example, Girard/Zide *-teʔ* ~ *-t̪eʔ*, Nagaraja *-te* inalienable possession; Nagaraja *ra* ‘roar, break’, Girard/Zide *raʔ*- ‘roar, crow (of a cock), cry (of an animal)’, Nagaraja *ura* ‘house’, Girard/Zide *uraʔ*, Nagaraja *ira* ‘to return’ (intransitive), Girard/Zide *iraʔ*, etc.

2.4 Syllable structure and phonotactics

There are no initial or final consonant clusters in Korku in the older language. Initial clusters in loanwords were usually reduced from the presumed source words in various ways depending on the consonants involved (e.g. Brahman, *bra(h)man* > *bamman*/*bammàn*, or *st̪eshan* (the Hindi dialect source form may not have been much different from the Korku) > *tesan* ‘station’, etc.

Nagaraja treats the semivowels as consonants so that he analyzes more and longer consonant cluster sequences. I treat them as vowels, the CV₁ in, for example, *tiaʔ*- being a short syllable, the lightest – least weighted – syllable type in Korku. Nagaraja gives a long list of examples of sequences of consonants. I consider Nagaraja’s syllable-final *y* to be the vowel *-e*, and *i-* and *u-* before V (usually *a*) as vowels, not consonants, and come up with a different and shorter list. See also my (1960) comments on the short vocalic link that he takes as ‘zero’ – no vowel (e.g. *koc[e]re* ‘eggshell’), and that I take as a minimal vowel. Nagaraja lists as units (‘words?’) forms that I would analyze differently, that is, as different kinds of intra-word structures: noun phrases, compounds with ‘plus juncture’, etc. The C₁ C₂ sequences that we both would agree on include a number of borrowed forms, for example, *mas̪tar* ‘school teacher’. Nagaraja in his lists does not distinguish C₁-C₂ sequences across morpheme boundary from morpheme-internal consonant sequences.

2.5 Intonation/stress

Stress is largely automatic, determined by the syllable weight (closed syllables, those with final consonant, are heavier than open syllables) and position, final position syllables having more weight than initial ones. The earlier mention of longer – that is, half-long or longer than simple single stops, but less longer than geminate stops (for example, *ti(tt)ipi* above), can be considered to reflect phonemically stressed initial syllables. Initial syllables of reduplicated verb forms perhaps generally get more (not necessarily phonemically) stressed, but they are – phonemically – stressed where the syllable weight would not otherwise give them the requisite stress. Medial CV syllables are short, particularly where the following syllable begins with a continuant. *kath(ə)lâ?* ‘armpit’. One can treat – as Nagaraja does – these as reduced (or nonexistent) zero vowels. The analysis of medial consonant clusters obviously depends on such a decision.

Various juncture phenomena are noted, in connection with various types of compounding: where two (possible) words are not considered separate words, but are written with a plus juncture, for example, with commonly paired words (pair of nouns and verb infinitives). We distinguish the following types (these distinctions apply widely, and not just to Korku): the first type (type-I) where both Word₁ and Word₂ can occur independently, but in some cases, for example, *ayom-ba* ‘mother-father’ the forms in the compound are not the regular free forms in which *mae* is usually for ‘mother’ and *aba* for ‘father’. In type-II Word₂ does not occur independently and does not fit with types III and IV. It is often an archaic form. Type-III is onomatopoeic/sound-symbolic or belonging to the domain of forms often called expressive formation, for example, *qoē+qoē* ‘dragonfly’, *rep+rep* ‘to whimper (of a baby)’, *kusu+musu* (A type-III form, since *kusu* does not occur independently) ‘whisper’. Lastly, there are type-IV ‘echo’ forms like *sita+mita* ‘dogs and the like’. Echo forms are derived by adding an initial *m-* to Word₂ words beginning with V, or replacing an initial consonant with *m-*.

Practically nothing is known about intonation in Korku and this remains a subject for future research.

2.6 Morphophonology

The (pre-)glottalized consonants – weakly voiced ‘*b*’, ‘*q*’, ‘*j*’ – are a salient and distinctive characteristic of (almost all) the Munda languages. These occur word-finally, syllable-finally, and word-medially in open juncture. Pre-glottalized ‘*g*’ occurs in some Munda languages, but not in Korku; but the morphophonemic alternation of ? with -*g*- (found elsewhere in Munda with ‘*g* and *g*-’) occurs: the -*g*- occurs before V in close, word-internal juncture, for example, *ura?*, ‘house’, -*èn*, locative, *uragèn* ‘in the house’. Note also that stem-final ? behaves differently in its morphophonemics from affix-final ?. These are the only permitted noncontinuant (i.e. nasals, and *l*, *r*, and *ʃ*) final stops in the older language.

Monosyllabic verb-stems with (underlying) final *b*, *q* or *j* or ? when reduplicated in their final (second) syllable become low-toned, for example, *kab-* ‘to bite, hold in the mouth’, *bid-* ‘to sow seed’, reduplicated *kakàb* (not **kakhàb*), *bibid*. For polysyllabic verb-stems, the morphophonemic rules are similar. The glottalization, wherever ‘basically located’ and whatever its domain, has the same lowering effect as an aspirated initial would. *s* functions as an aspirate, a completely regular aspirated

equivalent to (unaspirated) *c* in some dialects, so that the reduplication of *sum-* would be *cusùn*; *sod-*, *cosòd* (see Ramsay). There is a reduction of aspirates rule (only one permitted, something akin to Grassmann's Law in Indo-European studies). With voiceless aspirates (e.g. *kh*, *s*) the aspirate occurs medially (*cusùn*, *kakhàd*); with voiced aspirates the aspirate occurs initially (*ghagàl*). The reduction-of-aspirates rule occurred in some forms where only cognate forms elsewhere show an original aspirate, that is, *s* (example: *cakhàn* 'firewood', Santali *sahan*). Note also Korku *samì* 'small', where the plural is marked with the $-(V)p-$ infix – *caphimì* 'little (ones)', that is, $Ch_1-VC_2-VCV > C_1-VC_2h-VCV$. In other dialects there is no *s/c* alternation, that is, *s* is a 'simple consonant' exhibiting no alternation with any other consonant, thus *sod-*, reduplicated form *sosòd*. No stems of form CV^2C were recorded.

Unaspirated stops can be geminated. Zide also noted half-length – longer than simple *C*, shorter than geminated *C* – voiceless stops, occurring in one restricted morphological environment – verb reduplication, for example, *tipi-*, 'to inform', reduplicated: *tif[tt]ipi*.

There is very little word-internal consonantal assimilation across morpheme boundaries. Among the very few examples are those with the auxiliary *lab-* (used in forming progressive tenses) before *-ken* (*ki-en*) and *-jen* (*ya-en*) > *lakken*, and *lajjen*, $V+lakken$ 'is V-ing', $V+lajjen$ 'was V-ing'. For example,

- | | | |
|-----|-------------------------|-------------------------|
| (3) | <i>kukullaken</i> | <i>kukullajjen</i> |
| | <i>ku-kul-lab-ki-en</i> | <i>ku-kul-lab-ya-en</i> |
| | REDPL:SEND-PROG-T/A-ITR | REDPL:SEND-AUX-T/A-ITR |
| | 'is sending' | 'was sending' |

Note that Drake (1903) preserved the earlier–pre-assimilation'–forms: *llabjen/* and *llabken/*.

Most morphophonemic complication (as pretty much everywhere else in Munda) occurs in verb forms and demonstrative derivatives, almost always with verb suffixes or demonstratives, particularly with those demonstrative stems with a final vowel, for example, in forms exhibiting vowel harmony (vowel harmony is found as elsewhere in NM in demonstratives): *in-èn* > *enèn* 'here', *qi-èn* > *qèèn* 'there', from *in-* 'this', *di-* 'that', *-èn* (locative), *i(n)to?* > *eto?*, *-to?* 'quantity', 'this much/many (of)'. Front vowels (i.e. the Dem stem *i(n)*, and the back vowel *u* in *hu-* 'a further that, yonder', 'this' harmonize (are lowered) to *e* and *o* respectively before (more heavily phonetically stressed) following syllables with the vowels *-e* or *-o*, for example, *hu-to?* > *oto?* 'that many'. For more detail on the morphophonemics of demonstrative forms, see the section on demonstratives (section 3.1.7).

The initial syllable changes to harmonize with the more stressed final. Glottal stop is lost everywhere but before #. In verb morphology vowels on $-CV$ suffixes contract with and are reduced when preceding suffixes with initial vowels, usually the $-è?$ transitive past, or $-en$ intransitive (participial) past, and $-à?$ genitive, $-èn$ locative. Examples include $-ki-e?$ > *khè?*:

- | | |
|-----|-----------------------------|
| (4) | <i>kulkhè?nèj</i> |
| | <i>kul-ki-è?(n)ej</i> |
| | SEND-INTNS/TLOC-PST.TR-3OBJ |
| | 'sent him' |

Glottal stop finally in affixes ($-à?$ genitive, $-è?$ past transitive) is lost (as in the above example) except before #, for example, *ijn-à?* *jumu* 'my name' > *ijnà jumu*, but its

morphophonemic alternate *-g* is found before a following vowel, where the *?* is not an affix-final sound (e.g. in the genitive, past transitive, etc.) but the final consonant of a noun or verb, for example, *urag-à doroja* ‘the door of the house’, *ḍoḍòg-à pèèla* [REDPL:look/see-GEN before] ‘before seeing/looking’. For some speakers for some nouns the *-g-*, originally an alternate of *?* when final on a verb or noun-stem is extended, that is, reanalyzed, it is treated as part of the genitive affix (thus, a new allomorph, *-gà?*), in a few examples: *yisu-gà?*, ‘Jesus’-reconfigured genitive, *yisugà?* *dherom* ‘Jesus’s religion (dharma)’ instead of the expected – earlier – *yisù dherom* on the analogy of *urag-à?* *doroja* ‘the door of the house’. Nagaraja records *byaugasutu*, that is, *byau-gà?-sutu?* from *byau* ‘wedding’, *-gà?*- GEN- and *sutu?* ‘before’ and ‘before the wedding’. The new genitive allomorph *-gà?* occurs, as above, on V-final morphemes for some speakers, but is probably spreading. This marks genitivity more prominently. Note that, as mentioned above, when a name ends in a vowel, the only morphological indicator of genitivity is the low tone, for example, *dadù ura?* (from *dadu-à?* *ura?* ‘Dadu’s house’.

3 MORPHOLOGY

3.1 Nominal morphology

Nominal morphemes in Korku are not marked by classifiers or class markers of any kind. In the non-borrowed vocabulary Korku simple – monomorphemic – nouns can be monosyllabic, dissyllabic, and – not often – longer than that. There is little derivational morphology (older derivational affixes can be identified in a few forms). A great deal of Korku vocabulary is borrowed from local Indo-Aryan, and a perusal of Girard’s lexicon shows that as many as half the nouns she lists are borrowed.

3.1.1 Number

Inanimate nouns are not usually marked for number (see also the sections on pronouns and demonstratives, sections 3.1.6 and 3.1.7) and quantifying adjectives are used. Animate nouns – particularly humans – require number marking: dual with *-kij*, and plural with *-ku*. As noted above, at least one Hoshangabad Korku dialect has lost the dual. For some kin-terms there is a non-singular infix *-(V)p-*, for example, *kup(a)rar-ku* ‘grandchildren’ from *kurar* ‘grandchild’. These contrast with ‘simple’ duals and plurals. They are used with third (and possibly second) persons. The *-(V)p-* infix also occurs with two adjectives, *khad* ‘big’ and *sani* ‘small’. The same (?) infix is used with verbs to form reciprocals (see section 3.2 on verb morphology). The *-(V)p-* is found in very few distributives derived from demonstrative stems, for example (Girard) *copta?*, *copthà?*, the distributive of *coto?* ‘how many’, that is, ‘how many each’ (cf. Hindi *kitne kitne*). Distributives are usually formed by repeating the word, for example, *je je* (*je* ‘who’) ‘who all’. Note also the *-Vm-* infix in some South Munda languages, for example, Gutob *b-um-uyay* ‘a set of (each other’s) brothers’ from *buyay*, ‘older brother’. For demonstrative derivatives we find the suffix *-(n)ḍùr* for non-singular (see below). Nagaraja points out that when we find (animate) plural suffixes with certain inanimate nouns, the word means the ‘inanimate’s people’ – people somehow associated with the noun, for example, *ura?-ku* (*ura?*, ‘house’), ‘the members (residents) of the house’.

Other postpositions with the genitive *-à?*: location and taking the locative suffix *-èn*

- (8) (a) *-à?-mera-èn* ‘near’ for example *ratanàmeran* *Ratana-à?-mera-èn* ‘near Ratana’ (the use of *-mera-* also can occur with possessive meaning, like for example Hindi *mere pas* ‘near me’, but also ‘my (but usually ‘my, meaning with/on me’). Korku *ijn-à?-mera-èn. ijnàmeran*
 (b) *-à-ljn-èn* ‘on, on top of’
 (c) *-à?-italiṭa-èn* ‘under, below’
 (d) *-à?-PP-èn ura?-à?-samma-èn* → *uragàsamman!* ‘in front of the house’
 cf. Hi *-ke sammne*
 (e) *-à?-tau-èn* ‘behind’
 (f) *-à-sutu?* before (temporal)
 (g) *-à-baadòn* after (temporal)

3.1.3 Person

Person is not widely expressed in the Korku nominal system. Inalienable possession is marked by *-te?* ~ *-ṭe?* (a cognate affix to which is found in most of the Munda languages) for some kin-terms and body parts, for example, *kimin-te?* ‘daughter-in-law’, *gal-te?* ‘cheek’, used with third and second persons, for example, but not the first person in these examples, that is, not of the speaker’s (daughter-in-law, cheek, etc.). It may follow a genitive marked possessor.

- (9) *aḍi* *dijà?* *ijṭe?* *par* *ṭuyè?*
 dij-à? *ij-ṭe?* *ṭu-ya-è?*
 current he-GEN shit-3 completely carry.away-TLOC-PST.TR
 ‘The current/flood carried away all his shit.’

3.1.4 Definiteness

Not explicitly investigated for this chapter in Korku.

3.1.5 Class/gender

Gender is weakly expressed in the Korku nominal system. There is a feminine suffix *-jei* (*-jay*, *-je*) found with a small number of kin-terms, for example, *kon-jei* ‘daughter’ vs. *kon* ‘child, son’. There are forms borrowed from Hindi where the masculine ends in long *a*, the feminine with *-i* (long *i*). Similarly with adjectives, for example, *lanṭa* masculine vs. *liṅṭi* feminine ‘lame, limping; lame person’.

3.1.6 Pronouns

The pronouns of Korku exhibit slight dialect differences. In Nagaraja (slightly edited), we find the following set of basic pronominal forms for the first person singular, dual, and plural:

- (10)
- | | Singular | Dual | Plural |
|------------|---------------|-------------------------------------|---|
| 1st person | <i>ijn(j)</i> | <i>alaj</i> INC.
<i>alij</i> EX. | <i>abuj</i> INC.
<i>ale</i> EX.
<i>alam</i> (not specified) |

Girard's materials on the other hand yield the following slightly different set:

(11)		Singular	Dual	Plural
	Girard	<i>ij</i> INCL EXCL	<i>alom/lalam</i> <i>alij</i>	<i>abuj/labū</i> <i>ale</i>

Both have a similar set of second person pronominal forms

(12)		Singular	Dual	Plural
	2nd person	<i>am</i>	<i>apij</i> (<i>apo</i> honorific)	<i>ape</i>

Girard's third person pronouns are the following *in-ij*, *in-kij*, *in-ku*; Nagaraja gives the same forms in his own transcription, and calls *qi* – inaccurately – neuter.

(13)		Singular	Dual	Plural
	Masc/Fem	<i>qic*/in</i>	<i>qi-kijj</i>	<i>qi-ku</i>
	Neuter (Girard, Zide)	<i>qi*</i>	[<i>qi-kijj</i>]	[<i>qi-ku</i>]
	Reflexive (Girard)	<i>heclj</i> <i>hej</i>	<i>i-kij</i>	<i>i-ku</i>

Zide agrees with Girard, and also notes occasional – older – *alaŋ* for 1st dual inclusive from which *alom* and *alam* derive.

I include here also the pronominal stem *to(n)-* ‘which (one)’, – paralleling *i(n)-*, *qi-*, and *hu-* (see below), its pronominal derivatives and the adverbs derived from *to[n]-/to(n)-/co-*, *je*, ‘who’, and miscellaneous other forms.

(14)	Singular	Dual	Plural	
	<i>ton-ej</i>	<i>ton-kij</i>	<i>ton-ku</i>	<i>ton-e</i> (INAN) ‘which (one)’

Another interrogative stem similar to *ton/ton-* appears as *tu[u]*, yielding *tùùdej* [SG.] *tukijè?* [DL] and *tukuè?* [PL]. *je* ‘who’ takes case suffixes, and also occurs as an adjective, ex. (Nagaraja) *je kon-te?* ‘whose son’ (Nagaraja) *je-ka* ‘anyone’ *je je* (distributive) ‘who all’. The difference between *je* and *tone* (as adjectives) parallels that of Hindi *kya*, ‘what’ and *kaun-(sa)*, the former meaning ‘what Noun’, the latter ‘which Noun of an understood set’.

Another interrogative pronoun is *coj* ‘what’ (Nagaraja *coch* Girard *coj/co?*). Girard points that *coj* is now used by some speakers (on the model of Hindi *kya*, ‘what’) as a question marker, usually at the beginning of a sentence. *coj* as a pronoun takes the expected case suffixes. *co?-èn* (*-èn* LOC) ‘why’, *coj-à?*, ‘why’. It is not clear under our current state of understanding what exactly is the difference in meaning between these two variant forms of ‘why’.

Interrogatives derived from *to(n)-* include: *ton-èn* ‘(at) where’ – paralleling *enèn* ‘here’ *qèèn*, ‘there’ *huèn* ‘yonder’. Another interrogative element is (Girard) *tujan/ tūān* (Zide) *tūgàn*, (Nagaraja) *tojan* ‘to where, whither’ also *tuwon*. Note Girard also has the form *tujan-e*, ‘where is it’ (see also the comment on interrogatives as verbs below, cf. *cuphàr*, ‘(do) how’).

(15)	<i>am cophar-khe</i> {cf. Zide <i>cuphàr-ki-è?</i> ‘do how, in what way?’}
	you do.what-PST
	‘What did you do? (Nagaraja 1999)

But the members of that set are expressive. On the set with *iùde?*, *huùde?*, etc. see below. The suffix *-bi-* from Hindi *bhi:* (emphatic) ‘too’ is used to form indefinite interrogative ‘ever’ forms, often with the emphatic *-ka*, (Nagaraja) *co:la-ka-bi* ‘whenever’.

The interrogative stems *co(-j)* and *to(-n)* are probably related – earlier perhaps even the same morpheme – but are now separate stems with a difference of meaning (more or less the one offered for *je* vs. *to(n)* above). Those derived from *co(j)* in addition to those taken up under demonstratives: (*coto?* *cuphàr!*(Girard) *cophar*, *cuphìn/cuthìn*) we find (Girard) *cupta?*, (Zide) *cupthà?*, the distributive of *coto?* with infix *-p-* (see *-p-* in kin-term plurals), that is, ‘how many each’, and forms such as (Girard/Zide) *cutthà/cutthà*, (N) *coffa* ‘what sort of, who’ (*-ta* as an obscure pronominal suffix, perhaps a nominalizer in North Munda); *coðla* ‘when’ (*-la* not identified elsewhere), and (Zide) *coðj(-)à(?)* (Girard) *coja* (Nagaraja) *coja*, (Girard) *coja-ki* ‘because’ (cf. Hindi *kyō* ‘why’ *ki*). Note that the lengthening in *coðja?*, and *coðla* is ‘expressive’ in nature.

3.1.7 Demonstratives

The demonstrative stems are *i(n)* ‘this’, *qi-* ‘that’, *hu-* ‘yonder’ (a further ‘that’) and *ton-* ‘which’. They take – as nouns and adjectives – the number suffixes, for example, for animates: sing *-ej*, dual *-kijn*, and plural *-ku* and for inanimates (not marked for number, but usually singular) *-e*, and non-singular *-(n)qùr*. Note that expected **hu-[n]qùr* was not recorded (see also (Zide 1991)).

Additional forms were recorded by Nagaraja, and there are also other – possibly archaic – less productive forms. These can be considered pronouns as well as demonstrative forms. Nagaraja gives several forms with final *-e* not noted elsewhere *no:je*, (near, SG. INAN), *hu-je* (remote, SG. animate). Nagaraja distinguishes very near/near/far and remote. Nagaraja’s morpheme boundaries (e.g. *i-nij* as well as the last two) are misplaced. What these *-e* forms mean is not stated. Perhaps these are equivalent to Zide’s ‘exclamatory verbs’ with demonstrative stems, for example, *huùde?* (remote, SG. animate) ‘there (yonder) he is’ *i-kijn-è?* ‘here these two are!’ (Nagaraja does not record glottal stops), with slightly different morphophonemics. What is interesting in Nagaraja’s demonstrative table – and not noted elsewhere are the ‘very near’ (*ni-*) – contrasting with ‘near’ (*no-*, *i(n)-* forms), and the ‘remote’ vs. (remote) ‘pinpointing’ forms. He gives *ha-lho-* and *hu-* as remoter stems, and usually *ho-* (once *hu-*) in the ‘pinpointing’ forms. The *-e* suffix occurs with both the remote non-pinpointing forms as well as (two of the) pinpointing forms. There is considerable local variation in demonstrative stems and derivations. Girard, surprisingly, does not go into detail on the subject. Zide finds different – apparently archaic – forms in his ‘exclamatory verbs’ – as does Girard.

Zide finds what might be called ‘pinpointing’ demonstrative stem modifiers, *na* occurring before *i(n)-* and *han* before *qi-*.

(16) *na ini koro* ‘this very man’ *han qi japae* ‘this very/particular woman’

The demonstrative derivations with *-è?*, not neatly matched with the regular pronominal and demonstrative nouns and adjectives, are as follows:

(17) inanimate (near)
naàne? ‘here it is/they (INAN) are!’ *haàne?* ‘there it is/they are’

- (18) animate (near) *ni-li-* (no difference in meaning noted)
 sing *niide?/liide?* dual *nikijè?, ikijè?* plural *nikuè?, ikuè?*
 (far) *hu-* SG. *huùde?* dual *hukijè?* PL *hukuè?*

Note that there is vowel harmony (lowering or height harmony), with the demonstrative stems in *i-* or *u-* becoming *e-* or *o-* before a manner suffix with *-e* or *-o*, for example, *i(n)-to?* > *eto?*. Note also that the *-n-* in *i(n)-* seems to be an old nominalizer, cf. Santali. But in Korku it seems to be just a linking or liaison phoneme, appearing where *i(n)-* is followed by a vowel, for example, *in-ej* > *inij*, with a ‘raising’ harmonic pattern, but it is not found in the ‘exclamatory verb’ demonstrative derivatives, that is, *iide?* {*i-ej-è?*}.

Adverbs are regularly derived from these demonstratives, at least from a pair of them usually *i(n)-*, ‘this’ vs. *qi-*, ‘that’ or from a set of three. The adverbial derivatives that commonly appear in such formations are offered below.

- (19) (i) *-phùn* ‘manner’ – also *-thìn* (Nagaraja/Zide) *i-phùn qi-phn* (+Zide *hu-phùn*), (Nagaraja) *i-thìn, qi-phìn*
 (ii) *-phàr* manner (Nagaraja) *i-phar*, (Nagaraja/Girard) *qi-phar* Zide *ĩdàr/ hĩdàr, qĩ-dàr/qhĩdàr* but interrogative *cu-phar* ‘how’

Note that Nagaraja distinguishes *-phar* ‘like this’ from *-phin* ‘this way’. Note also that for Zide’s speakers the initial syllables of demonstrative derivatives where the second – final, stressed – syllable is low-toned, frequently aspirate the first syllable. Note that some of these adverbs can be used as verbs, for example, *cuphàr*.

- (20) *ĩj cuphàre?* {*cu-phar-e*}
 I HOW-MANNER-INAN.OBJ
 ‘I how it?’ that is, ‘How shall I do it?’
 (iii) *-to?* ~ *-to?* Quantity as in *e-to?* ‘this many’ *qe -to?* ‘that many’ *o-to?* ‘a greater than many’ that is, still more than *qeto?*

Nagaraja gives an example with a number suffix after *coto?*, *coto?-kij*, ‘how many (dual)’.

Finally, there are the adverbial demonstratives of direction or movement toward (the transcriptions vary, some use *ŋ*, some [*~g*] (i.e. nasalization + velar obstruent vs. velar nasal); here I have transcribed all of them with *~g*, for example, *ĩgàn* ‘hither’, etc.

- (21) (Nagaraja) *ĩgalĩgan qĩgalqĩgan* (Girard) *ĩgan qĩgan ũgan/hũgan* (Zide) *ĩgàn/ hīgàn qīgàn hugàn*.

3.1.8 Numerals

The numeral system of Korku is complex as it is in many Munda languages. Numbers for ‘11’ and above are mainly from Indo-Aryan (and from what could be remembered of these, from ‘30’ and above there was a vigesimal system as in many Munda languages), but the first decade exhibits (or did 40–50 years ago) a clear Munda base. The basic set is

- (22) *miã?* ‘one’ (also *mijà?*)
bari ‘two’

(i.e. reduplicated or copied), and (iv) echo-words, where Word₂ derived phonologically from Word₁.

Finally, note that my interpretations of the Korku materials differ from those of Nagaraja in the placement of word and clitic boundaries, and in hearing the distinguishing features thereof, that is, N-sfx vs N+X (X is usually Noun₂). I have here distinguished compound nominals (bound forms) from multi-word noun phrases. Sometimes – rarely – the two are hard to distinguish, and the choice is arbitrary, but not in most of the cases. Girard and I are usually in agreement in what we have heard.

3.1.11 Adjectives

Adjectives are intransitive verbs in most cases – there are a few exceptions, where the adjective occurs only as prenominal adjective. The intransitive verb adjectives can be transitivized/causativized by adding transitive verb inflections. The adjective *khad*-‘big’, when it occurs as a prenominal adjective it takes the plural $-(V)p-$ > *khaped*, noted above in section 3.1.1 in connection with *sani*, *caphini*, ‘small’, but not when it is used as a verb-stem. An example of a predicate ‘adjective’ in Korku may be seen in (24).

- (24) *dij doroja khad-kh-è?*
 S/he door big-INTNSV-TR.PST
 ‘He/she enlarged, made big(ger) the door.’

Korku, unlike the South Munda languages has no causative marker (usually realized as verb prefix). Instead, the transitive- or transitivized-verb is used with syntactic rules marking agent, object(s), etc. with the appropriate case suffixes. See the functional classification of verb-stems below, and for the contrast between (when both are possible) N Adj<verb> #, ‘N is Adj’, and N Adj-*bà*- also translated as ‘N is Adj’ (with *-bà* the verbal nonpast predicator, which is realized as *-ò* in Lahi (Hoshangabad) Korku). Such distinctions – and the ‘copular’ constructions involved – are expressed in various ways, depending, often, on the existence of a (borrowed, more and less integrated and productive) copula, ‘to be’ – sometimes only on the awareness of the distinction in the – dominant – source language before borrowing of an overt morpheme takes place. Korku has locational copulae, that is, ‘to be (somewhere)’, but not an existential one – except, as in some of Nagaraja’s examples, where *ho-* has been borrowed. Some North Munda languages (probably all the Kherwarian languages), for example, Santali, have an entrenched positive copula (note also that all have negative copulae), so such forms and constructions are not to be reconstructed for proto-North Munda. Some Korku idiolects seem to have moved the emphatics *-ka?* and *-se* into (incipient) copular roles.

Unfortunately, a detailed description of the grammaticalization machinery involved cannot be provided. The areal pressure in a bilingual community where the dominant outside community uses a copula will probably strengthen its advent and use in all the various Korku dialects.

3.1.12 Adverbials

Adverbials generally precede the verb, adjective, or adverb they modify in Korku. Examples may be found in cited forms throughout the chapter.

3.2 Verbal morphology

As is common for most Munda languages, (basic) nouns, adjectives, and adverbs in Korku are used as verb-stems, with the following (usually intransitive) meanings: with Nouns, 'to become/act as the Noun', with adjectives 'to be/become Adjective', with Adverbs 'to what/how', that is, 'to be what' or 'to do-how'. There are many onomatopoeic verbs, usually of the form Verb₁+Verb₁, for example, a type of (probably expressive, augmentative or repetitive) reduplication or stem-copy: *kefej+kefej* 'to clatter', *reñ+reñ* 'to whimper of a very young baby'.

With regard to verb-stem shapes, most Korku (non-borrowed) verb-stems are monosyllabic or dissyllabic (the few longer stems seem to be borrowed). Nagaraja finds no simple verb-stems of more than two syllables. I found a few, all borrowings. One such stem is *phijito* – (from a local Hindi noun, *lphajii(ha)t* (Urdu *faziihat*), in Korku 'trouble, annoyance, persecution'. As a verb it is reduplicated to *pipijito*, 'to persecute, to annoy'. Another stem of this type is *bhaõta* 'to scratch an itch', with its reduplicated form *bhabàõta*. Such borrowed verb-stems in Korku are usually reduplicated when transitive.

There are many monosyllabic and dissyllabic simple – monomorphemic – stems. The 'extension' morpheme *-e* is added to some basic stems (see below on a few verbs of motion – *sen-*, *niñ-*, etc. (for Nagaraja and Drake, but not for Zide) where no real object is to be understood, and in derivation from Nouns, Adverbs, etc. where some of the V-*e*- stems again have no 'understood object'.

Transitive verbs mark animate and usually animate (direct) objects with pronominal markers (mainly) simply derived from the pronoun baseforms, that is, *-jn* 'me', *-mi* 'you' (the only object marker not simply derived from the base nominative form), *-ej* 'him/her', *-e* 'it', etc.

Like nominal compounding, double verb-stems are of several kinds: (i) Verb₁–Verb₂ both occur independently and the combination is 'more than the sum of the parts'. (ii) where Verb₂ is an echo form, derived by a phonological rule from Verb₁. (iii) Onomatopoeic stems with limited morphology where the two pieces – joined by a plus juncture – are identical or a copy one from the other. Most of these have limited 'Mode', Object, etc. morphology, apart from the tense/aspect markers.

- (i) (Nagaraja) *asi-jom-* 'to beg and eat'. This he says is a simple sequence with the combination meaning no more than two related successive actions with their regular meanings, that is, a type of serialized formation. (Nagaraja) *jom-kab-* (lit. eat-bite), roughly 'eat and drink'. *ljojom+numul*, 'to eat, drink – and carry on.' There are also constructions with Verb₁ Verb₂ that are not double verbs in this sense. Nagaraja gives *giñij-sene-sutu* (Zide: *Isutu?*): 'sleep-go-before', 'before going to sleep'. I take this as not a more tightly bound double verb, but as a sequence of verbs, specifically an infinitive of a Verb that is serving as a complement of a verb of motion, here meaning 'to go (off) "to V", – or 'for the purpose of V-ing', that is, 'before he went to sleep'. One commonly finds in Korku such formations as the following:

- (25) *dij giñij olen dij jojom olen*
 he sleep go:PST:ITR he REDPL:eat go:PST:ITR
 'He went (off) to sleep.' 'He went off to eat.'

where the reduplicated stem of the monosyllabic verb (at least in this construction) is clearly marked as a complement of the verb 'go', appearing in this dependent stem allomorph.

- (ii) The phonological formation rule for Verb₂ ‘echo verbs’ closely parallels that of echo forms of nouns: copy the stem and add /*m-*/ when the verb begins with an initial vowel, or replace the initial consonant of a verb when there is one with /*m-*/. This sort of formation and its meaning is pan-Indian, with a bilabial (‘prefix’) in a number of Indian languages as the added consonant/ *kusù+musù-bal*, ‘to whisper’ is a pseudo-echo formation because there is no free /*kusù*/ to derive the /+*musul*/ from. The meaning of this sort of double stem is ‘to Verb etc.’, or ‘to Verb and do associated activities’ /*saṛub+maṛubl*, ‘to run, etc.’; /*lkhù+mukhùl* ‘to hide – and perform actions related in a particular context’.
- (iii) There are many of the copied verb-stem, some speakers using them frequently. Examples:

l keṭej+keṭejl, ‘to clatter’ (of hooves); /*bhen+bhenl*, ‘to buzz or whine – of certain insects’. Some of these expressive words stretch the usual sound pattern of the language. /*reṇ+reṇl*, ‘to whimper of a very young baby’. These verb-stems don’t take Object or ‘Mode’ suffixes and are most common in the Present/Future form.

3.2.1–3.2.2 Subject and object

Korku regularly marks object on transitive verb forms. Unlike the other North Munda (Kherwarian) languages it does not regularly mark subject on transitive verbs. The few exceptions: Nagaraja notes subject-marking for a few examples with only one verb, /*jom-*/, ‘to eat’, and only with certain pronouns. How these forms are understood as marking subject rather than object is not clear, perhaps it is due to contextual semantic interpretation, since (here anyway) one does not eat people, so subject is understood. Perhaps the (not clearly identifiable) ‘mode’ (or tense/aspect form, see below) preceding the subject is distinctive, that is, it is a morpheme indicating: switch from object to subject (a type of ‘inverse’?). One of his forms /*jom-ja-pijl*, ‘you (DL) eat, will eat’ has the ‘object’ following /-*jal*. Is this /-*jal* with a (new) person-role switching meaning or a new mode marker, /-*jal*, with that particular function? Drake (1903), Girard (1965), and Zide (field notes) record no examples of subject-marking in their Melghat materials. However, I worked very briefly with a Lahi (Hoshangabad) informant and did find a few such forms. This suggests that subject-marking may vary in a dialect chain across North Munda; remember that Melghat Korku is the westernmost North Munda language.

One systematic exception to this lack of subject-marking in Korku verbs is that subjects are marked (in all grammars) for a few locative verbs, in particular the common /*ṭàkha(?)l*, ‘to be (located)’. Nagaraja translates it, confusingly, as ‘have’, indicating that one common use of the verb is in locative constructions (corresponding to English subject + have constructions). So, for example, the normal way to say ‘he has four sons’ in Korku is literally ‘in him four sons are’:

- (26) *dij-èn uphùn kon-ku ṭàkha-ku*
 3-LOC four SON-PL LOC.COP-3PL
 ‘He has four sons.’

The 3 PL /-*ku*/ marks the plural subject, ‘sons’. This verb /*ṭàkha(?)l* occurs with third person subjects only, and there is considerable variation in the forms, for example, /*ṭàkha(?)* INAN, /*tijkà?* ANIM SG., /*ṭàkhakij* ANIM DL /*ṭàkhaku*, ANIM PL. Drake (1903)

finds a few first person forms but says they are dubious, ‘not worth considering’. His dual and plural forms are based on *tijka*: *tijkakiŋ* and *tijkaku*. The inanimate form is, in the P/F (with *-bà!*) *Itaakaayebal*, that is, *Itaàkha-e-bà*) such tense forms with *-bà* forms are not common or found at all elsewhere. As well as *taakha[ʔ]* we find *taakhaa* and *taakaa*.

The *l-j-l* or *l-ijl* in the singular presumably represents the usual 3 SG. ANIM *l-ej/*-infix here – in demonstrative derivatives and as the 3 SG. ANIM object suffix in transitive verb forms.

The other verb constructions with marked third person subjects are: Place-LOC-Subject (27), and in ‘exclamatory’ (or interjectional) verbs in Demonstrative-Subject-*-èʔ* constructions (29). In the former instance, such constructions occur with ‘place words’, for example, house, Nagpur (a city), and one can, with animate subjects, suffix the usual singular, dual, and plural suffixes to the Place-LOC formation.

- (27) *ɖi-ku urag-èn-ku*
 He-PL house-LOC-3PL
 ‘They are in the house.’

ɖij Nagpur-èn-ɖan
 He Nagpur-LOC-PST:AUX
 ‘He was in Nagpur.’

For inanimates the simple locative is used:

- (28) *kitaabo urag-èn*
 book house-LOC
 ‘The book is in the house.’

With a limited set of old, irregular demonstrative stems, plus the third person suffix, the exclamatory verbalizer *l-èʔ/* is added (this set for Zide includes *li-*, *lu-l*, and for Nagaraja *n-i-*, *i-*, ‘this’ *hu-* ‘that, yonder’, and *u-* interrogative, ‘which (person)’. The *l-èʔ/* in these forms is not to be identified with the past transitive *l-èʔ/* but with the emphatic ‘verbalizer’ in the emphatic negative copula *bannèʔ* {*ban-* (*n*)*èʔ*}. The *ln-* in *ln-i-l* is a prefixed element (compare the demonstratives in Kherwarian, see Ghosh (this volume); Osada (this volume); Zide (1972)).

- (29) *i-kij-èʔ am-àʔ tarei-kij,*
 DEM-DL-VBLZR YOU-GEN daughter-DL
 ‘Here they are! Your (dual) daughters/girls.’
- huùɖeʔ, ij-àʔ aba*
hu-ej-èʔ
 DEM-3-VRBLZR I-GEN father
 ‘There (yonder) he is!, my father.’

Object-incorporation of pronouns. The incorporated forms of the pronouns are simply derived from the full forms in almost all cases (see section 3.1.6) but for third persons the usual elements in the demonstratives are found: *-e*, INAN, *-ej* ANIM/3 SG., *-kij* ANIM/3 DL, *-ku* ANIM/3 PL; *-ijn* 1 SG. and *-mi* 2 SG., etc. Note that for Nagaraja *-mi* is commonly used for not only second singular but, curiously, first singular as well. For the other pronouns *ape*, *ale alom*, etc. the *a-* is deleted and the incorporated/monosyllabic form of the pronoun is used, for example, *l-pe*, *-le*, *loml*, etc. The

inanimate *-e* is used in verb forms without a following ‘mode’ suffix. With a mode suffix inanimate remains unmarked, that is, is marked with Ø *tol-ki* ‘tie it/them’.

- (30) *baki tol-e*
 NEG tie-INAN
 ‘Don’t tie it/them.’

Note that the suffix *-e* also marks the imperative for intransitive verbs, for example, *haj-e*.

My materials have a few intransitive verbs that take */-e/* before the Present/Future */-bà/*, that is, *senebal*, ‘to go’, not **senba* (but Drake seems to allow some such forms)> among the verbs that do this *haj-lhej-*, ‘to come’, and */sen-/* ‘to go’ should be mentioned. Also noted was *niṛ-* ‘to run off, escape’. Note that for */niṛ-/* but not */sen-/* or *haj-lhej-* there is a reduplicated form, *niniṛ-*, ‘to cause to, get to run off’.

Nagaraja has more intransitive verb *-e* forms (I am not speaking of those derived from nouns, adverbs, etc.), for example, *biḷ-*, ‘to rise, awaken’. These are verbs of motion, the */-e/* perhaps marking or ‘reflecting’ that. Perhaps there is some phonological motivation as well, for example, a preference for dissyllabic bases occurring before *-bà*.

3.2.3–3.2.8 Tense, aspect, mood, direction/orientation, voice, and finiteness

The categories of tense, aspect, mood, directionality, voice, and finiteness are all interconnected in Korku in a complex manner and thus not differentiated by sections explicitly in this presentation. I do attempt however to roughly follow the order of presentation of these facts, when possible, with those sub-headings found in other chapters in this volume.

There are two tense systems in Korku, Past and Present-Future (PRS/FUT), as there are in some other Munda languages (and, presumably, Proto-Munda). Some Munda languages have – independently – innovated Futures. Korku has not. The only strictly Present (and not Present-Future) verb is the borrowed auxiliary *hoy* (see below). The Past and PRS/FUT systems are not symmetrical.

There is no PRS/FUT tense marker. That tense is regularly indicated by the Predicator or finitizer suffix *-bà* (realized as *-ò* in the Hoshangabad (Lahi) dialect); this is probably cognate with the finitizer or predicator suffix *-a* found in Kherwarian languages, but there it is found with (almost) all finite tense/aspect forms, not just those of the PRS/FUT series. Nagaraja (1999) calls it a tense suffix marking the PRS/FUT, but I for various reasons – for example, the loss of *-bà* in participles and adjectival forms – sees it as something separate, looser – less tightly bound – than a tense suffix. The Past is marked by two tense suffixes: *-en* for the Intransitive (ITR) and *-è?* (with variants *-e*, *-en* in Nagaraja and Drake who do not transcribe tone) for Transitive (TR). The ITR too has variant forms, usually in defective verbs: */-an*, *-on/*

There are two large inflectional classes of verbs: a simple stem class (Class-I), where the (intransitive) verb takes the predicator *-bà* directly, for example, *biḷ-bà*, ‘gets up, arises’ or *iraḷ-bà*, ‘returns (ITR), *suban-bà*, ‘sits’, and a derived stem class (Class-II), where there is a (reduplicated) transitive stem and a derived intransitive in *-ù?*, for example, *gujùba* {*goj-ù-bà*}, ‘to die’. A small set of intransitive verbs (limited in my materials to *sen-* ‘to go’ and *hej-lhaj-* ‘to come’ and a few others, but found with a larger set in Nagaraja and Drake) that take *-e* in intransitive forms, that is, *sen-e-bà*, ‘goes’, *hej-e-bà*, ‘comes’, *niṛ-e-bà*, ‘runs/goes away’.

There is some variation among the sources as to which inflectional class a given stem belongs to. For example, Nagaraja has *biq-* in the Verb-*e* class but Zide does not. Drake calls the *-e/* an increase or increasing element, Zide calls it an extension. It is common in deriving verb-stems from nouns, for instance (see Drake for many examples). For the intransitive infinitives (used in the prohibitive, for instance) Zide has the regular intransitive (pre-*bà* forms), for example, */biq, ira?, guj-ù?, sen-e/*, but, apparently not */*niṛ-e/*. Here the transitive infinitive R-V (*/niniṛ/*) may be used.

Class-II transitives, where corresponding Class-I verbs occur, are marked by [C]V-stem reduplication, */bid/* and */bibid/*, */ira?/* and */iira?/*, */suban/* and */susuban/* and have causative meaning. Thus we find */bid-l/*, 'to arise, get up/' and */bibid/*, 'to wake up someone, get someone up/', */ira?/* 'to return' and */iira?/*, 'to return something/someone', and */suban/*, 'to sit' and */susuban/*, 'to seat someone, get someone to sit'. There are no reduplicated forms for 'come' (*/hej-*) and 'go' (*/sen-*). As to which verbs are Class-I, which take the form Verb-*e*, and which Verb-*ù?*, and the semantic notions underlying the selection, nothing much can be said. For */goj/*, the meaning 'to kill' is somehow 'primary', and */guj-ù?/*, 'to die', 'be killed' is derived. Thus, this appears to be a kind of passive-type derivation but not all stems with this inflectional class marker share these functional semantics. Note that there is no form **/goj-bà/* possible.

A schematic of the inflectional classes in both past and PRS/FUT forms are offered in (31).

(31)	<i>Intransitive</i>	<i>Transitive</i>
PRS/FUT	Verb- <i>bà</i> Verb - <i>e-bà</i> Verb -<Mode>- <i>ù?-bà</i>	REDPL-Verb- <i>bà</i> Verb-<Mode>-Object- <i>bà</i>
PST	Verb -Mode- <i>en</i>	Verb -<Mode>-Object- <i>è?</i>

The Mode suffixes are one of the characteristic features of the Korku verb system, with parallels in Kherwarian languages as well. There are five commonly used ones of these, namely *-ki-*, *-ya-*, *-li-*, *-wa-*, and *-ṭhà-*. These occur after the Verb stem and before Object suffixes (including zero) and the predicator suffix-*bà* in transitive verbs, and more restrictedly with some intransitives as well. A brief description of the same follows.

The first two Mode suffixes are common and recognized by Nagaraja and Drake, although their functions are differently analyzed (and in some cases the suffixes function differently in their material as well). I don't find *-ki-* to be a transitivizer as they do. For Nagaraja the *-ki-* (his */khe/*) forms seem to be the simple – unmarked – forms, not markedly intensive, because the simple Modeless forms do not exist, or are found very rarely. It is an intensive occurring with intransitives and transitives, more commonly with transitives. The element *-li* occurs as a Mode suffix with cislocative meaning, infrequently in Nagaraja and Drake, but Nagaraja tends to see it at a separate verb, which I think is wrong. I explain these in my data briefly and then have some remarks about these morphemes and related ones in Drake and Nagaraja. Suffixes *-ki* and *-ya* also occur before *-ù?* (for more, see the discussion of *ù?* that follows, my analysis of *-yù?* and *-khù?* is open to question).

With respect to verbal directionality or orientation, we find the *-li-* cislocative opposed with the *-ya-*, translocative. Note that *-ya-* also participates in another opposition in the past: *-ya-* vs. *-ki-*, in which *-ki-* indicates a recent past action, where *-ya-* an action performed in the remoter past. Nagaraja gives examples of *-ki-* with a past action not completed. *-ya-* before Object of the shape -VC- (i.e. *-ej,*

3sg., and *-jn*, 1sg) takes a second Mode suffix, *-ki-*, for example, not the expected and semantically simple */-en/* **kulyejbà*, but *kulyakhèjbà*. This is the only example of a Mode₁–Mode₂ sequence noted and remains unexplained. In the Past we find the expected Verb–Mode–Tense–Object order – */kulyènej/* {*kul-ya-è?*(>*n*)-*ej*}.

The Mode suffix */-thà/* is common and has a tentative modal meaning, for example, Verb-*/thà/*- thus for ‘try (eat) it-and-see’ we get *jom-thà-* while *qo?thà-* means ‘look and see’.

Zide also records the Mode suffix *-wa-*, a benefactive – that is relatively uncommon, with the beneficiary noun phrase appearing in the accusative (i.e. with */-khè?/*). Girard also records this form ‘for someone else’s benefit’.

The simple Past is used to translate ‘has sent’ as well as ‘sent’, the use of the Mode */-kil/* and */-ya/* indicate degrees of recency or remoteness, */-kil/* meaning ‘just sent, sent very recently’, or, as Nagaraja states, for some verbs an action begun but not necessarily completed. Thus, the ‘ongoingness of effect’ implied with ‘has V-ed’ is to some extent conveyed by the */-kil/* recent past.

Other Mode suffixes may be attested as well. For example, Nagaraja has */-pan/* where Zide has */-pen/* (<**ya-en* after nasal). Nagaraja also has a unique suffix after vowels */-pl/*, which seems to be the past intransitive – from */-en/*, or more likely, */-yen/*. Examples from Nagaraja (1999): *ukhu-n* ‘concealed’, from *lukhu-l*, ‘to conceal’ or *pura-n* ‘became complete’, from *lpura-l*, ‘to become complete’.

We need to mention here the very few verbs – and presumably archaic verb forms – that take {*-en*} directly with no preceding Mode suffix. Again there are differences in the set of these in the different grammars. All have the common form *ol-en*, ‘went’, *lol-l* being the suppletive past stem for PRS/FUT */sen-l*, and {*mhen-an*} – sometimes *lmenan/* and *lmenon/*, ‘it is said’, ‘they said’, a very common narrative introducer.

The inflectional element U */-ù?/* (also realized as *-Mode-ù?*) is ‘passive-potential’, with one or both connotations in the differing uses with different kinds of verb. It occurs with both transitive and intransitive verbs, in three realizations, namely */-ù?/* */-yù?/*, (< {*-ya-ù?*}), and */-khù?/* presumably from {*-ki-ù?*}. The form *-khù?* has a distinct but similar meaning for Girard and Zide, but is used differently in Nagaraja’s materials (see below). The analysis of */-khù?/* into *-ki-ù?* is morphophonemically neat, but no semantic interpretation of the morpheme combination is obvious.

In derivation from nouns, adverbs, etc. for Drake *yu* and only *yu* (not *u*) is used. For Zide there is semantic overlap, but *-ù?* is used for (stative) ‘to become Adjective’ with *-yù?* more ‘can (potentially) be’, is ‘to be Adjective’. The *u?* is used in forms meaning ‘to become X’ (in Nagaraja it is *-yu*). Note that verb-stems with final vowel or ? (from *j-y*) have alternative realizations of *-yù?*. We find also *li?*. Thus *gola-yù?-bà* comes out as either *golayùba* or *golaùba*; for *lgoj-ya-ù?bà/* we get only *lgo?ùbal*. That they form one kind of set is clear: in the Negative the set reduces to Verb-*ù?*, that is, in the negative, the contrast of Verb-*ù?*, Verb-*yù?*, and Verb-*khù?* are neutralized in favor of Verb-*ù?*: ‘Remain seated’ is *subankhù?* but ‘don’t “remain seated”’ is *baki subanù?*.

With transitive verbs the passive is formed with *-ù?*, for example, *dij mūdàyuba* {*mūdà-ya-ù?-bà?*} ‘he/she will be beaten/is beaten’. The form with */-yù?/* in place of */-ù?/* means, ‘He can (potentially) be beaten, he is to be beaten.’ With intransitives, for example, from *lira?*, ‘to return’ we find *ira?yùba* which has the potential transitive meaning. How the meaning of *V-ù?-bà* differs from the */-yù?/* form is less clear. Informants usually say that there is no difference in meaning. The difference is, perhaps for the *-u?* form ‘will, may Verb’ as opposed as contrasted with the *-yu-* form ‘is to/can Verb’.

/-ùʔ/ and */-yùʔ/*. Here one should note – as not all the linguists who have written on Korku have – that for Class-I intransitives there is a three-way difference between (Verb-*bà*), {Verb-*ùʔ-bà*} and {Verb-*ya-ùʔ-bà*}, for example, */subanbà/*, */subanùbal/*, and *subanyùbal* (*/suban/*, ‘to sit (down)’). Examining the contrasting meanings here should be particularly useful for understanding *-ùʔ* and *-yùʔ*. But we also have here the transitive/causative *susùbanba* derivatives, with *-ùʔ* and *-yùʔ* forms homonymous with those derived from the intransitive. These homonyms complicate the picture, that is, although distinctive meanings can be given for all five forms a preference for one (e.g. the *-ùʔ* form derivative from the transitive with passive meaning) may make the homonymous nonpassive intransitive */-ùʔ/* derivative of the intransitive a less favored interpretation of */subanùbal/* rare, or at least less favored. The simple infinitival stem – for example, */iraʔ/* has the usual – definite, customary – meaning of infinitives.

Drake finds a ‘potential’ */-ye/*, something the other dialects (grammars) do not have so that the ‘potential’ properties of *-ùʔ* and *-yùʔ* are not made use of, if in fact they existed in that Korku dialect. If we take *-yùʔ* as derived from {-*ya-ùʔ*-}, the general ‘movement away from’ (in time or space) is only in fanciful ways useful in characterizing the difference between the two.

Note that for Nagaraja in a number of cases in the PRS/FUT for intransitive verbs, *V-ùʔ* is translated as ‘will V’ where *-yùʔ* is (not systematically) translated more ‘potentially’ with the meaning ‘can V, may V’ also with transitives */ùʔ/* ‘will be V-ed’ vs. */-yùʔ/*, ‘may be V-ed’, ‘is to be V-ed’.

The element */-khùʔ/*, on the other hand, is functionally quite distinct. For Zide and Girard it is a durative, occurring usually with verbs where the action results in a state that can be continued, or statives, for example, *subankhùʔ*, ‘remain seated’ < *suban* ‘sit’; cf. *susùban* ‘to seat, to make/get someone (to) sit’; *katankhùʔ* ‘remain quiet’ < *katan*, ‘be(come) quiet, shut up’. Nagaraja has several examples of */-khùʔ/* with the continuous/progressive auxiliary */lak-/* (from *lab-*; the *lab-* is preserved in Drake, but has assimilated to *laj-* and *lak-* in Girard, Zide, and Nagaraja) which have durative meaning,

- (32) *jamlakkhùba*
 {*jam-lak-khù-bà*},
 weep-PROG-DUR-FIN:PRS/FUT
 ‘will keep weeping’ (Nagaraja 1999)

But */jamkhùbal/* is translated by him as ‘would weep’, not a meaning found in Zide and Girard. A few other examples with */-khùʔ/* he calls optative and translates (roughly) ‘may you Verb’ and ‘let me Verb’ are not found in Drake.

Korku also recognizes a number of modal formations in its basic verb structure. The probabilitive modal suffix */-kil/* (or *ki₂* so as not to confuse it with the intensive or recent past Mode suffix, among other morphemes) occurs in the PRS/FUT as the final morpheme before */-bà/* with the meaning ‘probably’. It is more common with transitive verbs but can occur with intransitive stems as well.

As alluded to above Drake has */-ye/* as a potential or capabilitive mood marker.

- (33) *dij tol-ye-bà*
 He bind-CAP/POT-FIN:PRS/FUT
 ‘He can bind’ or ‘He will be able to bind.’ (Drake 1903)

Drake also has the ‘permissive’ prefix *wa-* which similarly was not found by the other investigators of Korku.

- (34) *dich ing-ken wa-shen dan*
 he I-ACC PRMSSV-GO PST
 ‘He was letting me/permitting me to go.’ (Drake 1903)

A yes/no question in Korku is marked by another *-ki* (*-ki₃*); this is the final morpheme of a verb form in the interrogative and not part of its internal structure. It really functions as a clitic and probably is a loan element.

- (35) *am sene-bà-ki*
 you GO-FIN:PRS/FUT-Q
 ‘Are you going/will you go?’

A confirmative construction is found (only?) in the Lahi (Hoshangabad) dialect of Korku and not in any of the published materials. Here the final pre-*bà* syllable – the penultimate syllable of the verb form is repeated in a reply to a sentence proposing or asking something where the respondent agrees or confirms what is proposed, asked, stated. The syllable is the pronominal Object suffix if it begins with a consonant, but if the Object begins with a vowel (*/-ej/*, 3SG. or */ijn/*1SG.) the syllable consists of Mode-Object (note that *-ki₂* cannot occur in these forms).

- (36) *am dij-khè? kul-khèj-o-ki {kul-ki-èj-ò-ki₃}* (Lahi has */-ò/* for */-bà/*)
 you he-ACC send-INTNSV:3ANIM-FIN:PRS/FUT-Q
 ‘Will you send him/her?’

- (37) Confirmative answer:
ha?ā, kul-khèj-khej-o
 yes send-INTNSV:3ANIM-INTNSV:3ANIM-FIN:PRS/FUT
 ‘Yes, I’ll send him.’

The imperative of an intransitive verb is marked by the suffix *-e*; those forms derived from *U*-infinitives take */-ù?/* imperatives. From *suban* ‘to sit’ we can form *subāe* {*suban-e*}, ‘sit (down)’ and also *subāù?* {*suban-ù?*}. The latter, a more polite form, means, literally, ‘be seated’. The usual transitive imperative can take Object suffixes, but not commonly Mode ones. The negatives – see section 3.2.9 – have the prohibitive copula */bakil* (from */ban-kil*) with an infinitive, usually REDPL-Verb, Verb, or for irregular intransitive, for example, */hej-l*, ‘to come’ (where the infinitive takes the ‘extension’-*e*), Verb-*e*, that is, ‘don’t come’ is *baki heje* Verb-*e*. For ‘go’ in the positive *sen-e*, a politer imperative occurs with the potential */yù?/*.

There are defective imperative stems, verbs that just occur in the imperative, for example, */bol*, ‘come on (let’s go)!’ */nal*, ‘take it!’.

- | | |
|--|--|
| (38) Positive | Negative |
| INTRANS Verb- <i>e</i> ; Verb -(<i>ya-</i> , <i>-ki-</i>)- <i>ù?</i> | <i>ba-ki</i> Verb; <i>ba-ki</i> Verb - <i>ù?</i> |
| TRANS Verb-(Mode)-Object | <i>ba-ki</i> REDPL-Verb |
| | <i>ba-ki</i> Verb-(Mode)-Object |

Girard writes that the imperative forms also serve as subjunctives, but with no examples of either use in the Appendix to her dictionary. Drake finds the ‘imperative forms with overt “subject” for example and an invariant verb form’ (as above).

- (39) *dij biq-e* or *dij bi-biq*
 He rise-IMP he REDPL-rise
 'He may/should rise/get up.'

The transitive meaning of */bi-biq/* would be made explicit by adding an object to the REDPL-Verb form:

- (40) *dij poera-khè? bi-biq*
 He boy-ACC REDPL-get.up
 'He should/may get the boy up.'

The emphatics */kà?~kà/* and */-sè/* (*-sen* in Drake) can give 'tense/aspect meaning', but the subtleties of such meaning have not been explored. *dij olen-kà?* (or */-sè/*) would be translated by Drake as 'he has gone' (with tense/aspect meaning) where Zide and Girard would translate (with emphatic meaning) 'He DID go/ he certainly went/ indeed he went'. Certainly these emphatics 'have increasingly taken on verbal value', but to describe what could be meant by this – with numerous examples – is more than we can do here. */-kà?/* is commonly used with nouns as well; it is the usual translation of Hindi */-hiil/*.

Apart from the complex of elements clustering around passive mentioned above, the other voice morphemes in Korku include a productive reciprocal and a lexicalized causative, both with parallels in other Munda languages (Anderson 2007). A bilabial reciprocal morpheme is old in Munda. It is */-p-/* in Korku. Examples: */juù?/*, 'to give', */jupui?/*, 'to exchange, give each other'; */müðal/*, 'to hit, beat' */mupüðal/*, 'to fight, beat each other'. These are transitive verbs formed from simple transitive verbs. They take no Mode form or object agreement.

The */a-/* causative is old in Munda, but like other prefixes gone or almost in Korku. In fact, there are only two verb-stems that take *a-* (compare the richer set of such forms elsewhere in North Munda in Mundari). These are *{a-jom}*, 'to feed, give to eat', from */jom/*, 'to eat', and *{a-nu}* 'to give to drink', from */nul/*, 'to drink'. The preservation probably owes something to the very common (and somewhat irregular) causatives of 'eat' and 'drink' in Hindi */khilaanaal/* from */khaanaal/* 'to eat', and */pilaanaal/* from */piinaal/*, 'to drink'.

The intransitive past forms with *-Mode-en* neutralize transitivity, that is, *{goj-ya-en}* > */go?en/* and *{goj-ki-en}* > */gojken/* mean both 'died' and 'was killed' as finite verb forms and as non-finite adjectives. *-en-* occurs after the Mode suffixes *-wa* and *-li*, but not commonly. There is an archaic adjectival form using the simple stem (e.g. */goj/*, 'dead', */dejl/*, 'broken') for only a small number of verbs where a state with no history of agent or action (dying or killing) is implied. Thus, *goj sim*, 'dead chicken' contrasts with *go?en sim* and *gojken sim*, both for – more literally – 'died chicken' and 'killed chicken'. The difference between these two – the difference indicated by the Mode suffix */-ya-/* in one and */-ki-/* in the other is of relative pastness: */gojken/* for recent past ('just died, been killed') and */go?en/* for remoter past. Nagaraja points out that in some cases the */-ki-/* form also indicates that the past action has begun but is not complete. There are a number of allomorphs of */-ya-en/*, the common ones */-nen/*, and */-jen/* being simply phonologically conditioned, the first occurring after stems with final nasal, and the second after stems with final glottalized consonants *-b* and *q*, while glottalized *-j-y-* > *ʔ*: *{goj-ya-en}* > *go?en*. After stem-final vowels various glides are found in Nagaraja: *-a(y)en*, *-i(y)en*, *-o(w)en* and *-u(y)en*.

Pinnow talks of several kinds of reduplication, full and partial ((C)V-). Korku lacks the iterative CV-reduplication found in some South Munda languages, and has no full syllable or full stem reduplication in the verb-stem. Some apparently simple stems are partially – CV-reduplicated in form, and may derive from an old reduplication function (e.g. *lraraml*, ‘to feel cold’). Full stem copies in the form of repeated, for example, *lsene senel*, ‘going along’ – non-finite – participial forms are common as elsewhere in India for continuing action. In onomatopoeic verbs (with limited morphology) the whole first piece is repeated, for example, *lkeʔej+keʔejl*, ‘to clatter’.

The usual – and very common – reduplication (discussed above in presenting the simple tense forms for verbs, transitive and intransitive) is used in forming the transitive infinitive – ‘customary’ when used as a verb-stem, for example, *ljojombàl*, ‘eats’ (customarily), ‘will eat’.

Class-I intransitives can be reduplicated to give transitive/causative meaning, for example, in the examples above: *lsusùbanl* from *lsubanl*, ‘to sit’; *li-iraʔl* to return (something) from *liraʔl*, ‘to return, come back (ITR)’. Class-II verbs are basic transitives where no same stem Class-I intransitive is found, for example, *lgogojl*, ‘to kill’, *lqedej*, ‘to break’. Nagaraja claims that ‘grammatically reduplicated’ verbs with initial vowels are homonymous with their unreduplicated forms. He is mistaken. He has-hears- *liraʔl* and *liiraʔl* as the same. They are not the same.

3.2.9 Negation

Nagaraja’s Korku materials exhibit more differences from Girard and Zide here than almost anywhere else in his grammar. Of particular interest is the surprising retention of the present negative *ban* in Past sentences, so that the sentence has two negatives, the semantically appropriate Past Negative *-dìm*, and the no longer semantically appropriate *ban*. For Girard/Zide, the negative copula used in the indicative, that is, the present/future *ban ~ baj*; the emphatic *bannèʔ* is usually found only in sentence-final position. The *l-èʔl* in the latter form is probably to be identified with the *l-èʔl* in the exclamatory locative verbs (see above), not with the Past Transitive *l-èʔl*. The emphatic prefix *lhe-l* can be used to form *he-ban*, *he-bannèʔ*. Nagaraja’s speakers use *lhe-l* forms more commonly than Girard’s and Zide’s. The negator *ban* can occur either before or after the verb it is negating. A monosyllabic lexical verb appears in its reduplicated stem allomorph.

(41) *dì-ku jikʔa-àʔ jilu jo-jom (he)bannèʔ*
 He-PL porcupine-GEN meat REDPL-eat NEG.COP
 ‘They don’t eat porcupine(-s) meat.’

or

dì-ku jikʔa-àʔ jilu (he)ban jo-jom
 He-PL porcupine-GEN meat NEG.COP REDPL-eat
 ‘They don’t eat porcupine(-s) meat.’

or

dì-ku jikʔa-àʔ jilu jo-jom (he-)ban
 He-PL porcupine-GEN meat REDPL-eat NEG.COP
 ‘They don’t eat porcupine(-s) meat.’

In the Past (replacing tense markers */-en/* or */-èʔ/*) the enclitic – *qùn* is used which follows a bare verb-stem, with no ‘Mode’ or Object suffix.

- (42) *qiku jikɾa-àʔ jilu jom-qùn*
 He-PL porcupine-GEN meat eat-NEG.PST
 ‘They didn’t eat porcupine meat.’

qi-ku am-kheʔ qoʔ-qùn
 He-PL you-ACC see-NEG.PST
 ‘They didn’t see you.’

qi-ku iraʔ-qùn
 He-PL return-NEG.PST
 ‘They didn’t return.’

Since the stem occurs with no Mode or Object or Reduplication, the negative of the reduplicated transitive/causative is also *-qùn*.

- (43) *qij kitaabo iraʔ-qùn*
 she book return-NEG.PST
 ‘She didn’t return the book.’

In the prohibitive *ban* is found with the intensive Mode suffix, that is, */ban-kil > baki*. This is used with a reduced set of Present/Future infinitive verb forms. The various transitive infinitives are comparatively rare in these constructions, the Mode suffix */-kil/*, the intensive, is more common than the others (e.g. the directionals, i.e. */-li/* and */-ya/*). The intransitive infinitives are */V-ùʔ/* (for all three positive intransitives: those with *V-ùʔ*, *V-ya-uʔ* and *V-ki-ùʔ/ V*, and for *V-e*, *V-e*).

- (44) *baki oq*
 PROHIB go.out
 ‘Don’t go out!’

baki oòq
 PROHIB REDPL:take.out
 ‘Don’t take (something) out!’

baki oq-ej
 PROHIB take.out-3
 ‘Don’t take him/her out!’

baki kul-kh-èj
 PROHIB send-INTNSV-3
 ‘Don’t send him/her.’

baki sene
 PROHIB go:INF
 ‘Don’t go!’

baki kul-li-n {kul-li-ijn}
 PROHIB send-CLOC-1
 ‘Don’t send me (here, this way, in this direction).’

Nagaraja has apart from the distinctive double negatives other differences in his negative forms: he writes *baw* (not found in Drake, Girard, or Zide) for negated propositions. It occurs in verbless constructions, for example, *inija batɛ baw* ‘this

boy has no father' (Girard/Zide have *he-ban*, *he-baj*, *ban*, (*he*)*bannè?*). This distinction – a special Negative Copula in these cases probably owes something to the reorganization of the system following on the innovating – borrowing – of a positive copula *hoy*, not known to Drake's, Girard's and Zide's speakers which would be used in the positive of the above sentence. The other grammars have nothing – a zero copula – in Nagaraja's *hoy* position.

The earlier use in Past sentences of forms with */ban/* implied no Present/Future meaning. It was only after the later introduction of */-dùn/* and the tense specialization in some dialects, that is, */ban/* only with Present/Future negation, and */-dùn/* only with Past – that the double negation would be felt as strange and contradictory by speakers of Korku dialects who were familiar only with a completely regularized tense specialization; that is, */ban/* is always used in Present/Future forms and derivatives, and */-dùn/* is always used with Past forms.

- (45) *bhuri siṅrup heba ira-dun-ka*
 bird evening NEG return-NEG.PST-EMPH
 'The bird did not return in the evening.' (Nagaraja 1999)

in amen heba ji-dun-ka
 I you-LOC NEG give-NEG.PST-EMPH
 'I had not given you.' (cf. *in am-èn he-ban ji-dùn-ka?*), (Nagaraja 1999)

The form *aṭikà?* (Girard *aṭikhha*, Nagaraja *aṭhika*) means 'not yet' and is used with a past verb form. The expression of 'not-yet-ness' is common in the Munda languages. Gutob has two contrasting negative copulas: *lura?*, 'to not be', and *lorojl* 'to not yet be'. Note also the negative verb form *baṅgon*, presumably some sort of derivative of *baṅ*. The verb means 'to not do', 'to refuse to do'.

- (46) *urag-àten oḍe! in baṅgon,*
 house-ABL come.out:IMP I NEG.VERB
 'Come out of the house!' 'I won't', 'I refuse to'.

A partial paradigm in the positive/affirmative and negative of the verb *tol* 'tie' is offered in the following example.

- | | | |
|---------|--|--|
| (47) | Positive | Negative |
| PRS/FUT | <i>toṭolbà, tolejbà</i>
'ties, will tie' | <i>toṭol ban(n)-è?</i>
<i>toṭol ban(n)-è?</i> |
| | <i>tolùba</i> 'will be tied' | <i>tolù? bannè?</i> |
| | <i>totoldṅan</i>
'used to tie, has tied' | <i>totol bannèḍan</i> |
| | 'formerly tied' | <i>tolù? ban[n]èḍan</i> |
| PST | <i>tolken</i> 'was tied' | <i>toldùn</i> |
| | <i>tolkhènej</i>
'tied something/someone' | <i>toldùn</i> |
| ITR | <i>tolkenḍan</i> 'had been tied' | <i>toldùnḍan</i> |
| TR | <i>tolkhènejḍan</i> 'had tied' | <i>toldùnḍan</i> |

3.2.10 Derivation

The many verbs derived from Indo-Aryan – in one common derivational pattern – add *-atij* to the source language stem, usually the transitive stem, sometimes the intransitive, and sometimes for a few stems – both (with no apparent difference in the Korku meaning).

For example, *kholna*: (I use the standard Hindi forms, dialectal forms for these purposes are much the same) ‘to open’ (TR) and *khulna*: ‘to open (ITR)’ yield Korku *kholatij-* and *khulatij-*, for some speakers meaning the same thing. This is not a recent borrowing. Such forms are common in Drake. There are dozens of examples, and nonce forms with *l-atijl* also occur. Some examples of it use include the following:

lk(h)ukhùlatij-k(h)okhòlatijl;lkhulatijebàl, ‘opens it, will open it’, *lkhulatijùbal*, ‘is opened’, ‘will be opened’, *lkhulatij(n)ùbal* ‘may be opened’ {*khul-atij-ya-ù?-bà*}, *khulatijnen*, *khulatijken* (past participles), etc. Even onomatopoeic stems (apparently borrowed), can take *l-atijl* for example, *phaṛ+phaṛ-* ‘to flutter, quiver’ > *lphaṛ+phaṛatijl*.

Note the full syllable repetition is onomatopoeia, not the usual CV-reduplication.

Some Hindi verb-stems are borrowed directly – with no additions – as Korku stems (with minor phonological adjustments and reduplicated), particularly when they are derived from nouns (Nagaraja) *khyal*, ‘thought’ > *kikhyal-* ‘to think (about)’, *muṭi* a handful → *mumuṭi* ‘to grasp’. In some cases transitive uses are unexpected, for example, *kunum*, ‘urine’ (not borrowed). That the verb *kunum-*, ‘to urinate’ would take an inanimate object {*kunum-e-bà*} was expected, but that *kunum-ej-bà* occurred was not. This turned out to mean ‘urinates on someone’.

A common – rough – observation about the Munda languages, particularly the North Munda languages, is that anything can be verbalized; any nonderivational or inflectional or other special functional morphemes, can be a verb-stem. A nice example of that comes in comparing the two early Santali dictionaries, that of Bodding and the other one of Campbell. Campbell defines the first dictionary entry simply as ‘the grunt of a water buffalo’, an interjection. Bodding, scrupulous about such matters, classifies it not only as an interjection but as a verb(stem) – and more. Certainly many – most – nouns can serve as verb-stems. To list some from Drake (many of his nouns are loanwords), the nouns and the verbs and their meanings (one can argue the direction of the derivation, but for many of these, the noun form is primary):

aṭkom, egg, as Verb ‘to lay an egg’, *koyo[?]* ‘wind, to blow of the wind’ *jo*, ‘fruit’, V, ‘to yield fruit’; *chaaso* (loan), ‘breath’, to sigh; *luṭo*, (loan) loot, ‘to loot’ (but note that *luṭnaa* is a verb in Hindi). The meaning of the derived verb is usually fairly predictable.

Certain adverbs are verbalized, for example, *lcuphàrl*, ‘how, in what way’, from ‘interrogative-manner’. In all four grammars ‘to how’ is to do what was referred in the previous sentence ‘in what way’. For example, ‘to yoke a bullock’ is the action in the previous sentence, *dij cuphàrej* ‘how will he (yoke) him’. Similarly {*cuphàr-e*}, with an inanimate ‘object’.¹ Also like these are: *i-phìn-eli-phìn-ej* from *i-phìn* < *lin-phìnl*, ‘this-manner’: ‘to do (the action) in this way’; *qèèn*, ‘there’ < {*dì-èn*}, ‘that-loc’; as a verb (Drake) ‘to get there’. For Drake ‘to get there’ and ‘to become how’ are inceptive (‘reponent’) verbs. Their intransitives for Drake are formed with *l-yul*

{*ya-ù?*}); such Verb-*yu* forms have as infinitives *V-yù?*; the corresponding transitive infinitives take *-el*, for example, */cuphàrel*, etc. Verbs derived from adjectives for Drake – it is not clear that the distinction Adjective vs. Adverb is made in Drake’s material – also have *V-yu* infinitives in the intransitive and *V-e* infinitives in the transitive.

3.2.11 Noun incorporation and combining forms

Noun incorporation (and the use of combining forms of nouns) such as is found in various South Munda languages either productively or in a lexicalized manner (Anderson 2007) is not attested in Korku.

3.2.12 Auxiliary verb constructions

Many – if not most – of the verbal forms functioning as auxiliaries in Korku (like the postpositions) are borrowed, and their use is different in the various grammatical sources. Some are found in only one source. The auxiliaries in this section supplement those in the Negation section 3.2.9 (*-qìn* and (*he*)*ban/bannè?*) and the section on Tense/aspect (sections 3.2.3 and 3.2.4 (*-qàn* and *//lab*)).

The Korku auxiliary usually follows an infinitive form of the lexical verb which is formally reduplicated for certain stems. With the auxiliary *//lab-/* it may – dialects differ – follow Verb-Object and, more rarely, Verb-Mode-Object. Following a bare verb-stem is the Past Negative AUX *-qìn*, which neutralizes all Mode and Object distinctions on the verb elsewhere, for example, those in preceding sentences with common reference. Also, the bare verb-stem form is used with the permissive prefix *wa-*, only by Drake who states ‘the principal verb in its shortest possible form is found’.

No semantic characterizations of auxiliaries are offered although modal and causative notions are central. A number of the notions that are expressed in Indo-Aryan with auxiliaries – and that are obviously borrowed from Indo-Aryan into Korku, for example, the cislocative and translocative – are part of – internal to – ‘the basic verb form’ in Korku (e.g. cislocative (*-li*), translocative (*-ya-*), intensive (*-ki-*), delayed action (*-jom-*), and tentative (*-thà-*)).

Korku and probably most of the other Munda languages must, in earlier – pre-borrowing – times, have used various discourse features to convey some of these auxiliary – for example, modal – meanings. Most of the auxiliaries are borrowed. Unfortunately we do not have data on the eastern dialect(s) of Korku, Mowasi/Mawasi, to confirm the suspicion that they would exhibit a rather different set of auxiliaries.

The various auxiliaries are presented below with brief descriptions and examples. There are a few interesting differences in usage (e.g. concerning *//lab-/*). One could sort the auxiliaries by meaning, grouping together the different grammars’ auxiliaries (and non-auxiliary machinery), for example, permissive, causative, potential but that will not be done (explicitly) here. The auxiliaries to be discussed here (the descriptive labels/translations are mostly those used in the original sources):

- | | | | |
|------|------------------|---------------------------------|------------------------------|
| (48) | <i>tekk-</i> | (Drake, Girard, Zide, Nagaraja) | ‘to want to V’ |
| | <i>hona</i> | (Girard, Zide, Nagaraja) | ‘to have to V’ (obligative) |
| | <i>lag-atij-</i> | (Drake) | ‘to have to/be obliged to V’ |
| | <i>hola?</i> | (Girard, Zide, Nagaraja) | ‘to cause/force to V’ |

<i>ghaṭa-</i>	(Drake)	‘to manage/find a way to V’
<i>-ni bhala</i>	(Nagaraja)	‘should have Ved’
<i>-len</i>	(Nagaraja)	‘to (V) for someone else’
<i>-lab-</i>	(Drake, Girard, Zide, Nagaraja)	(a) progressive/continuous
	(b) inceptive	
	(c) and more (see below)	
<i>ḍa-</i>	(Drake, Nagaraja)	inceptive
	(Nagaraja)	‘to be about to V’
	(Drake, Girard, Zide, Nagaraja)	permissive
	(Nagaraja)	‘would V’

With regard to *tekko* ‘to want to’ (Drake *taku*), Nagaraja has it in his lexicon as a noun meaning ‘desire’, but without examples of its use. In Girard and Zide’s materials it takes an infinitive in the genitive.

- (49) *dij jilu jojom-à tekko-bà*
 he meat REDPL:eat-GEN AUX-FIN:PRES/FUT
 ‘He/she wants/will want to eat (the) meat.’

In Drake, along with the genitive construction, he finds a simple genitiveless infinitive. Thus he finds both.

- (50) *hadiru taku-ba* and *hadiru-ga taku-ba*
 arrive AUX-FIN:PRES/FUT arrive-GEN AUX-FIN:PRES/FUT
 ‘wants to (sometimes “intends to”) arrive’ (cf. Zide {(h)adīr-ù?-[à?]}
tekko-bà})

The auxiliary *hona* ‘to have to’ appears in an invariant form without *-bà*. The usage is borrowed from regional, not Standard Hindi. The form is lacking in Drake. It is used with the simple infinitive. In Girard’s, Zide’s, and Nagaraja’s texts, the ‘logical subject’ is in the nominative. (One might have expected from the equivalent structures in Standard Hindi and elsewhere in the region for it to be in the accusative, but it is not.)

- (51) *ḍiig^halya mumūḍāmit^haj ḍijà bagolten*
ḍi-e-g^halya mu-mūḍā-mi-t^hā-(e)j ḍij-à? bagol-ten
 that-INAN-REASON REDPL:beat-one-NMLZR-3SG.ANIM he/she-GEN place-ABL
- saṛub hona ḍo ḍijk^hè? mumūḍà hona*
saṛub hona ḍo ḍij-k^hè? mu-mūḍà hona
 to run to have to and him/her-ACC REDPL:beat/hit have.to
 ‘For that reason the hunter (lit. beater) has to run at the side/alongside him (the porcupine) to beat him.’

Drake has *lagatij-ya-en* and *lagatij-yu-bà* as an ‘obligative’, ‘to have to’, ‘to be obliged to’. Presumably this derives from Hindi *lag-naa*.

The auxiliary *hola?* (Girard *hoṛalhora*, Nagaraja *holā*; not found in Drake) is used with the infinitive form of the lexical verb.

- (52) *ij gaṛi cacalu hola?-kh-en-ej*
 I car REDPL:MOVE AUX-INTSNV-PST.TR-3
 ‘I started the car.’ (Nagaraja 1999)

This is an example of */-ej/* the third singular animate object marker used in Nagaraja more widely (as it is not in Girard and Zide), to refer to inanimate objects. This would be a ‘causative’ of the intransitive, that is, ‘I caused the car to move.’

- (53) *ij dij-khè? cuciri hola?-kh-en-ej*
 I he-ACC REDPL:steal AUX-INTNSV-PST.ITR-3ANIM
 ‘I caused/forced him to steal.’

As noted above the old Munda causative prefix *-/a-/* in Korku – is found only with *jom-*, ‘to eat’, and *nu-*, ‘to drink’. Causatives are commonly formed syntactically by taking a transitive sentence and adding the agent of the action caused – marked by the ablative, with no auxiliary.

- (54) *dij dadu-ten sita-khè? tol-kh-èn-ej*
 She Dadu-ABL dog-ACC tie-INTNSV-PST.TR-3ANIM
 ‘She had Dadu tie the dog.’

With *hola?* one could have both *Dadu* and *sita* in the accusative, or *Dadu* in the ablative. A ‘second causative’ – à la Hindi – for *a-nu* and *a-jom* is possible to form by treating the */a-/* causative, for example, *ajom*, ‘feed’ as a simple transitive – to translate ‘I had Dadu feed the child’, or with *hola?* ‘I forced Dadu to feed the child’, etc.

Drake finds an ‘acquisitive’ *ghaṭa(-)u*, ‘to manage to, to find a way to’ that is lacking as an auxiliary in other sources. Note that *ghaṭa* here takes not a simple infinitive but Verb-Object form.

- (55) *ḍoba?-khè? tol-ej dij-khè? ghaṭa-ḍin*
 Bullock-ACC tie-3.ANIM he-ACC AUX-NEG.PST
 ‘He did not manage to tie the bullock.’ (Drake 1903)

Nagaraja finds another borrowed auxiliary construction: Verb-*e-nibhala* meaning ‘should have’. The auxiliary *-nibhala* follows the augmented infinitive form in *V-e-*.

- (56) *paṛaṭij-e-nibhala saṛub-e-nibhala*
 study-INF-AUX run-INF-AUX
 ‘should have studied’ ‘should have run’ (Nagaraja 1999)

Another auxiliary element is the benefactive mood form in *-le[n]*. It is found only in the imperative. It marks action performed for benefit of another, for example, *ol-ki* ‘write!’ vs. *ol-ki-len*, ‘write for someone else’s sake’ (Nagaraja 1999).

As discussed in the section on tense/aspect etc. above (sections 3.2.3–3.2.8), one of the most common auxiliary elements in Korku is the (now fused) construction involving an element that appears to be underlyingly */lab-/* though it now is almost never realized in this form. The progressive is formed with the auxiliary *lab-*, preserved as *lab-* in Drake, but with the */b/* assimilated to the following */j/* or */k/* in the other three grammars. There is some disagreement about the reduced set of Verb-Mode-Object forms that occurs with */lab-/* progressive forms. Girard says */lab-/* – independently – means ‘to begin’. It for Zide is a progressive and not an inceptive, but it has inceptive meaning too elsewhere.

The forms are REDPL-Verb-*lab-* or Verb-Object-*lab-*. Almost none of the *lab-* forms have Mode suffixes. The forms that are realized include *-lakken* *{-lab-ki-en}* ‘is V-ing, is in the act of V-ing’ also with inceptive meaning ‘beginning to V’ in some areas. Examples include *jojomlakken*, *jomelakken* ‘is in the act of eating’. Other forms include the realization *-lajjen* *{-lab-ya-en}* ‘was V-ing, was in the act of V-ing’,

for example, *jojomlajjen*, ‘was in the act of eating’ and less common forms such as *jojom-lakkhùba* {-lab-ki-ù?-bà}, ‘go on’ – or ‘will go on’ – (being in the act of) eating’ and *jojom-lajjùba* ‘went on (being in the act of) eating’, also with ‘potential’ meaning, ‘went on being able to continue eating’.

As alluded to above, for Drake, Girard, and Nagaraja //lab-/ has both progressive and inceptive meaning. For Zide it is only progressive in the *-lakken* and *-lajjen* forms. For Zide – and, Girard (Drake) and Nagaraja *-jojomlakken //labken* means ‘is eating’ and, for Zide, *ljojomlajjen/* – {-lab-ya-en} means ‘was eating’, whereas for the others it has primarily an inceptive meaning, ‘began to eat’. Drake is more nuanced in finding both meanings and trying to ascertain which one predominates (where). Nagaraja has the emphatic *l-ka?* commonly in *l-labl* sentences with progressive meaning, that is, in */Verb:INF-ka?-lab-/* structures (see Drake and Nagaraja for details). Some examples of somewhat unexpected //lab-/ uses in Nagaraja include:

- (57) *kama:y-ku laj-j-en*
 {*kama:y-ku-lab-ya-en*}
 WORK-3PL-AUX-MODE-PST.ITR
 ‘It started to work (by many people).’ (Nagaraja 1999)

kama:y-mi-laj-ju-ba
 {*kama:y-mi-lab-yu-ba*}
 WORK-1/2-AUX-MODE-FIN:PRS/FUT
 ‘It starts/will start to work – by 1SG. or 2SG. Person.’ (Nagaraja 1999)

I am not sure what Nagaraja’s English translations mean. The pronominal marking on the verb, it seems, marks Subject, not Object (see also sections 3.2.1–3.2.2)!

Drake has *ēḍa-* as in inceptive much like //lab-/

- (58) *ura? jul-u? ēḍa-u-ba*
 house burn-ù? begin-ù?-FIN:PRS/FUT
 ‘The house will begin to burn.’ (Drake 1903)

This *ēḍa-* ‘to apply, attach (to)’ apparently does not occur as an auxiliary in the other grammatical sources. The equivalent form with //lab-/ would be {*lab-yù-ba*}.

One of the most varied and multifunctional elements in Korku, and one which differs the most among the various descriptions is the element *ḍa* –with, among other meanings–‘about to’. It may be used with a range of lexical verb forms (\pm reduplicated, etc.): *olkendàn*, ‘was about to write/wrote’; *nunukendàn*, ‘was about to drink/drank’.

Nagaraja’s material suggests an interesting difference between his Korku and the Korku of Girard and Zide (and presumably Drake): that for Nagaraja’s speakers the action – with */V-ki-en/* begins with an ‘about to’ phase and continues with the start of the performance of ‘the action proper’ which may or may not be completed. This contrasts with the */V-ya-en/* forms which describe remoter past actions with the action complete. For the other sources the action in Verb-*ken* forms begins with the start of the ‘performance of the act proper’ (no ‘about to’ phase) and is usually completed. For these speakers too */V-ya-en/* forms refer to remoter – completed – past actions.

His second ‘about to’ usage – with plain *ḍa* – has the element suffixed to the verb root directly, the ‘root’ in his three examples is either a transitive verb with

reduplicated infinitive *jojom-qa*, ‘is/was about to eat’, or Verb-Object-*qa*, *mūqà-mi-qa*, ‘is/was about to beat me’. (Note: for Nagaraja *l-mil* is commonly used for both first and second persons whereas for other researchers it is used only for second.) If *lqal* is the verb *lqal* – which I think it must be – this is a very odd construction with no Tense or Mode suffix on either *ljojoml* or *lqal*. Is *lqal* a verb-stem functioning as a kind of ‘Second Predicate’, contrasting with *l-bàl* in these structures? More research is required to make this clearer.

The element *l-qànl*, presumably from {*qa-en*}, where *l-enl* is the Past Intransitive. Thus *-qàn* – like *lolenl*, ‘went’, and *lmhenan~mhenen~mhenon~menanll*, ‘said’, a narrative introducer, is one of the very few old intransitive past forms lacking a Mode suffix. *-qàn* figures in the basic tense/aspect system occurring with both PRS/FUT and Past forms, positive and negative. Drake gives *lqaenl* as the Past where *qa-* is the main verb, and *qan* is found in functional compounds.

One also finds *qa* in Girard’s Dictionary *qada* ‘to do’, ‘*qa-ù?*’ ‘to be done’, ‘to happen’ (also ‘become’). This is an auxiliary used in forming permissives.

The ‘permissive’ for Nagaraja uses *lqawl* (presumably from *lqau?* elsewhere). It follows the infinitive, and the auxiliary takes Mode and Object suffixes.

- (59) *ijn dij-khe? sen-qaw-khe(n)-ej*
 I he-ACC go-AUX-INTNVS-TR.PST-.3SG.ANIM
 ‘I allowed him to go.’ (Nagaraja 1999)

naaka dij sen-qaw-ej
 now he go-AUX-3.SG.ANIM
 ‘Now, let him go’, that is, ‘he may go, is permitted to go’. (Nagaraja 1999)

Note that ‘He’ (*dij*) is in the nominative in the second example, that is, this is a ‘subjunctive’ but the auxiliary *lqawl* takes a 3.SG.ANIM marker, which presumably marks subject! In the first example *lqawl* is not the main verb, but an auxiliary with verb *lsenl*, ‘to go’.

For the most part, Drake attests the same structures as described above (from Nagaraja), but with some differences. Drake’s speakers can use the *l-e-* as well as the (passive-potential) *l-ù?* before *l-bàl*. With the *l-ù?* forms the ablative *l-àtenl* is used for the ‘logical subject’; with the *l-e-* forms, the nominative:

- (60) *ijn sen-e qa-e-ba*
 I go-AUGM/INF AUX-TR-FIN:PRS/FUT
 ‘Let me go.’ (Drake 1903)
- ijn-aten sen-e qa-u-ba*
 I-ABL go-AUGM/INF AUX-PSSV/ITR-FIN:PRS/FUT
 ‘Let me go.’ (Drake 1903)

OR

ijn-aten sen-yu qa-u-ba
 I-ABL go-PSSV/ITR AUX-PSSV/ITR-FIN:PRS/FUT
 ‘Let me go.’ (Drake 1903)

With *lqa-wa-l* for the permissive, the alternative constructions are (61) the ACC object of ‘let-go’, *l-ijnl*, ‘me’ is not marked on the (ITR) verb, (62) the *qawa-* marks the accusative object.

(61) *dij in-khe? sen-e dawa-ba (or da(a)ba)*
 He I-ACC go-AUGM/INF AUX-FIN:PRS/FUT
 'He lets/will let me go.' (Drake 1903)

(62) *dij in-khe? sen-e daw-ijn-ba*
 He I-ACC go-AUGM/INF AUX-FIN:PRS/FUT
 'He lets/will let me go.' (Drake 1903)

A type of past tense meaning is conveyed by the suffix or (former) AUX in *-dàn* in Korku. Used with a PRS/FUT infinitive or a Mode-marked verb it commonly serves to mark a kind of imperfect or past habitual. The element *dàn* is also used in past copular constructions.

(63) *dj-ku ape-khè? dō-dò?-dàn dj-ku ape-khè? dō?-khè-pe-dàn*
 They you.PL-ACC REDPL-see-PST.AUX They you.PL-ACC see-INTNSV-2PL-PST.AUX
 'They used to see you (pl). They formerly saw you.'

3.3 Expressives

Not explicitly investigated for this study. Some brief comments on expressive-type copy forms in verb-stems and noun-stems are offered in sections 3.1 and 3.2.

4 SYNTAX

The syntax of Korku remains largely unstudied. One can infer some syntactic features from, for example, the texts in Nagaraja (1999), but I leave this to the interested reader. Only a few brief comments are offered here.

The introduction of the copula *hoy* (see Nagaraja), in the Melghat area or much of it, probably came some time in the last 50 years. Nagaraja finds it all over the place, but it is unknown to Drake, Girard, and Zide. As a result of a reorganization of copular predications following the introduction of */hoy/* all the adjectives in Nagaraja's Korku now work identically; none of them takes *-bà*; they all take the introduced copula */hoy/*, 'is, are'. In the present – not the Present/Future (this distinction found only here is also a result of the introduction of */hoy/*) – this homogenized construction is now like Hindi, from some variety of which */hoy/* must have been borrowed.

(64) *dji simil hoy [Zide=similba] dj-ku lāgra hoy [Zide=lāgra Ø]*
 3 sweet COP 3-PL lame COP
 'They (inanimate) are sweet.' 'They are lame.' (Nagaraja 1999)

The element *dàn* is used in past copular constructions. In Nagaraja's Korku the negative copula */baw/* is distinguished from */ban~baŋ/* and */ban(n)-è?/* which have similar meanings. He writes 'it occurs in verbless (non-locative) constructions after the object and is used to negate equational constructions as well.'

(65) *inij-èn ba-te? baw*
 DEM-LOC father-3 NEG.COP
 'He has no father.' (Nagaraja 1999)

4.1 The simple sentence

The constituent order in the Korku sentence is the one that is common in the Indian area, that is, SOV. An indirect object precedes the direct object in ditransitive sentences. In simple transitive sentences, the (human or animate, rarely inanimate, referential) direct object is marked with the accusative *l-khèʔl* (for Zide, elsewhere *l-ke ~ khe ~ ken.*). In ditransitive sentences the indirect (animate, usually human) object usually takes *l-khèʔl*, the direct object usually being inanimate. When both are human both can take *l-khèʔl*.

In informal discourse, the subject, understood from the previous sentence(s), is often deleted, but in the dominant languages of the area (e.g. Hindi) – and this is true of the other North Munda – that is, Kherwarian – languages as well, the subject is marked on the verb as it is not in Korku. This permits a simple subject-less construction (or unspecified subject), a sentence with a transitive verb with a ‘passive meaning’ to occur in Korku:²

- (66) *sita-khèʔ ʔol-kh-èn-ej*
 Dog-ACC tie-INTNSV-TR.PST-3.SG.ANIM
 ‘The dog was tied (down)/somebody tied the dog.’

With an explicit subject, the only difference in the form is the overt subject.

- (67) *Ratana sita-khèʔ ʔol-kh-èn-ej*
 Ratana dog-ACC tie-INTNSV-TR.PST-3.SG.ANIM
 ‘Ratana tied the dog down.’

The subject can sometimes, with a ‘resumptive’ intonation, come at the end of a sentence, again, as in Hindi. Objects too in Hindi in conversation are not uncommonly deleted (sometimes with deleted subject as well in the same sentence). This sort of object dropping is not common in Korku, although the object, marked on the verb, would be apparent.

As noted earlier (but not in Nagaraja’s material) there is/was no positive copula, so that sentences like ‘this house is mine’ lack a verb as in *ini uraʔ ʔn-àʔ* [this house I-GEN]. In Nagaraja’s material a borrowed copula verb */hoy/* would be used: *ini uraʔ ʔn-àʔ hoy* [this house I-GEN COP].

Another simple sentence construction common in the area and in most Munda languages is the ‘dative subject’ construction; this also is found in Korku with the locative in *l-èn/* encoding the ‘dative subject’, with roughly similar semantics to the parallel Hindi sentences.

- (68) *ʔn-èn raram-bà*
 I-LOC REDPL:feel.cold-FIN:PRS/FUT
 ‘I will be/feel cold.’

In some cases along with the above construction, a calque on the Hindi using the verb ‘come’ – lit. ‘hunger will come to him’ – is also used. See (69a) vs. (69b).

- (69) (a) *dij-èn rangejbà* or *rangejùba*
 he-LOC hungry-FIN:PRS/FUT hungry-PASS/ITR-FIN:PRS/FUT
 ‘He will be/feel hungry.’
- (b) *dij-èn rangej heʔibal {hej-ya-ùʔ-bà}*
 he-LOC hunger come-MODE-PASS/ITR-FIN:PRS/FUT
 ‘He will be/feel hungry.’

based on *lhej-l*, ‘to come’. Nagaraja gives an example of this sort of construction with a ‘complement’ (?) (not his term, he doesn’t say anything about the grammar of *lqaʔ* ‘water’) in a sentence with *ltaʔam*, ‘to be thirsty’:

- (70) *dij-èn qaʔ taʔam-nen*
 He-LOC REDPL:water thirst-MODE:ITR.PST
 ‘He/she is (for) water thirsty.’

Adjectives – including possessives – precede nouns, numerals follow adjectives and precede nouns

- (71) *dij-àʔ aphàì kende sim[-ku]*
 S/he-GEN 3 black chicken[PL]
 ‘her/his three black chickens’

Note that either a singular or plural form of the noun may follow a numeral in Korku (singular or dual for ‘two’).

Roughly, adverbs – the details for the semantic classes of adverbs are not given here – precede the verb phrase; some adverbs – for example, of time – commonly occur at the beginning of the sentence.

Although examples of simple sentences in Korku are numerous throughout this chapter, I offer some more in the following text.

- (72) *am je miàʔ gao dan*
 you who one village PST.COP
 ‘Who are you?’ ‘(once) there was a village’, literally ‘one village was’
- dij coj-ghalya en-èn-ej*
 he what-for this-LOC-3.SG.ANIM
 ‘Why is he here?’
- sebei tarei-ku sisirij-lak-k-en*
 all girl-ANIM.PL REDPL.Sing-PROG-PST-ITR
 ‘All the girls are singing.’

4.2 Complex sentences, relative-type clauses

Zide, in his (still unpublished) review of Nagaraja (1999) was trying to make observations on language change in regard to complex sentence structure in Korku, juxtaposing Drake (1903) with Nagaraja (1999) and Girard and Zide (roughly, 1950s to 1960s), comparing three possibilities: the (older) adjectival modifiers (e.g. *ʔolken sita* ‘the tied dog/the dog that was tied’, and the Indo-Aryan like – relative-and-correlative – structure, and – by far the most common Korku arrangement – two simple sentences. I went through ~50 pages of Girard’s unpublished texts and all of Nagaraja’s texts and found no ‘relative-type clauses’ of the first two types at all. Then how can we account for those of the first type found in all four of the grammars. Is the presence of relative-type clauses of the first two kinds an artifact of the elicitation (in Hindi or Marathi, where the relative–correlative construction would be used, but if a relative is encouraged the first type – not the relative–correlative – is usually preferred), or of the kinds of texts recorded, or of the kinds of informants worked with? Could it be that monolingual Korku speakers never used or use ‘relative-type’ clauses? It is true that ‘instant calquing’ – for

the most part falling back on existing adaptations of the ‘upper’ language, rather than nonce invention – is common in a lot of ‘lower languages’. Working with Gutob in Orissa (where a knowledge of the ‘upper’ language, Oriya, was not as common among the Gutob as Hindi is among the Korku) it was clear that one could show off one’s literacy, and presumably one’s intelligence, by a command of the upper language, and importing features of it into the lower language where it was thought some might find it – their native language – wanting, and thus ‘primitive’. Clearly, a different kind of field study of Korku usage is needed to understand what is going on. What discourse – or cultural(?) – features are promoted by being able to say not just ‘Baṭu’s daughter brought the mangoes’, but also ‘It was Baṭu’s daughter who brought the mangoes’? Are there, perhaps, subtler discourse resources in (earlier) Korku to handle some such features?

The adjectival relative-type clauses common, or at least mentioned, in the more recent Korku material, and apparently common earlier in Drake’s material are of the following structure. (The structure is typologically common, found, in India, also in Dravidian.) The simple sentence can be converted into a modifier by preposing it before the appropriate noun (formerly the subject) with minimal modification.

- (73) *tarei sita-khè? ʔol-kh-èn-ej*
 girl dog-ACC tie-INTNSV-PST.TR-3
 ‘The girl tied (down) the dog.’

transforms to ‘the girl who tied down the dog’ by preposing the verb phrase to the former subject noun:

- (74) *sita-khè? ʔol-kh-èn-ej tarei*
 dog-ACC tie-INTNSV-PST.TR-3 girl
 ‘The girl who tied down the dog’ literally ‘the tied down the dog girl.’

Similarly – and as simply – for almost the whole range of verb forms, for example, ‘the just tied dog’, ‘the not tied dog’, ‘the to-be-tied tomorrow dog’, the ‘they-were-tieing-him-down-dog’, etc.

The Indo-Aryan relative–correlative structure, roughly literally translated, gives for ‘the dog that she tied down’ as (relative) ‘which dog she tied down’ (correlative), ‘that dog ... verb phrase (e.g. ‘died’, ‘which house we lived in ...’ ‘that house (burned down)’, etc. This construction was not common in any of the Korku grammar, perhaps it is more common now. Earlier there was variation in the morphemes used. For Zide, *ljel* (literally ‘who’) for ‘which (Noun)’, *lqil*, for ‘that’ (Noun) for the correlative. Nagaraja has *ljol* (which is the Hindi form) where Zide has *ljel*, and *lqil* for the ‘that’ introducing the correlative phrase, such as the following example from Nagaraja (1999):

- (75) *jo kitaabo tebal-à-ljñ-en qòð-k-en*
 which book table-GEN-ON-LOC put/keep-INTNSV-ITR.PST
qii kitaabo iñ-à? hoy
 that book I-GEN COP
 ‘The book put/kept on the table is mine.’

Compound coordinative sentences are common (formed in a common way) with *lqol* ‘and’ (see the texts, for example, sentence 6 of the porcupine story), and (the borrowed) *lyal*, ‘or’. Disjunctive coordination, that is, ‘but’ (Zide/Girard *phini*, *pini*,

Nagaraja *phene, phire*) again occurs where, in many languages, it commonly does. An example from Nagaraja (1999):

- (76) *ini dukan-àten saman sa?-jom, phene qi dukan-àten*
 this shop-ABL goods take/bring-TENTATIVE, but that shop-ABL
baki sa-sà?
 don't REDPL-take
 'Buy goods from this shop, but don't buy from that shop.'

For the 'adversative' use, that is, 'nevertheless' – an example comes from the porcupine text:

- (77) *jikra-khè? cofo?-kà-bi jorten mūdà-ej*
 porcupine-ACC however-EMPH₁-EMPH₂ strongly hit-3.SG.ANIM
pini dij ban guj-ù?
 but he not kill-passive...
 '(If) you hit him however hard, but (*pini*) he doesn't die, then ...'

'If' – sentences are formed with *Itaasò(n) ~ taàsò(n)*l, 'if', at the end of the antecedent clause (following the verb). If the antecedent clause is PRS/FUT the predicator is usually absent, but with past antecedents the tense is marked. Note the consequent also has no predicator (*-bàl*). The following example comes from Nagaraja (1999).

- (78) *am heje-taaso in dij-khè? ban mumūdà*
 you come-if I he-ACC NEG REDPL:beat
 'If you come I won't beat him.'

Other 'if' sentences (again as in Hindi, with only the 'then' (Hindi *to*) introducing the consequent clause present, but no overt 'if'; with *makhàn* as 'then'. Example from the jackal story *makhàn* also occurs as 'sequential then'. 'They did that. Then *makhàn* they did (something else)', etc.

- (79) *dij mhen-laj-jen: in-à? ij ban, makhàn in-khè? kaku je,*
 He say-PROG-MODE:PST.ITR I-GEN feces NEG.COP then I-DAT/ACC fish give
 'He was saying (to the river): "(if/since) my feces is-not (here), so/then give me a fish"'

Quotatives with the (final) verb *lmhenanl* (and less commonly *lmhen-l* in other verb forms), 'said' or (with no subject), 'it is said'/'they say' are followed by what is said (as said, no modifications). Example: *dij mhenan: in baàngon* 'he said: 'I won't/ I refuse to (do whatever was proposed in the previous sentence)'. See also the sentence above – in the 'if' section – with *lmhenlajjenl*.

Some common subordinate constructions are listed here (see Nagaraja for more examples and a more detailed treatment):

1 The repeated present participle (a restricted set of PRS/FUT forms without the predicator), the doubled – repeated – form meaning continued, protracted, repeated action, a common areal feature. Example from the 'jackal text':

- (80) *dij ij-e ij-e gaḍa paṛi-kh-è?*
 he defecate-INAN COPY river block-INTNSV-PST.TR.
 'He (the jackal) defecating (repeatedly) blocked up the river.'

- 2 Clauses introduced by ‘because of, -*ghalya*’; ‘after’ /-à?-*baado-èn*/; ‘before’ /*N-à?*-*suŋu?*/, .and many others, postpositions, the ablative, etc. Example *jojomàsufu?* (REDPL-eat-GEN-before), ‘before eating’.
- 3 ‘in order to’ the ‘customary’ reduplicated infinitive ‘to V’, ‘in order to V’: *jikŋa-khè? lalaga ...*, ‘to chase/follow the porcupine(s)-ACC ...’
- 4 ‘when’ with clause -final /-*ki*/ Example from the porcupine story:

- (81) *dhuiã aka+baka-en ki dji-ku porŋa-àten oled-bà...*
 smoke suffocate-PST.ITR-ki, they hole-ABL come.out-FIN:PRS/FUT
 ‘when they have suffocated (begun to suffocate), they will come out of their hole(s)...’

5 SEMANTICS/DISOURSE

Not explicitly examined for this study.

6 LEXICON

Although Korku has borrowed extensively from Hindi, with up to half or more of its nominal lexicon borrowed from this language, much of the Munda ‘basic’ vocabulary remains such as body-parts, verbs of motion, etc.

7 BRIEF ANALYZED TEXTS

7.1 Text 1: *kolia* ‘The Story of Jackal’

- (i) *miã? kolia dan*
miã? kolia da-en
 ONE JACKAL be-PST.ITR
 ‘There was a jackal.’
- (ii) *dij ije ije gađa paŋik^hè?*
dji-ej ij-e ij-e gađa paŋi-ki-è?
 DEM-3SG.ANIM defecate-VB.SFX COPY river block.up-INTNSV-PST.TR
 ‘He, defecating (continually, repeatedly), blocked up the river.’
- (iii) *dŋgàten bāđa ađi he?en*
dŋ-ga-àten bāđa ađi hej-ya-en
 that-direction-ABL flood, big wave current, flow come-TLOC-PST.TR
 ‘A big wave (flow, current) came there.’
- (iv) *ađi djià? ijte? par tuyè?*
ađi dji-à? ij-te? par tu-ya-è?
 current he-GEN shit-3 completely carry.away-TLOC-PST.TR
 ‘The current/flood carried away all his shit.’
- (v) *dusra din kolia gađa qoqò? olen*
dusra din kolia gađa qo-qò? ol-en
 next day jackal river REDPL:see go-PST.ITR
 ‘The next day the jackal went to see the river.’

- (vi) *dij* *gaḍak^hè?* *mhenan:* *gaḍa gaḍa,* *iñà?* *ij*
dì-ej *gaḍa-k^hè?* *mhen-an* *gaḍa gaḍa,* *iñ-à?* *ij*
that-3SG.ANIM river-ACC say-PST.ITR river river I-GEN shit
tūganè?
ʃon-ga-èn-e?
which-direction-LOC-VBLZR
‘He (the jackal) said to the river: where is my shit?’
- (vii) *dij* *mhenlajjen:* *iñà?* *ij ban*
dì-ej *mhen-lab-ya-en* *iñ-à?* *ij ban*
that-3SG.ANIM say-AUX-T/A-PST.ITR I-GEN shit not(be)
mak^hàn *iñk^hè?* *kaku* *je*
mak^hàn *iñ-k^hè?* *kaku* *i-e*
then I-ACC fish give-INAN.O
‘He then (lit. ‘was saying’) went on to say: my shit isn’t (there); then give me (a) fish (instead).’
- (viii) *dīikalaaka* *gaḍa* *miā?* *kakuk^hè?* *teṛpak^hènej*
dī-ka?-laaka *gaḍa* *miā?* *kaku-k^hè?* *teṛpa-ki-è?(n)-ej*
that-EMPH-time river one fish-ACC throw-INTNS-PST.TR-3SG.ANIM
‘Then (‘at that time’) the river threw (him) a fish.’
- (ix) *kolia kaku* *oja(w)en* *ḍo olen*
kolia kaku *oja-wa(?) -en* *ḍo ol-en*
jackal fish pick.up-BEN(?) -PST.ITR and go-PST-ITR
‘The jackal picked up the fish and left.’

7.2 Text 2: *jikṛa* ‘The story of the porcupine’

- (i) *jikṛa* *miā?* *jaato* *janoar* *ʃ^haàṛba*
jikra *miā?* *jaato* *janoar* *ʃ^haàṛ-bà*
porcupine one community animal be, exist-PRES/FUT
‘There is a kind of animal, the porcupine.’
- (ii) *ṣṅku* *ḍōgor-èn* *kapṛig^heràn* *popṛa* *harujom-bà*
ṣṅ-ku *ḍōgor-èn* *kapṛi-g^hera-èn* *popṛa* *haru-jom-bà*
this-ANIM.PL forest-LOC mountain:side:LOC den,hole build:AUX-PRES/FUT
ḍo *ḍeèn* *ʃ^haàṛba*
ḍo *ḍi-èn* *ʃ^haàṛ-bà*
and that-LOC dwell-PRES/FUT
‘They make their holes (dens) on hillsides in the forest and live there.’
- (iii) *jikṛakùm* *pak^ha* *ʃ^haàṛba* *pini* *ḍikùten* *ap^hiru*
jikṛa-ku-èn *pak^hà* *ʃ^haàṛ-ba* *pini* *ḍi-ku-àten* *ap^hir-ù?*
porcupine-PL-LOC quill be-PRES/FUT but this-ANIM.PL-ABL (make)fly-PSSV
k^hob bannè?
k^hob ban-[n]è?
much NEG-3/PST.ITR
‘Porcupines have quills in them, but they don’t make them fly [eject them] much.’

- (xi) *jikɾakuk^hè?* *lalaga* *ɖo mumũdàn*
jikɾa-ku-k^hè? *la-laga* *ɖo mu-mũ-dà-èn*
porcupine-ANIM.PL-ACC REDPL:follow/chase and REDPL-hit,strike-LOC
k^hob catraeten ɖikù tauten saɾub hona
k^hob catrae-ten ɖi-ku-à? *tau-ten saɾub hona*
much,very clever-INSTR that-ANIM.PL-GEN (place)behind-INSTR/ABL run have.to
‘To chase and kill (lit. “beat”) the porcupines (one, hunters) have to very cleverly/carefully run behind them.’
- (xii) *banbaken* *ɖiku kaɾ^hà* *phaɾphaɾ-atɪŋkhè?-ki*
ban-ba-ki-en *ɖi-ku kaɾ^hà* *phaɾphaɾ-atɪŋ-k^hè?-ki*
not-otherwise-INTNS-PST.ITR they quill(s) move, make.quiver-AUGM-PST.TR.NF
ɖikù kaɾ^hà tirk^hija oledbà ɖo ɕiŋrièn
ɖi-ku-à? *kaɾ^hà tir-k^hija oled-bà ɖo ɕiŋri-èn*
that-ANIM.PL-GEN quill(s) arrow-like come.out-P/F and leg-LOC
k^huɕiùba
k^huɕi-(ya)-ù?-bà
pierce-TLOC-PASS-P/F
‘Otherwise they move/make quiver their quills and like arrows they come out (are released) and can pierce the leg(s) (of the hunters).’
- (xiii) *kaɾ^hà sarkakà(?) taùn ɕiùyùba*
kaɾ^hà sarka-kà? *tau-èn ɕiŋ-ya-ù?-ba*
quill(s) straight-EMPH place.behind-LOC to shoot-TLOC-PASS-P/F
mhenan
mhen-an
say-PST.ITR
‘It is said the quills are shot straight behind (them).’
- (xiv) *ɖiŋg^halya mumũdàmít^haj ɖijà*
ɖi-e-g^halya mu-mũdà-mi-t^hà-(e)j ɖij-à?
that-INAN-REASON REDPL:beat-one-NMLZR-3SG.ANIM he/she-GEN
bagolten saɾub hona ɖo ɖijk^hè? mumũdà hona
bagol-ten saɾub hona ɖo ɖij-k^hè? mu-mũdà hona
PLACE-PROL/ADESS to run to have to and him/her-ACC REDPL:beat/hit have.to
‘For that reason the hunter (lit. beater) has to run at the side/alongside him (the porcupine) to beat him.’
- (xv) *jikɾak^hè? coɬoɬkàbi jorten mũdàej*
jikɾa-k^hè? co-ɬoɬ-kà?-bi jor-ten mũdà-ej
porcupine-ACC which-QUANTITY-EMPH-EMPH strength,force-INSTR beat-3SG.ANIM
pini ɖij ban gujù? pini kaparten miã? jirasan mũdà
pini ɖij ban goj-ù? pini kapar-ten miã? jirasan mũdà
but he not kill-PASS but head-INSTR one mild, slight hit,strike
g^haienki ɖeɬogènka[?]
g^hai-en-ki ɖi-ɬo?-èn-kà?
to.adhere.to,receive.a.blow-PST.ITR-WHEN that-quantity-LOC-EMPH
gujùba mhenan
goj-ù?-bà mhen-an
kill-PASS-FIN:PRS/FUT say-PST.INTR

‘[If] however hard (lit. with however much force) (they) beat him he does not die (is not killed) but (then if) he gets a light blow on the head that is enough to kill him, it is said (lit. from that much[emphatic] he is killed/dies).’

(xvi) *korku* *jikɾakù* *jilu* *jomebà*
koro-ku *jikɾa-ku-à?* *jilu* *jom-e-bà*
 person-ANIM.PL porcupine-ANIM.PL-GEN meat eat.ITR-P/F
 ‘The Korkus eat porcupine(s) meat.’

jikɾakù *nāgà+ti* *korkù nāgà+ti-k^hijabà*
jikɾa-kù-à? *nāgà+ti* *korku-à? nāgà+ti-k^hija-bà*
 porcupine-ANIM.PLU-GEN foot+hand person-GEN foot+hand-SIMIL-P/F

mhenan

mhen-an

say-PST.ITR

‘They say/it is said that the porcupines’ feet-and-hands are like humans’ feet-and-hands.’

ɖiig^halya *sikarminku* *ɖikù(?)* *nāgà+ti*
di-g^halya *sikar-mi-en-ku* *ɖi-ku-à?* *nāgà+ti*
 that-reason hunt.one-NMLZR-ANIM.PL that-ANIM.PL-GEN foot+hand

gēɖakiba *ɖo* *teɾpakibà*
gēɖa-ki-bà *ɖo* *teɾpa-ki-bà*
 cut.off-INTNS-P/F and throw-INTNS-P/F

‘For that reason hunters cut their (the porcupines’) feet-and-hands off and throw them away.’

NOTES

* I am grateful to Arjun Jade Patel, my Korku tutor and friend, for what I know – not enough – of that language. But for Arjun Guruji I would not have gone on in Munda studies, the University of Chicago Indo-American Munda Languages Project would not have happened, and, presumably, Greg Anderson would not have gone into Munda, and this volume would in all likelihood not exist. Thanks to the editor for making extensive revisions and formatting to this chapter without whose encouragement, persuasion and assistance this chapter would simply not have seen the light of day. Responsibility for the content rests solely with the author of course.

- 1 One finds ‘bullocks’ in some places (with some speakers) as animate, in some as inanimate.
- 2 See the discussion of U in the verb section (sections 3.2.3–3.2.8) above for the contrast in meaning of this subjectless transitive and the passive(-potential) with U. For both of these one can translate the sentence ‘the dog was tied (down)’.

REFERENCES

- Anderson, Gregory D.S. (2007) *The Munda Verb. Typological Perspectives*, Trends in Linguistics Studies and Monographs 174, Berlin: Mouton de Gruyter.
- Anderson, Gregory D.S. and N.H. Zide. (2002) ‘Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint’, in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.

- Drake, John (1903) *A Grammar of the Kurku Language*, Calcutta: Baptist Mission Press.
- Ghosh, A. (this volume, Chapter 2) 'Santali'.
- Girard, Beryl (1965) *Korku-Hindi-English Dictionary*, unpublished manuscript, Ramakhera and Derhtalai.
- Grierson, G. (1906) *Linguistic Survey of India Volume IV: Munda and Dravidian*, (S Konow did the Korku description in Volume 4).
- Nagaraja, K.S. (1989) *Austroasiatic Languages, A Linguistic Bibliography*, Pune: Deccan College.
- (1999) *Korku Language: Grammar, Texts and Vocabulary*, Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo: Tokyo University of Foreign Studies.
- Osada, T. (this volume, Chapter 3) 'Mundari'.
- Zide, N.H. (1960) *Korku Phonology and Morphophonemics*, PhD dissertation, University of Pennsylvania.
- (1972) 'A Munda demonstrative system: Santali', in Jacqueline M.C. Thomas and Lucien Bernot (eds) *Langues et techniques. Nature et Société*, Tom I, Paris: Klincksieck, 267–274.
- (1978) *Studies in the Munda Numerals*, Mysore: Central Institute of Indian Languages.
- (1991) 'Possible Dravidian sources of some Munda demonstrative bases', *Studies in Dravidian and General Linguistics* (Krishnamurti Festschrift), Hyderabad, 349–364.

SORA*

Gregory D.S. Anderson and K. David Harrison

1 INTRODUCTION

The Sora language, also known variously as Saora, Saura, Savara, or Sabar[a] is one of the most important Munda languages. The tribal name is known from the Sanskrit sources but has been applied to numerous people other than the Sora proper. The main concentration of Sora speakers are located in Ganjam, Gajapati, and Rayagada districts of Orissa, but speakers are found throughout this state and in the adjacent parts of northern Andhra Pradesh. It is difficult to gauge accurately the total number of speakers of Sora, but estimates vary between 150,000 and 300,000, but could be much more. In some areas with significant ethnic Sora populations however (e.g. Sambalpur district of Orissa) a language shift has occurred. Thus, Sora language appears to be endangered in some areas, but in others (e.g. Gumma block of Gajapati district), most women and children are monolingual Sora speakers.

The only language that Sora is clearly at all closely related to within Munda is Gorum (Parengi) with which it forms the Sora-[Juray]-Gorum subgroup; Juray is a poorly known language that largely appears to be a divergent Sora dialect, but future research will have to tease apart their exact relation. Some accounts (e.g. Gordon (2005)) suggest that the Lodha (or Lodhi) language of northern Orissa is related based in part on the use of the autonym Sabar[a] by at least some members of this 'criminal tribe'; the only extant linguistic materials on this language (Das Gupta 1978) suggest that the autonym may be misleading in this case as the language appears to be an Indo-Aryan variety, not an undocumented member of the Munda language family. Traditional classifications of Sora-Gorum include this within a larger genetic group called Koraput Munda, and this in turn within a group called South Munda. Anderson (2001) suggests an alternative classification, which has Sora-Gorum splitting off early from the Proto-(South)-Munda language (sharing certain features with North Munda and a number of others with various South Munda subgroups). Unfortunately, resolving this taxonomic issue and recovering an adequate understanding of Sora-Gorum pre-history must await the discoveries of future research.

There is no adequate map or study of Sora dialectology. Speakers we interviewed in 2007 reported awareness of regional Sora speech differences across a range of structures, for example, phonetic, morpho-syntactic, prosodic, etc. The Sora speech community is also distributed across Oriya- and Telegu-speaking areas, and one would expect to find different language-contact effects in the two parts. The variety described in this chapter is that spoken by Sora of the central Gumma Hills region, Gajapati district, Orissa state.

Sora remains primarily an oral language, not widely read or written even by speakers who may be fully fluent and literate in Oriya and English. Though the Sora Bible may be found as a prized household possession in some Christianized

villages, it is not typically read. When highly literate speakers in Gajapati district of Orissa do read it, they report that it contains many unfamiliar words, and that the orthographic system is not completely familiar to them. Thus, the dialect differences in Sora are such that the fledgling literary standard is clearly non-ideal for all possible users of the language and that the norms for its orthographic rendering have not been sufficiently disseminated to make Sora literacy truly viable in its current state.

The following is a sample text from the Sora Bible, Psalm 23 (page 827, verse numbers as in original, Sora orthography is shown on the first tier):

- (i) GAMANGtungan na Gupa Mar-hen, henate er asuige
gəmaŋtuŋən-na gupa[:]mar-nen nen-ate er-asu-ige
 rich.god+EMPH shepherd:HUM-1 I-? NEG-be.ill-AUX:T/A
 ‘The Lord is my shepherd, I shall not be in want (be ill).’

- (ii) Anin d’ong-hen ledenggab lingan abtabmutih;
anin do?[o]ŋ-nen ledəŋgab-leŋən
 he BODY-1 tender.green.grass-LOC:N.SFX
ab-tabmu-t-ijŋ
 CAUS-lay.down[doll/child]-NPST-1
 ‘He makes me lie down (like a sheep) in the tender green pastures;’

Anin lagad hana d’an adam adamban d’ong-hen urungtih
anin lagad nana da?a-n adəmadəm-ban do?[o]ŋ-nen
 3SG peaceful/docile? water-N.SFX along-ALL OBJ-1
uruŋ-t-ijŋ
 convey-NPST-1
 ‘He leads me beside quiet waters.’

- (iii) Anin pwrarda-hen absamagte;
anin purarda-nen absamag-te
 3SG life-1 CAUS:RESTORE-NPST
 ‘He restores my soul;’

Anin ahimdaman apsele roitadgod lingan d’ong-hen urungtih.
anin aŋim-damən əpsele royta²dgo²d leŋən do?oŋ-nen
 he name-PURP PURP righteousness=road-LOC:N.SFX OBJ-1
uruŋ-t-ijŋ
 convey-NPST-1
 ‘He guides me in paths of righteousness for His name’s sake.’

The Sora Bible employs a Latin-based orthography with a number of Sora-specific conventions. Sora has also been rendered in the Oriya script in Orissa and in the Telugu script in Andhra Pradesh, as well as a modified phonetic alphabet in Ramamurti’s grammatical materials and dictionary. The use of and knowledge of Sorang Sompeng (N. Zide 1996), the indigenous script, appears to be quite limited.

In many areas Sora remains a vital and thriving language, but one that has no state or institutional support (sermons and materials are increasingly in Oriya in Gajapati district which we observed and were told about in Christianized Sora communities). In other areas, Sora is reportedly being or indeed has already been replaced by Telugu or Oriya. So, although not an endangered language *in sensu stricto*, Sora (except in

the areas shifting to Telugu and Oriya) like other similarly sized and underdeveloped minority languages of India, may be considered threatened overall.

2 PHONOLOGY

2.1 Vowel inventory

Sora has vowel phonemes and contrastive vowel length. Schwa [ə] is never stressed and has no long counterpart. Vowels [a] [o] [i] [u] [e] [ɪ] [ʊ] may be stressed or unstressed.¹ There appears to be significant variation in vowel quality across Sora dialects, not yet adequately documented.

Vowel length is (probably) not phonemic in Sora, but may be used to create expressive formations with certain stems, for example, *sura* ‘big’ *surra* ‘really big’. According to Ramamurti (1986), vowels may be short, half-long, or long, but the phonemic status of vowel length requires further study. In data drawn from Ramamurti, we preserve his length marking.

Vowel assimilation processes have not been adequately studied, but vowels may undergo assimilation to their preceding consonant for rounding:

- (1) *li* > [u] / [p] — *yur'pid* ~ *yur'pud* ‘to shake off’

2.2 Suprasegmental phenomena (tone, register)

Tone and register or voice quality have not been reported in the literature on Sora but impressionistically, there appeared to be certain features of this type associated with certain words for at least some speakers. This has not been tested experimentally. Note that Gorum, Sora’s only close sister language, has a creaky voice feature (see Anderson and Rau, this volume).

Vowels may be glottalized or glottally interrupted. This strategy is particularly noticeable in the case of monosyllabic nouns, for which the glottal stop may serve phonologically to add a mora to the syllable, thus satisfying a minimal word template. Evidence of this can be seen in the following minimal pair:

- m²u* ‘nose’ (lexical form) vs. *mu* ‘nose’ (combining form)

2.3 Consonant inventory

(2)	Labial	Dental	Retroflex	Palatal	Velar	Glottal
Obstruents:						
Voiceless	p	t		c	k	ʔ
Voiced	b	d		j	g	
Fricatives		s				
Nasals	m	n		ɲ	ŋ	
Flaps			r	ɽ		
Lateral		l				
Glides				y		

Consonants may also be pre-glottalized, for example, [ʔb], [ʔm] a phenomenon known from other Munda languages, and one that awaits instrumental analysis in Sora.

Assimilation of both progressive and regressive types may be observed. Geminate consonants may result from these processes.

(3)	Assimilation	Attested variants	Gloss ²
	/b/ > [m] / __ [m]	<i>'mib'mib ~ mim'mib</i>	'to be sleepy'
	/b/ > [r] / __ [r]	<i>'jənar'ruj < lj<ən>ab-ruj/</i>	'to churn'
	/b/ > [s] / __ [s]	<i>rabsa:d'o:lən ~ rassa:d'o:lən</i>	'dry leaves'
	/b/ > [k] / __ [k]	<i>rakud- ~ rakkud-</i>	'to make narrow'
	/d/ > [b] / __ [b]	<i>rib'bidən ~ rid'bidən</i>	'powdered salt'
	/d/ > [s] / __ [s]	<i>'pudsu: ~ 'pusse:</i>	'to persuade'
	/d/ > [r] / __ [r]	<i>'radrad ~ 'rarrad</i>	'chop'
	/d/ > [p] / __ [p]	<i>raɖ'pui ~ 'rap'pui</i>	'to fall off'
	/d/ > [l] / __ [l]	<i>id'loj ~ il'loj</i>	'to be contracted'
	/n/ > [m] / __ [b]	<i>an'baŋ ~ am'baŋ</i>	'to hear, listen'
	/m/ > [n] / __ [t]	<i>santab ~ samtab</i>	'to set apart'
	/ŋ/ > [m] / __ [b]	<i>'iliŋ 'bo:ŋən ~ 'ilim 'bo:ŋən</i>	'rainbow'
	/n/ > [l] / __ [l]	<i>'punla:n ~ 'pulla:n</i>	'chaff'
		<i>jalla ~ janla</i>	'to be slow'
	/r/ > [m] / __ [m]	<i>mommō:ri:n ~ mormō:ri:n</i>	'mist'
	/b.s/ > [t.t]	<i>ərabsuis:i:n ~ ərattuis:i:n</i>	'pointing hand'
(4)	Epenthesis of linking consonant		
	Ø > d / [n]__ [r], [l]	<i>kənri ~ kəndri</i>	'rice gruel'
		<i>kanla:n ~ kandla:n</i>	'a leaf-stitched cup'
		<i>bən'ra:b- ~ bənd'ra:b-</i>	'be angry'
	Ø > g / [ŋ]__	<i>eŋra:n ~ eŋgra:n</i>	'prickly cucumber'

Sora has true geminates for all consonants except the glottal stop and the glide [y]. From our limited set of relevant field data, it appears that geminate stop consonants are (nearly) twice the duration of simplex ones. In the two examples that follow, the geminate [gg] tokens have a duration of 0.14 and 0.12 seconds, respectively, as compared to a typical duration of 0.07 seconds for tokens of singleton [g].

(5)	(a) <i>ʔag-gaʔ</i>	(b) <i>ʔamin ʔa'g-ga</i>
	CAUS-eat:IMP	you CAUS-eat
	'Make him eat', that is, 'feed him'	'Make him eat'

2.4 Syllable structure and phonotactics

The velar nasal does not occur in syllable or word-initially, except in rare cases. Ramamurti lists just seven words with initial [ŋ], all expressives, for example, *ŋak'ŋak* 'dead silence, as at midnight' (Ramamurti 1986 (1933):185).

The following heterosyllabic consonant clusters are found in Sora lexemes.³

(6)

	<i>p</i>	<i>b</i>	<i>m</i>	<i>t</i>	<i>d</i>	<i>n</i>	<i>ɲ</i>	<i>y</i>	<i>j</i>	<i>s</i>	<i>r</i>	<i>l</i>	<i>k</i>	<i>g</i>	<i>ŋ</i>	<i>ʔ</i> ⁴
<i>p</i>	<u>✓</u>			✓					✓	✓						
<i>b</i>	(✓)	<u>✓</u>	✓	✓	✓				✓	✓		✓	✓			✓
<i>m</i>	✓	✓	<u>✓</u>	✓	✓		✓	✓	✓	✓			✓	✓		✓
<i>t</i>				<u>✓</u>								✓				✓
<i>d</i>	✓	✓	✓	✓	<u>✓</u>	✓					✓	✓	✓	✓	✓	(✓) ✓

<i>n</i>	✓	✓		✓	✓	✓		✓	✓		✓	✓	✓		
<i>ɲ</i>		✓			✓		✓				✓		✓		(✓)
<i>y</i> ⁵	✓	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓		
<i>j</i>									✓						✓
<i>s</i>									✓						✓
<i>r</i>	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	(✓)
<i>l</i>	✓	✓	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓	✓
<i>k</i>												✓			(✓)
<i>g</i>													✓		✓
<i>ŋ</i>	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
?														✓	

Note

✓ indicates geminates. (✓) indicates rare, found only in expressives, compounds or loanwords.

- (7) *p*
kuppa: ‘fat’
raptij- ‘to be able to do’
ap’sɛle ‘for, on account of’ (postposition)
- (8) *b*
gabtur ‘to pinch between fingers’
babje ‘to console’
jabda- ‘to suck or munch with noise’
jabjab- ‘to adjure, abut’
rabsa:d’o:lən ‘dry leaves that have dropped from a tree’
rabkud- ‘to make narrow’
b?o:b ‘head’
ab’baja- ‘to hasten, vex’
b?ai ‘to turn into seed’
dub’mad- ‘to shut one’s eyes’
ob’leŋ- ‘to reply, cross-examine’
pijebpijeb ‘peep’ (onomatopoetic)
pabjab- ‘to smite’
- (9) *m*
kəmbud ‘to be carried away by a bear’
gumte ‘to dream’
kimmad ‘to close the eyes’
kumsi ‘to hold one’s fist’
kimpoŋən ‘belly’
jumjum ‘cover body with cloth’
’tamd- ‘to wash (clothes)’
ɲam’ɲam- ‘to eat’ (children’s language)
ɲamjo- ‘to fish’
padum-gamle ‘suddenly, with a thud’ (onomatopoetic)
m?u:n ‘nose, beak of bird’
- (10) *t*
kutta:n ‘a pen for sheep’, ‘a vegetable garden’
’tʔalən ‘spleen’

- (11) *d*
ənartud^lpel 'matchbox'
kudta:n 'a horse'
ɲidɲid bæren 'balderdash'
rad^lmoy 'to wring out'
raddu 'strong'
rid^lbidən 'powdered salt'
sud^lsud- 'to mix up'
sud^lgum- 'to be drenched'
dʔu:l- 'to suck'
pad^lkad- 'distribute buffalo flesh'
- (12) *n*
gunturən 'a large rat'
ondo:ta 'forwards'
gunlun 'snail'
minnum 'year'
an^lreɲ 'already'
anbənar^ltiki 'afterwards'
dum^lpulən 'projected navel, one who has such'
ensoy 'alone'
- (13) *n*
raɲjam- 'to become dry or withered'
aɲnam- 'to name'
aɲja 'barren'
dɲɲbo:ɲ- 'to cook the flesh of a buffalo'
dɲɲda:r^lpadən 'pot in which rice is cooked'
nʔo:i nʔo:i 'the bellowing of a buffalo'
- (14) *y*
yʔo:ɲən 'a stranger'
pui^lra:b 'to put forth a new sprout'
pui^lda- 'to peel off in thin layers'
na:i:n 'river'
moɲkiyla: 'not uniform in size'
pɑ:ygaɲən 'the halo around the moon'
- (15) *j*
bajje- 'to console'
jʔun 'to be fit, to suit'
- (16) *s*
sʔo:r- 'to bubble up'
sussub 'a fib'
- (17) *r*
urɲaɲ 'to unstring the bow'
ersin 'crime, sin'
ɲaɲaɲarloge 'with a quaver'
irdo^ldoi- 'to rebound'
jarre 'again'

<i>rʔa:n</i>	‘elephant’
<i>ənartud^lpel</i>	‘matchbox’
<i>ar^lgadi</i>	‘alliance, friendship’ (~ <i>ar^lgadi</i>)
<i>ar^lda:-</i>	‘to ladle water’
<i>ar^lpə:d-</i>	‘to unfold’
<i>ar^lmodən</i>	‘cobwebs’
<i>aur^lgab-</i>	‘to chew the cud’
<i>bənar:rjumən</i>	‘dinner given in recognition of service’
<i>erlə^lnu:</i>	‘undomesticated, wild’
<i>er^lna:nə:n</i>	‘name of Sora deity’
<i>er^lsi:n</i>	‘crime, evil menstruation’

(18)

<i>l</i>	
<i>molloi</i>	‘five’ (~ <i>monloi</i>)
<i>^lta:lmad-</i>	‘to see with fixed eyes’
<i>tal^lsi:</i>	‘palm of hand’
<i>al^ldub-</i>	‘to wrench the neck’
<i>al^lber-</i>	‘to converse’ (Reciprocal)
<i>al^lgadi</i>	‘alliance, friendship’ (~ <i>ar^lgadi</i>)
<i>alpad^lded-</i>	‘to quarrel with one another’
<i>al^lrode-</i>	‘to quarrel or dispute with one another’
<i>al^lkalla-</i>	‘to be astounded, dumbfounded’
<i>dul^ltid-</i>	‘to beat in return’
<i>^ldulna:</i>	‘instead of’

(19)

<i>k</i>	
<i>ŋakŋak</i>	‘dead silence’
<i>^ltakko-</i>	‘to deduct’

(20)

<i>g</i>	
<i>miggāl</i>	‘twelve’
<i>g^lwr-</i>	‘to ripen’

(21)

<i>ŋ</i>	
<i>mʊŋka:n</i>	‘face’
<i>keŋgʊ:ŋən</i>	‘woodpecker’
<i>joŋjoŋ</i>	‘to pour rice’
<i>jenəŋde:n</i>	‘however’
<i>əme:ŋ^lpurən</i>	‘raw meat sacrifice’
<i>əso:ŋ^lneb</i>	‘a kind of tree’
<i>əso:ŋmed^ltulən</i>	‘Bengal gram, or peas’
<i>əso:ŋbo:b^lmarən</i>	‘insolent man’ (lit. ‘shit-head’)
<i>^ldʊŋroi</i>	‘early in the morning’
<i>dʊŋ^lsi:-</i>	‘to carry something in the hand’
<i>diŋlu:d-</i>	‘to have ears hanging down’
<i>ŋaŋŋaŋ</i>	‘to teach’
<i>saŋŋa:ŋ-</i>	‘to keep at a distance’

Tri-consonantal clusters are rare, including

(22)

<i>nts</i>	
<i>əntsə^lriŋ</i>	‘thereafter’

2.5 Intonation/stress

The system of stress assignment in Sora remains largely unstudied. It may appear in different syllables in a word. Some morphemes seem to attract stress while others do not accept it at all (see Ramamurti 1931:7–8). This suggests at least a partially morpho-lexically conditioned system.

2.6 Morphophonology

Both consonants and vowels may show some (conditioned?) alternations perhaps based on stress-shifts or perhaps morphologically conditioned. The details remain to be worked out. Note in this regard the following possessive paradigm for the word meaning ‘eye’, that is, ‘my eye’, ‘your eye’, etc. Here we find allomorphs *maʔlmaʔd*, *mɔd*, and *moʔ*.

- (23) (a) *ɲen maʔ-ɲen* (b) *maʔ-nɔm* (c) *maʔ[d]-len*
 I eye-1 eye-2 eye-1PL
 ‘my eye’ ‘your eye’ ‘our eye(s)’
 (d) *a-mɔd-an* (e) *amben moʔ-bim*
 3-eye-N.SFX you (PL) eye-2PL
 ‘his/her eye’ ‘your (PL) eye(s)’ [OG]

The past tense marker *-l-* variably (idiolectally?) may undergo assimilation to a preceding stem-final *-r*.

- (24) (a) *anin-a suʔuy-ban iar-lɛ*
 He-GEN house-ALL go-PST
 ‘he (elephant) went to his (big-frog’s) house.’ [Text-1, line 11]
 (b) *uan a-ier-re ən-n-ɔləŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-PST-NEG-PL
 ‘“where did you go?”, they didn’t answer him.’ [Text-1, line 16]

In rapid speech, there may be a place assimilation across word boundaries, for example, between the final nasal in the first or third person pronoun and the initial obstruent of the following word, for example, from *ɲen* ‘I’ or *anin* ‘she/he’ we hear such forms as the following:

- (25) (a) *ɲem bajar-ən ʔyer-t-e*
 I market-N.SFX go-NPST[:1/TLOC]
 ‘I am going to the market.’ [OG]
 (b) *anin giʔy-t-ey*
 she/he see-NPST-2/3
 ‘she/he sees’ [OG]

3 MORPHOLOGY

Sora has an elaborate derivational apparatus that is seen in the formation of nominal forms. From an inflectional perspective, Sora nouns may realize number, a range of case and quasi-case adpositional categories, and person/possession. In addition to nouns, nominal forms include pronouns (including demonstratives, indefinites, and interrogatives) numerals, adjectivals, and adverbials.

3.1 Nominal morphology

The maximally expanded form of a noun in Sora has not yet been investigated. Stems may be compounded yielding fairly long constructs. How these are constrained remains a topic for future research. At least three stems can be found compounded in certain common words.

- (26) *daiburyuən* ‘sunrise’ (Lit. climb-hill-sun)
dai ‘climb’ + *bur(ən)* contracted form of *bəru:n* ‘hill’ (Ramamurti 1933:114)

A quick look through Ramamurti’s Sora dictionary’s numerous three-part and indeed four-part compounds, for example, *əgodlobum*, *əjajra:da:tid*, *əartuddaype:l*, *anəŋgatidrud*, and *abdimmaddar:em*, etc. (Ramamurti 1986 (1933):5–6, 11, 21, 33). Long forms can arise from simple two-part compounding (or perhaps izafet constructions) with derived nominal forms that have reduplication (Ramamurti 1986 (1933):10).

Some common inflected morphological noun templates attested in our materials include

- (27) Stem-Stem-Noun.Suffix-Plural
sənnə-mər-ən-ji
 little-person-N.SFX-PL
 ‘the little guys’
- (28) Noun-1/2POSS-Plural/Case
alaŋ-ben-ji
 tongue-2PL-PL
 ‘your (pl) tongues’ [OG]
- (29) 3POSS-Noun-Noun.Suffix-Plural
a-on-ən-ji
 3POSS-child-N.SFX-PL
 ‘his children’
- (30) 3POSS-Noun-Noun.Suffix-Case
a-kako-n-adɔʔəŋ
 3-older.brother-N.SFX-OBJ
 ‘his older brother’

Formal nominal derivational and inflectional processes in Sora may be realized through prefixes, reduplication, suffixes, at least one common infix, and even a process of circumfixation.

3.1.1 Number

Plural in both nouns and (third person) verbs is marked by the suffix/enclitic *-ji*. It follows the multipurpose [ə]n noun-suffix and possessive markers in the Sora noun-word template.

- (31) (a) *gam-le kun sənnə-mər-ən-ji bɔibɔi ɛʔaka-le-n-ji*
 say-PST DEF little-person-N.SFX-PL very rejoice-PST-ITR-PL
 ‘the little guys said and rejoiced greatly’ [Text-1, line 33]

- (b) *əntɛ ɛttele oʔa-leŋ-ən əteŋ kəndud-ən-ji dəku-le*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 ‘like that there were many frogs in the ditch’ [Text-I, line 14]
- (c) *amben alaŋ-bm-ji*
 you (PL) tongue-2PL-PL
 ‘your (PL) tongues’ [OG]

There is no adjective noun concord, but adjectives may be nominalized and then marked for plural.

- (32) (a) *anin kuddib kəndud-ən-ji sireŋ anin suʔa*
 he all frog-N.SFX-PL from/than he big
 ‘he is bigger than all the other frogs.’ [Text-I, line 7]
- (b) *kudub-ən-ji daʔa-leŋ-ən gəlu-le-ji*
 all-N.SFX-PL water-LOC-N.SFX fall-PST-PL
 ‘(they) all fell into the water.’ [Text-III, line 14]

Numerals may take a similar set of markers as well in Sora.

- (33) *aninji bagu-n-ji bənlunəŋ dəku-le-ji*
 they two-N.SFX-PL clan/NMLZ/clan COP-PST-PL
 ‘both of these were two brothers (of one clan).’ [Text-2, line 2]

While animate nouns generally (but non-obligatorily) take the plural suffix, inanimates often may not.

- (34) *anlen si-len si-len ənlen si-len*
 we hand-1PL hand-1PL we hand-1PL
 ‘our hand(s)’ ‘our hand(s)’ ‘our hand(s)’ [OG]
- si-ben*
 hand-2PL
 ‘your (PL) hand(s)’ [OG]

Two coordinated nouns marked for plural may appear in an asyndetic/juxtaposed form with plural only on the rightmost conjunct. However, both may appear plural-marked as well, so plural is not obligatorily phrasally marked in Sora.

- (35) *uan-ji kəniət-le-n-a tiki bagu-n-ji jənnəŋ*
 father-3PL die-PST-N.SFX-GEN after two-N.SFX-PL field
səʔba-n-ji mailen bara-le-ji
 paddy-N.SFX-PL together work-PST-PL
 ‘after their father died, they both worked in their fields and paddies together’
 [Text-2, line 6]

Except after numerals, where the singular form of the noun is generally used, after quantifiers, (animate) nouns may or may not appear pluralized but inanimates generally do not.

- (36) (a) *aboj tuləb-leŋ-ən dəjeŋ arsi-n-ji dəku-le-ji*
 one forest-LOC-N.SFX several monkey-N.SFX-PL be-PST-PL
 ‘a number of monkeys were in a certain forest’ [Text-3, line 1]
- (b) *dəjiŋ meŋim ettegoŋ bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX/ITR-PL-GEN after

a-sanna-mər sukku-n a-kako-n-adɔʔɔŋ
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX=OBJ

sarɔba-n er-ti-lə-be anson bara-eten
 paddy.field-N.SFX NEG-give-PST-NF:W/O himself work-T/A:3
 '[after] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field' [Text-2, line 7]

- (c) *ənte ettele oʔa-ləŋ-ən əteŋ kəndud-ən-ji dəku-lə*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 'like that there were many frogs in the ditch' [Text-1, line 4]

- (d) *nen 'ba:gu 'mandra 'giʔ-l-ay*
 I two man see-PST-1
 'I saw two men.' [OG]

3.1.2 Case

The unmarked, basic, or uninflected form of the [pro]noun serves the function of the subject, primary (direct/indirect) object and goal arguments as well in Sora. Note that as in all nominal forms, the suffix in *-ən*, so common in nominal forms in Sora, appears frequently on nouns in any of these functions. This is, however, not a case marker *per se* at least synchronically. For a discussion (without much resolved) on (some of) its use(s), see Starosta (1967:255–256) (1981) and Ramamurti (1931:16–17).

- (37) (a) *aninji anin daʔa-n tiyʔ-te-ji*
 they she/he water-N.SFX give-NPST-3PL
 'they give him/her water' [OG]
 (b) *nen bo-ma'dra...bo-man'dra 'giʔ-l-ey*
 I one-man one-man see-PST-1[:3]
 'I saw a man.' (Note /n/ deletion in rapid speech) [OG]
 (c) *nen bajar-ən 'yer-t-e*
 I market-N.SFX GO-NPST[:1/TLOC]
 'I am going to the market.' [OG]

The enclitic *-dɔʔɔŋ* 'body' (in a possessed form) may function as an oblique object of an animate (or narratively elevated to quasi-animate) primary object argument of a verb that cannot be expressed morphologically in it (the complex system of verb agreement in Sora is discussed in sections 3.2.1 and 3.2.2). Thus, this element may be considered an oblique object or dative marker synchronically in Sora. Historically it is probably a (possessed form of the) word meaning 'body' (see section 3.1.3)

- (38) (a) *nen 'deyvid 'adɔŋ 'giʔ-l-ay*
 I David OBJ see-PST-1
 'I saw David.' (emphatic)
 (b) *nen [ŋ]ka-mər-ən a'dɔŋ 'giʔ-l-ay*
 I tall man-N.SFX OBJ:3 see-PST-1
 'I saw the tall man'
 (c) *nen boʔonsil-aɔŋ je'lu-n tig-ʔe-n je'lu-n*
 I one woman-OBJ meat-N.SFX give-PST-NF meat-N.SFX
 'I gave the meat to the woman, meat' [OG]

- (d) *ubban-adɔʔɔŋ* *gam-etan* *amən* *etenasən*
 younger.brother-OBJ say-T/A-N.SFX/ITR you why
səɾɔba-n *ət-ti-ŋ*
 paddy.field-N.SFX NEG-give-1
 ‘He said to his younger brother ‘why (do) not (you) give me some paddy field.’ [Text-2, line 9]
- (e) *kun asən kun sənna-dud-ən-ji* *raʔa-n=adɔʔɔŋ*
 DEF for DEF small-frog-N.SFX-PL elephant:N.SFX=OBJ
gij-an *gij-le* *bətɔŋ-le* *iersed-le-ji*
 see-N.SFX see-PST be.frightened-PST go.AUX-PST-PL
 ‘Because of seeing the elephant, the small frogs were frightened and ran away.’ [Text-1, line 12]
- (f) *bindɔ a-gaʔin-ji gam-le-ji ian-gamle aŋgaʔ-ən-adɔʔɔŋ təb-be*
 but 3-friend-PL say-PST-PL how! moon-N.SFX-OBJ take.out-1PL
 ‘“But how can we get the moon out?” her friends said.’ [Text-3, line 7]
- (g) *nen* ¹*doŋ-nam* *dʒiʔdʒi-ti-n-e*
 I OBJ-2 believe-NPST-RFLXV-1
 ‘I believe you.’ [OG]
- (h) *nen* ¹*doʔoŋ-nam* *gə¹lam-t-ai*
 I OBJ-2 know-NPST-1
 ‘I know you.’ [slowly]
- (i) *nen* ¹*doʔoŋ-nam* *ʔa¹gəlam-ai*
 I OBJ-2 NEG-know-1
 ‘I do not know you.’ [OG]

A different but similar adessive oblique ‘object’ *maŋ-* construction is found with verbs like ‘believe’ in forms like the following with a negative marked (and overtly intransitive) verb. Whether this is an idiosyncrasy of this speaker from Gajapati district or an example of the kind of object ‘demotion’ one finds in such languages as Russian or Estonian in negative formations is unknown (to genitive/partitive).

- (39) (a) *nen* ¹*doŋ-nam* *dʒiʔdʒi-ti-n-e*
 I OBJ-2 believe-NPST-RFLXV-1
 ‘I believe you.’
- (b) *nen* *maŋ-nam* *ʔa¹-dʒiʔdʒi-n-e...* *ʔa¹-dʒiʔdʒi-n-e*
 I ADESS-2 NEG-believe-RFLXV-1 NEG-believe-RFLXV-1
 ‘I do not believe you.’
- (c) *nen* *maŋ¹-nam* *ʔa¹-dʒiʔdʒi-n-e...*
 I ADESS-2 NEG-believe-RFLXV-1
 ‘I do not believe you.’ [OG]

When it has its directional or locational meaning it may be augmented by *-ba-n*, especially in non-pronominal complements in the form *-a-məŋ -ba-n*, for example, *ənselo-n-a-məŋ-ba-n* ‘to (near) the woman’ (Starosta 1967:164).

The clitic or suffix *a* may appear with pronouns and some nominal formations to mark a genitive-like or possessive relation between a [pro]noun and a noun. The second noun may be a relational noun functioning as a postposition (and with nominalized verbal forms, adverbial clausal subordinators as well). It is here glossed -GEN. It is likely a very old feature of Sora, with parallels in Juang and North

Munda. In some instances its function seems to be something akin to the *izafet* vowel of Persian, serving as an attributive link between nominals.

- (40) (a) *kuna gailo-n-a jattə abəj ɔɣa-n dəku-le*
 DEF road-N.SFX-GEN below one/INDEF ditch/drain-N.SFX COP-PST
 ‘There was a ditch/drain below that road.’ [Text-1, line 2]
- (b) *eɣale kun gəɣan-ə məndra:-ji ə-lo-n-a məneɣ*
 how DEF town-GEN person-PL PFX-road-N.SFX-GEN beside
ɔɣe gailun-a məneɣ ɔɣa-n daku
 or river-N.SFX:GEN beside ditch/drain-N.SFX be[come]
 ‘How is it that the town people have a drain/ditch either beside a river or by a road?’ [Text-1, line 3]
- (c) *anin-a suɔuɣ-ban iar-re[...iar-le]*
 he-GEN house-all go-PST
 ‘He (elephant) went to his (big-frog’s) house.’ [Text-1, line 11]
- (d) *dəjiɣ meɣnim etteɣoy bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX/ITR-PL-GEN after
a-sənna-mər sukku-n a-kako-n-ado?ɔɣ
 3-young-man Sukku-N.SFX-PL 3-older.brother-N.SFX=OBJ
sarəba-n er-ti-lə-be anson bara-eten
 paddy.field-N.SFX NEG-give-PST-NF:w/o himself work-T/A:3
 ‘[after] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.’ [Text-2, line 7]
- (e) *etteɣoy anlen-a alale-n-ji ted-an ted-le-n*
 like.this we-GEN REDPL:tail-N.SFX-PL hold-N.SFX hold-PST-N.SFX/ITR
aɣɣaj-ən-ado?ɔɣ a-təb-n-ai-bə
 moon-N.SFX-OBJ 1PL-get.out-n-CLOC-PL:IMP
 ‘In this way let’s hang by our tails and get out the moon.’ [Text-3, line 10]

The complex case suffix *-leɣ-ən* sometimes just *-leɣ* marks inessive, general locative, and illative functions in Sora.

- (41) *-leɣ-ən*
- (a) *puttar-leɣ-ən jəpba-leɣ-ən tilti-li-ɣ-ji*
 hole-LOC-N.SFX mire-LOC-N.SFX REDPL:bury-PST-ITR/RFLXV-PL
 ‘They buried themselves in holes and mud.’ [Text-1, line 13]
- (b) *aboj tuləb-leɣ-ən dəjeɣ arsi-n-ji dəku-le-ji*
 one forest-LOC-N.SFX several monkey-N.SFX-PL be-PST-PL
 ‘A number of monkeys were in a certain forest.’ [Text-3, line 1]
- (c) *kun arsi-n gam-eten bo-mənra a-kəndar-leɣ-ən*
 that monkey-N.SFX say-T/A:3 one person OBJ-branch-LOC-N.SFX
a-təd-nə-ba
 1PL-hold/hang-n-PL:IMP
 ‘That monkey told them “let one of us guys hold/hang from the tree branch.”’ [Text-3, line 8]
- (d) *ənte ettele oɣa-leɣ-ən əteɣ kəndud-ən-ji dəku-le*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 ‘like that there were many frogs in the ditch’ [Text-1, line 14]

- (e) *dɔ bɔ dinna kuna gailɔ-leŋ abɔj raʔa-n*
 so one day the road-LOC one/INDEF elephant-N.SFX
iar-eted
 walk-T/A:3
 ‘so one day an elephant was walking along the road’ [Text-1, line 9]
- (f) *e gəŋŋ-ji a-gij-ba aŋgaj-en daʔa-leŋ-ən*
 hey friend-PL 1PL-SEE-PL:IMP moon-N.SFX water-LOC-N.SFX
gəlu-le gamle ɔppuŋ-eten
 fall-PST QUOT tell-T/A:3
 ‘She told them “Hey friends, let’s go have a look; the moon has fallen in the water.”’ [Text-3, line 5]

Like many locative elements across the world’s languages, this case may appear with time nouns in Sora as well.

- (42) *ənt aiəm-leŋ-ən anin suʔa-dud-ən ət-dəku-əd*
 that time-LOC-N.SFX he big-frog-N.SFX NEG-COP-NEG[PST:3]
 ‘At that time he, the big-frog, wasn’t there.’ [Text-1, line 10]

A number of postpositional elements (and relational noun formations, see section 3.1.9) exist in Sora. Some of these appear to be drawn or have already been drawn into a now-increasing local case system. One such element is the allative marker *-ban*.

- (43) (a) *anin-a suʔuŋ-ban iar-rɛ[...iar-lɛ]*
 He-GEN house-ALL go-PST
 ‘He (elephant) went to his (big-frog’s) house.’ [Text-1, line 11]
- (b) *anin lagadɲana daʔa-n adəmadəm-ban*
 3SG peaceful/docile?? water-N.SFX along-ALL
doʔ[o]ŋ-ŋen uruŋ-t-ijn
 OBJ-1 convey-NPST-1
 ‘He leads me beside quiet waters.’ (Sora Bible)

3.1.3 Person

Possession of certain nouns may be marked by enclitic, possessive, and pronominal forms for first and second person (and third plural?) and prefixally/proclitically for third person singular. Sometimes, not insignificant morphophonological shifts occur in stems in possessed forms (see the paradigm for ‘eye’ below). Some pronouns may occur with the genitive/possessor enclitic *a*. Whether this is the same as the homophonous third person proclitic is unknown at present although they seem to not (usually?) co-occur. Common forms like ‘their hands’ are pronounced as one word usually anyway or use the alternative suffixal/enclitic third plural possessor strategy on the noun itself, like first and second person forms.

- (44) (a) *ŋen-a si-ŋen* (b) *ŋen si-ŋen*
 I-GEN hand-1 I hand-1
 ‘my hand’ ‘my hand’
- (c) *amən-a si-nam* (d) *anlen si-len*
 you-GEN hand-2 we hand-1PL
 ‘your hand’ ‘our hand(s)’

- (e) *si-ben*
hand-2PL
'your (PL) hand(s)'
- (f) *aninj-a-sî[-n]*
they-GEN-hand-N.SFX
'their hand(s)'
- (g) *aninj-a-sî*
they-GEN:3-hand
'their hand(s)'
- (h) *maʔ[d]-jen*
eye-1
'my eye'
- (i) *maʔ-nam*
eye-2
'your eye'
- (j) *a-məd-an*
3-eye-N.SFX
'his/her eye'
- (k) *maʔ-len*
eye-1PL
'our eye(s)'
- (l) *moʔ-bim*
eye-2PL
'your (PL) eye(s)'
- (m) *a-məd-an*
3-eye-N.SFX
'their eye(s)'
- (n) *alaŋ*
[3:]tongue
'his/her tongue'
- (o) *alaŋ-jen*
tongue-1
'my tongue'
- (p) *alaŋ-nəm*
tongue-2
'your tongue'
- (q) *ənlen alaŋ-[l]ən-ji*
we tongue-N.SFX/1PL-PL
'our tongues'
- (r) *amben alaŋ-bim-ji... amben alaŋ-bim-ji*
you (PL) tongue-2PL-PL... tongue-2PL-PL
'your (PL) tongues' [OG]

The variation between prefix/proclitic/enclitic/suffix *a* between the possessor and possessum may be seen in the following pair of sentences with the words/phrases meaning 'their father'.

- (45) (a) *bo dinna aninji-a-uan kəniət-le*
one day they-GEN-[3:]father die-PST
'one day their father died' [Text-2, line 5]
- (b) *uan-ji kəniət-le-n-a tiki bagu-n-ji jənnəŋ*
father-3PL die-PST-N.SFX-GEN after two-N.SFX-PL field
səwəba-n-ji mailen bara-le-ji
paddy-N.SFX-PL together work-PST-PL
'After their father died, they both worked in their fields and paddies together.' [Text-2, line 6]

Nouns that begin in *a-* have an unmarked third (singular) possessive form.

- (46) (a) *anim-jen opino gomango*
name-1 Opino Gomango
my name is O. G. [OG]
- (b) *bo-mənra-n aŋim sukku bar bo-mənra aŋim mənŋiɾa*
one person-N.SFX 3:name Sukku and one person:GEN 3:name Mangaɾa
'one was named Sukku one was named Mangaɾa' [Text-2, line 3]

The possessive marker *a-* appears most frequently with kin-terms.

- (47) (a) *bijndɔ a-kako-n kan-ate əmdəŋ-ən əmdəŋ-le dɔ*
 but OBJ-elder.brother-N.SFX that-PRTCL hear-N.SFX hear-PST and
gij-an gij-le bɔibɔi barab-le
 see-N.SFX see-PST very get.angry:EMPH-PST
 ‘But the elder brother got very angry when he heard and saw (all) that
 (his brother was doing).’ [Text-2, line 8]
- (b) *bar bɔndi-lə-n-ji-na tiki dəjin dinna de-le*
 and put.in.jail-PST-N.SFX-PL-GEN after few day become-PST
anin a-dukəɽ-ŋ dɔ a-on-ən-ji jinaŋ doləjən batte
 he 3-wife-N.SFX and 3-child-PL also hunger-N.SFX SOC/INS
kəniət-lə-ji
 die-PST-PL
 ‘and a few days after they put him jail, his wife and children also died
 of starvation.’ [Text-2, line 13]
- (c) *bijndɔ a-ubban bɔibɔi barab-le ier-an ier-le*
 but OBJ-younger.brother very get.angry-PST go-N.SFX go-PST
anin adɔʔɔŋ tuəb-eten
 he OBJ thrash?-T/A:3
 ‘But the younger brother got very angry went to him and thrashed[?]
 him.’ [Text-2, line 10]

Nouns in the possessive *a-* form may appear in any function in the sentence, for example, subject, various kinds of objects.

- (48) (a) *kun arsi-n kun aŋgaj-ən-adɔʔɔŋ gij-an gij-le*
 that monkey-N.SFX that moon-N.SFX-OBJ see-N.SFX see-PST
a-gaɽin-ji-adɔʔɔŋ gam-eten
 3-friend-PL-OBJ tell-T/A:3
 ‘That monkey saw that moon and told her friends.’ [Text-3, line 4]
- (b) *bindɔ a-gaɽin-ji gam-le-ji ian-gamle aŋgaj-ən-adɔʔɔŋ təb-be*
 but 3-friend-PL say-PST-PL how! moon-N.SFX-OBJ take.out-1PL
 ‘“But how can we get the moon out?” her friends said.’ [Text-3, line 7]

Sometimes the use of what appears to be the *a-* possessive prefix does not appear to be motivated or at least typical of its ‘usual’ uses. In some instances, it may be functioning more as a definite marker, as in its use in the word *a-səmma-mə* below.

- (49) *dəjiŋ məŋim ettegoy bara-le-n-ji-a tiki a-səmma-mə*
 several year like.this work-PST-N.SFX/ITR-PL-GEN after 3-young-man
sukku-n a-kako-n-adɔʔɔŋ sarəba-n
 Sukku-N.SFX 3-older.brother-N.SFX=OBJ paddy.field-N.SFX
er-ti-lə-be anson bara-eten
 NEG-give-PST-NF:W/O himself work-T/A:3
 ‘[after] they worked like this for several years, Sukku the younger brother
 cultivated the paddy-field himself without giving his older brother the
 paddy-field.’ [Text-2, line 7]

In the following example, its use is even more curious. The phrase *a-kəndar-leŋ-ən səreŋ* ‘from on the branch’ is a postposition *səreŋ* and a *-leŋ-ən* (locative) case marked noun of a stem *kəndar* ‘branch’. It is not yet clear what motivates this other than possibly the same kind of ‘definite’ marking that seems to characterize its use in the example above.

- (50) *aninji kən gam-ən gam-le ajaŋij-nə a-kəndar-leŋ-ən*
 they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX
səreŋ alan-ji təd-le-n təd-le-n aninji aŋgaj-ən-adɔʔɔŋ
 from [3-]tail-PL hang-PST-NSFX hold-PST-NSFX they moon-N.SFX-OBJ
a-nidəb-ben-asən ɔskai-le-n-ji
 PFX-pick.up-INF-FOR prepare-PST-[ITR/RFLXV]-PL
 ‘They thus discussed and then really began to pick up the moon hanging
 from the branch from each other’s tails.’ [Text-3, line 12]

3.1.4 Definiteness

Definiteness does not have a particular realization *per se* in the morphology of the Sora noun (phrase). The characteristic suffix *-ən* may appear with all kinds of nouns and does not appear to be restricted with regard to definiteness with our current state of understanding. An indefinite specific referent in narratives is usually qualified with the numeral *abɔj* ‘one’.

- (51) (a) *abɔj gailɔ-n dəku-le*
 one/INDEF road-N.SFX COP-PST
 ‘There (once) was a road.’ [Text-1, line 1]
 (b) *kuna gailɔ-n-a jattə abɔj ɔŋa-n dəku-le*
 DEF road-N.SFX-GEN below one ditch/drain-N.SFX COP-PST
 ‘There was a ditch/drain below that road.’ [Text-1, line 2]
 (c) *dɔ bɔ dinna kun a-gailɔ-leŋ abɔj raʔa-n iar-eted*
 so one day DEF OBJ-road-LOC one elephant-N.SFX walk-T/A:3
 ‘So one day an elephant was walking along the road.’ [Text-1, line 9]
 (d) *abɔj tuləb-leŋ-ən dəjeŋ arsi-n-ji dəku-le-ji*
 one forest-LOC-N.SFX several monkey-N.SFX-PL be-PST-PL
 ‘A number of monkeys were in a certain forest.’ [Text-3, line 1]

3.1.5 Class/gender

Class or gender is not an active part of Sora morphosyntax. There are both covert classes seen in the realization of various types of morpho-syntactic phenomena (e.g. the restriction of the ‘oblique/dative object’ element to animates or narratively energized inanimate referents (e.g. ‘the moon’ that is serving as a character in a story)). Relics of noun classes may be found in the prefix that typify Sora noun structure (see section 3.1.10) but the semantics of these remain an object of intensive future study. Gender of masculine/feminine elements can be signalled by various indigenous compounding means [=mar (etc.) ‘man, person’, =boj (female)] as well as in a limited set of loanwords via the Indo-Aryan suffixal contrast *-a* (masculine) *-i* (feminine) – a similar system to those found across the Munda languages. Some words combine the two means *dəŋgəda=mar* and *dəŋgədi=boj* ‘young-man’ and ‘young-woman’, respectively.

3.1.6 Pronouns

The personal pronouns of Gajapati Gumma Sora are offered in (52) below. It is a simple 3×2 system.

(52) Pronouns

	Singular	Plural
1	<i>ɲen</i>	<i>anlen</i>
2	<i>amən</i>	<i>ambəɲ</i>
3	<i>anin</i>	<i>aninji</i>

- (53) (a) *ɲen kulu? jaʔr-an ɲem-t-ay*
I green snake-N.SFX catch-NPST-1
'I am catching the green snake.'
- (b) *ɲen ɲam-yim-ti-n-ay*
I catch-chicken-NPST-RFLXV/ITR-1
'I am chicken-catching.'
- (c) *amən ɲam-yo-ti-ɲ*
you catch-fish-NPST-ITR:2/3
'You are catching fish.'
- (d) *anin ɲam-yo-ti-ɲ*
she/he catch-fish-NPST-ITR/RFLXV:2/3
'She/he is catching fish.'
- (e) *anlen a-ɲəm-yo-ti-n-ay*
we 1/2PL-catch-fish-NPST-ITR-1
'We are catching fish.'
- (f) *ambim á-ɲəm-yó-ti-ɲ*
you (PL) 1/2PL-catch-fish-NPST-ITR:2/3
'You (PL) are catching fish.'
- (g) *aninji ɲam-yo-ti-n-ji*
they catch-fish-NPST-ITR/RFLXV-3PL
'They are catching fish.' [OG]

Reflexive pronouns may be formed with the reflexive *-dəm*.

- (54) *ɔsəntan aninji-dəm tləblubaj-le-ji*
in.vain they-RFLXV drown/CAUS/-PST-PL
'In vain they got themselves drowned.' [Text-3, line 16]

Starosta lists a set of emphatic first and second plural pronouns *anlenji* and *ambenji* (1967:276).

Third person pronouns can be used as definite markers in noun phrases in Sora along with demonstratives, adjectives, and nouns.

- (55) (a) *əntannəɲ kuni anin suʔa-dəd-ən ier-ai-ted*
then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
'Then that one, him, the big-frog, came back.' [Text-1, line 14]
- (b) *aninji sənnə-dəd-ən-ji gam-le-ji amən sireɲ boiboi suʔa*
they small-frog-N.SFX-PL say-PST-PL you from very big
'The small frogs (they) said "much bigger than you".' [Text-1, line 23]
- (c) *anin kəddib kəndəd-ən-ji sireɲ anin suʔa*
he all frog-N.SFX-PL from/than he big
'He is bigger than all the other frogs.' [Text-1, line 7]

Wh-questions in Sora involve a variety of elements. These include *ian* ‘how’ *ete-n* ‘what’ *iando* ‘why’ *ete-n-asən* ‘why’, and *iaŋ* ‘who’.

- (56) (a) *anin ian de-le* (b) *iando tid-t-iŋ*
 he how become-PST why hit-NPST-1
 ‘What happened to him?’ ‘Why are (you) hitting me?’

Interrogative pronouns with the additive focus marker are used as indefinites in Sora as in many other Eurasian languages and indeed many Munda languages, for example, Gta? (Anderson, this volume).

- (57) *ete-jə ti-n*
 what-ADD give-1
 ‘Give me anything!’ (Starosta 1967:288)

3.1.7 Demonstratives

Sora has a developed deictic and demonstrative system like most other Munda languages. Some common elements found include *ənt-* ‘that’ and *kun[a]* ‘this/that’, *kəni* ‘this’, *te?ne* ‘here’, *te?te* ‘there’, etc.

- (58) (a) *əntannəŋ kuni anin suʔa-dud-ən ier-ai-ted*
 then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
 ‘Then that one, him, the big-frog, came back.’ [Text-1, line 14]
- (b) *ənt aiəm-leŋ-ən anin suʔa-dud-ən ət-dəku-əd*
 that time-LOC-N.SFX he big-frog-N.SFX NEG-COP-NEG[PST:3]
 ‘At that time he, the big-frog, wasn’t there.’ [Text-1, line 10]
- (c) *ənte ettele oʔa-leŋ-ən əteŋ kəndud-ən-ji dəku-le*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 ‘Like that there were many frogs in the ditch.’ [Text-1, line 4]
- (d) *əntəpsele anlen kən begiʔa a-sə-le-n-ai*
 therefore we DEF different.place 1PL-[go.]hide-PST-CLOC/1PL
 ‘Therefore we went and hid in (the[se]) different places.’ [Text-1, line 19]
- (e) *kuna gailə-n-a jattə abəj oʔa-n dəku-le*
 DEF road-N.SFX-GEN below one ditch/drain-N.SFX COP-PST
 ‘There was a ditch/drain below that road.’ [Text-1, line 2]
- (f) *kun asən kun sənna-dud-ən-ji raʔa-n-advəŋ giŋ-an*
 DEF for DEF small-frog-N.SFX-PL elephant:N.SFX-OBJ see-N.SFX
giŋ-le bətəŋ-le iersed-le-ji
 see-PST be.frightened-PST run.away-PST-PL
 ‘Because of seeing the elephant, the small frogs were frightened and ran away.’ [Text-1, line 12]
- (g) *əŋale kun gəʔan-ə mənɗra:ji ə-lo-n-a mənəŋ oʔe*
 how DEF TOWN-GEN person-PL PFX-ROAD-N.SFX-GEN beside or
gailun-a mənəŋ oʔa-n daku
 river-N.SFX:GEN beside ditch/drain-N.SFX be[come]
 ‘How it is is that the town people have a drain/ditch either beside a river or by a road.’ [Text-1, line 3]
- (h) *də bə dinna kun a-gailə-leŋ abəj raʔa-n iar-eted*
 so one day DEF OBJ-ROAD-LOC one elephant-N.SFX walk-T/A:3
 ‘So one day an elephant was walking along the road.’ [Text-1, line 9]

- (i) *kun arsi-n* *kun anɣaj-ən-adɔʔɔŋ* *gij-an gij-le*
 that monkey-N.SFX that moon-N.SFX-OBJ see-N.SFX see-PST
a-ɣaʔin-ji-adɔʔɔŋ *gam-eten*
 3-friend-PL-OBJ tell-T/A:3
 ‘That monkey saw that moon and told her friends.’ [Text-3, line 4]

Starosta (1967:203–204) lists the following demonstrative forms for Sora using the elements *-ne* (corresponding to *kən*) and *-te* (corresponding to *kun*) in a roughly proximal/distal opposition.

- (59) *ɛʔne* ‘like this’ *ɛʔte* ‘like that’
ɛʔnegɔʔ ‘like this, this way’ *ɛʔtegɔʔ* ‘like that, that way’
dəkiyne ‘this (little)’ *dəkiyte* ‘that (little)’
dəkəʔne ‘this (big)’ *dəkəʔte* ‘this (big)’
teʔne ‘here’ *teʔte* ‘there’
areʔne ‘around here’ *səleʔte* ‘at that time’

3.1.8 Numerals

Sora employs a base-12, base-20 system that is productive up into the thousands. It shows some signs of restructuring or variation. The following sets were collected in March 2007, from Oruncho Gomango, aged 50 in Gumma block, Gajapati district, Orissa, India.

- (60)
- | | | | |
|---------------------------|------|--------------------------|--------|
| <i>aboy</i> | ‘1’ | <i>bago</i> | ‘2’ |
| <i>yagi</i> | ‘3’ | <i>unji</i> | ‘4’ |
| <i>mɔnlɔy</i> | ‘5’ | <i>tudru</i> | ‘6’ |
| <i>gulji</i> | ‘7’ | <i>tʰamji</i> | ‘8’ |
| <i>tinji</i> | ‘9’ | <i>gelji</i> | ‘10’ |
| <i>gelmuy</i> | ‘11’ | <i>migel</i> | ‘12’ |
| <i>migelboy</i> | ‘13’ | <i>[migelbagu]</i> | [‘14’] |
| <i>migelyagi</i> | ‘15’ | <i>migelunji</i> | ‘16’ |
| <i>migelmonloy</i> | ‘17’ | <i>migeltudru</i> | ‘18’ |
| <i>migelgulji</i> | ‘19’ | <i>bokuri</i> | ‘20’ |
| <i>bokuri aboy</i> | ‘21’ | <i>bokuri bagu</i> | ‘22’ |
| <i>bokuri yagi</i> | ‘23’ | <i>bokuri unji</i> | ‘24’ |
| <i>bokuri monloy</i> | ‘25’ | <i>bokuri tudru</i> | ‘26’ |
| <i>bokuri gulji</i> | ‘27’ | <i>bokuri tʰamji</i> | ‘28’ |
| <i>bokuri tinji</i> | ‘29’ | <i>bokuri gelji</i> | ‘30’ |
| <i>bokuri gelmuy</i> | ‘31’ | <i>bokuri migel</i> | ‘32’ |
| <i>bokuri migelboy</i> | ‘33’ | <i>bokuri migelbagu</i> | ‘34’ |
| <i>bokuri migelyagi</i> | ‘35’ | <i>bokuri migelunji</i> | ‘36’ |
| <i>bokuri migelmonloy</i> | ‘37’ | <i>bokuri migeltudru</i> | ‘38’ |
| <i>bokuri migelgulji</i> | ‘39’ | <i>bakuri</i> | ‘40’ |
| <i>bakuri aboy</i> | ‘41’ | <i>bakuri bagu</i> | ‘42’ |
| <i>bakuri yagi</i> | ‘43’ | <i>bakuri unji</i> | ‘44’ |
| <i>bakuri monloy</i> | ‘45’ | <i>bakuri tudru</i> | ‘46’ |
| <i>bakuri gulji</i> | ‘47’ | <i>bakuri tʰamji</i> | ‘48’ |

<i>bakuri tinji</i>	'49'	<i>bakuri gelji</i>	'50'
<i>bakuri gelmuy</i>	'51'	<i>bakuri migel</i>	'52'
<i>bakuri migelboy</i>	'53'	<i>bakuri migelbagu</i>	'54'
<i>bakuri migelyagi</i>	'55'	<i>bakuri migelunji</i>	'56'
<i>bakuri migelmonloy</i>	'57'	<i>bakuri migeltudru</i>	'58'
<i>bakuri migelgulji</i>	'59'	<i>yakuri</i>	'60'
<i>yakuri aboy</i>	'61'	<i>yakuri bagu</i>	'62'
<i>yakuri yagi</i>	'63'	<i>yakuri unji</i>	'64'
<i>yakuri monloy</i>	'65'	<i>yakuri tudru</i>	'66'
<i>yakuri gulji</i>	'67'	<i>yakuri t^hamji</i>	'68'
<i>yakuri tinji</i>	'69'	<i>yakuri gelji</i>	'70'
<i>yakuri gelmuy</i>	'71'	<i>yakuri migel</i>	'72'
<i>yakuri migelboy</i>	'73'	<i>yakuri migelbagu</i>	'74'
<i>yakuri migelyagi</i>	'75'	<i>yakuri migelunji</i>	'76'
<i>yakuri migelmonloy</i>	'77'	<i>yakuri migeltudru</i>	'78'
<i>yakuri migelgulji</i>	'79'	<i>unjikuri</i>	'80'
<i>unjikuri aboy</i>	'81'	<i>unjikuri bagu</i>	'82'
<i>unjikuri yagi</i>	'83'	<i>unjikuri unji</i>	'84'
<i>unjikuri monloy</i>	'85'	<i>unjikuri tudru</i>	'86'
<i>unjikuri gulji</i>	'87'	<i>unjikuri t^hamji</i>	'88'
<i>unjikuri tinji</i>	'89'	<i>unjikuri gelji</i>	'90'
<i>unjikuri gelmuy</i>	'91'	<i>unjikuri migel</i>	'92'
<i>unjikuri migelboy</i>	'93'	<i>unjikuri migelbagu</i>	'94'
<i>unjikuri migelyagi</i>	'95'	<i>unjikuri migelunji</i>	'96'
<i>unjikuri migelmonloy</i>	'97'	<i>unjikuri migeltudru</i>	'98'
<i>unjikuri migelgulji</i>	'99'	<i>bo sua</i>	'100'
<i>bosua migel</i>	'112'	<i>bagusualbagusoa</i>	'200'
<i>monloysua</i>	'500'	<i>bo ajar</i>	'1000'

No plural form of the noun is used after a numeral in Sora but it may trigger (semantic) plural verb agreement.

- (61) *dijen-antide sua sanna-dud-ən daku-le-ji*
 how.many 100 little-frog-N.SFX COP-PST-PL
 'How many hundreds of little frogs there were!' [Text-1, line 5]

Numeral stems themselves are commonly either found in nominal compounds, or may be nominalized and themselves take plural marking.

- (62) (a) *bagu-mər-an-ji əmele kata-n-a+ber*
 two person-N.SFX-PL about talk-N.SFX-GEN+word
 'The story of two people' [Text-2, line 1]
 (b) *aninji bagu-n-ji blən/lyənə daku-le-ji*
 they two-N.SFX-PL clan/NMLZR/clan COP-PST-PL
 'There were two brothers (of one clan).' [Text-2, line 2]

Unlike many other Munda languages, there does not appear to be a difference, or at least not a rigidly maintained distinction, between animate and inanimate forms of the word 'one'; for example, both are *aboj/abəj* in the sentences below.

- (63) (a) *aboj tuləb-leŋ-ən dəjeŋ arsi-n-ji dəku-le-ji*
 one forest-LOC-N.SFX several monkey-N.SFX-PL be-PST-PL
 ‘A number of monkeys were in a certain forest.’ [Text-3, line 1]
- (b) *də bə dinna kun a-gailə-leŋ aboj raʔa-n*
 so one day DEF OBJ-road-LOC one elephant-N.SFX
iar-eted
 walk-T/A:3
 ‘So one day an elephant was walking along the road.’ [Text-1, line 9]

3.1.9 Adpositions

A range of variably clitic or free-standing postpositional elements are characteristic of the Sora nominal system. Some of these appear with the noun in the basic, unmarked form, while with others, so-called relational nouns, the noun complement appears in a dependent/genitive form.

Some examples of postpositions taking the basic form of the noun, which is the largest class by far, include the following:

- (64) (a) *əntannəŋ kuni anin suʔa-dud-ən ier-ai-ted*
 then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
 ‘Then that one, him, the big-frog, came back.’ [Text-1, line 14]
- (b) *jen je'lu-n a'san boiʔəŋsə'ləʔ ə'puŋ-l-aiʔ*
 I meat-N.SFX for woman tell-PST-1
 ‘I told the woman about the meat.’
- (c) *jem 'boʊnsəloʊ ba'tiy 'tuʔib-ən 'yeʔ-ʔ-er*
 I woman with forest-N.SFX go-PST-1
 ‘I went to the forest with the woman.’
- (d) *tuʔib-ən seʔŋ 'boʊnsəloʊ ba'tiy jen 'yeʔ-ʔ-ar*
 forest-N.SFX from woman with I go-PST-CLOC/1
 ‘I came from the forest with the woman.’
- (e) *'eŋjuwim ba'tiy aʔna'doŋ=e-l-l-ay*
 axe with tree-N.SFX-OBJ=cut-PST-1
 ‘I cut down the tree with an axe.’ [OG]
- (f) *bagu-məʔ-an-ji əmele kata-n-a+ber*
 two person-N.SFX-PL about talk-N.SFX-GEN+word
 ‘The story of two people’ [Text-2, line 1]
- (g) *aninji kən gam-ən gam-le aʔaʔij-nə a-kəndar-leŋ-ən*
 they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX
səreŋ alan-ji ted-le-n ted-le-n aninji
 from [3-]tail-PL hang-PST-N.SFX hold-PST-N.SFX they
aŋgaj-ən-adəʔəŋ a-ŋidəb-ben-asən əskai-le-n-ji
 moon-N.SFX-OBJ PFX-pick.up-INF-PURP prepare-PST-[ITR/RFLXV]-PL
 ‘They thus discussed and then really began to pick up the moon hanging from the branch from each other’s tails.’ [Text-3, line 12]

On the other hand, some relational noun adpositions prefer their noun complement to be in the genitive case form.

- (65) (a) *kun a-gailɔ-n-a jattə abɔj ɔʔa-n dəkɔ-lɛ*
 DEF PFX-ROAD-N.SFX-GEN below one ditch/drain-N.SFX COP-PST
 ‘There was a ditch/drain below that road.’ [Text-1, line 2]
- (b) *eɲale kun gəʔan-ə məndra:-ji ə-lo-n-a mənɛŋ ɔʔe*
 how DEF TOWN-GEN person-PL PFX-ROAD-N.SFX-GEN beside or
gailu-na mənɛŋ ɔʔa-n daku
 river-N.SFX:GEN beside ditch/drain-N.SFX be[come]
 ‘How is it that the town people have a drain/ditch either beside a river
 or by a road?’ [Text-1, line 3]

This pattern is also seen with subordinate clause complements of various types, thus certain clausal subordinators appear to be a sub-type of adpositional element (relational nouns (i.e. ones that take genitive complements) that take clausal complements).

- (66) *bar bəndi-lə-n-ji-na tiki dəjin dinna de-le anin*
 and put.in.jail-PST-N.SFX-PL-GEN after few day become-PST he
a-dukəʔɪ-ŋ dɔ a-on-ən-ji jinaŋ doləjən batte kəniet-lə-ji
 3-wife-N.SFX and 3-child-N.SFX-PL also hunger-N.SFX SOC/INS die-PST-PL
 ‘And a few days after they put him in jail, his wife and children also died of
 starvation.’ [Text-2, line 13]

There are also the case-like adpositions which appear to be a class of relational nouns. This includes the oblique object marker *dɔ* [ʔɔ] *ŋ* the adessive *maŋ*, etc. see section 3.1.2.

3.1.10 Derivation

As alluded to above, Sora makes extensive use of root/stem-compounds and lexicalized derivational elements in the creation of its nominal lexicon. At the heart of this is the use of combination with a set of largely monosyllabic combining forms or nominal roots and a host of compounding and derivational processes that serve to derive inflectable free-standing syntactic elements or words/noun phrases. All South Munda languages show this system to some degree but it is most pronounced in Sora and Gta? (see Anderson, this volume).

In Sora, the means used to derive the syntactically free-standing full forms from their corresponding combining forms include reduplication, prefixation, infixation, suffixation, and compounding. In a small number of instances there are suppletive combining form/free form sets (A. Zide 1976, Starosta 1992). A discussion of this is found in Anderson (2007), on which the following is based. One common means of deriving free forms of nouns from monosyllabic roots in Sora was via prefixation of an original syllabic nasal, synchronically realized as [ə] (67).

- (67) ə- < *N-
- | Full form | Combining form | Gloss |
|-------------|----------------|-------------|
| <i>əday</i> | = <i>day</i> | ‘bee-hive’ |
| <i>ələb</i> | = <i>leb</i> | ‘wild-goat’ |
| <i>əsoŋ</i> | = <i>soŋ</i> | ‘dung’ |
- (Starosta 1992:85–86, Ramamurti 1931:69ff.)

Another common prefixal element that was used to derive full forms from corresponding monosyllabic (combining form) elements was **kVN-*, where V is either *-i-* or *ə* (sometimes *u* assimilated to a following *u*).

- (68) **kVN-*
- | Full form | Combining form | Gloss |
|---------------|----------------|-----------|
| <i>kinsod</i> | = <i>sod</i> | 'dog' |
| <i>kəmbud</i> | = <i>bud</i> | 'bear' |
| <i>kəntuj</i> | = <i>tuj</i> | 'owl' |
| <i>kimpuy</i> | = <i>puy</i> | 'stomach' |
- (Starosta 1992:85–86, Ramamurti 1931:69ff.)

A small number of other prefixes may be evinced when comparing Sora combining forms and free forms. These include **VN-* realized as a non-high vowel and a usually assimilating nasal, *on-* or *u-*.

- (69) (a) **VN-*
- | Full form | Combining form | Gloss |
|--------------|----------------|--------|
| <i>enjum</i> | = <i>jum</i> | 'axe' |
| <i>aŋgaj</i> | = <i>gaj</i> | 'moon' |
- (b) *on-*
- | Full form | Combining form | Gloss |
|-----------------|----------------|--------|
| <i>on[d]rej</i> | = <i>rej</i> | 'rat' |
| <i>ontid</i> | = <i>tid</i> | 'bird' |
- (c) *u-*
- | Full form | Combining form | Gloss |
|-------------|----------------|-------------------------------|
| <i>uab</i> | = <i>ab</i> | 'vegetable' |
| <i>umud</i> | = <i>mud</i> | 'smoke' |
| <i>usal</i> | = <i>sal</i> | 'skin' (Ramamurti 1931:69ff.) |

A very small number of nouns seem to be compounds of the noun–noun shape, where the dominant root element (i.e. the one used as a combining form) is the second member. These include the following:

- (70) *X-√*
- | Full Form | Combining Form | Gloss |
|---------------|----------------|----------------------------------|
| <i>bomaŋ</i> | = <i>maŋ</i> | 'chameleon' |
| <i>gorzaŋ</i> | = <i>zaŋ</i> | 'village' (Ramamurti 1931:69ff.) |

Another means of deriving bimoraic/bisyllabic full forms of nouns in Sora from corresponding monomoraic/monosyllabic combining forms is reduplication (71). This takes the shape of a CV(C) copy of the stem.

- (71) **Reduplication-X*
- | Full form | Combining form | Gloss |
|---------------|----------------|-----------------------------------|
| <i>saysaŋ</i> | = <i>saŋ</i> | 'turmeric' |
| <i>tujtuj</i> | = <i>tuj</i> | 'star' |
| <i>tittin</i> | = <i>tin</i> | 'tamarind' (Ramamurti 1931:69ff.) |

Inflection processes are commonly made use of in deriving bisyllabic/bimoraic free forms of nouns. One infix of relatively high frequency in Sora is *-ʔ-*.

(72) *-ʔ-

Full form	Combining form	Gloss
<i>daʔa</i>	= <i>da</i>	‘water’
<i>jeʔeŋ</i>	= <i>jeŋ</i>	‘leg’
<i>raʔa</i>	= <i>ra</i>	‘elephant’
<i>oʔon</i>	= <i>on</i>	‘child’
<i>suʔuŋ</i>	= <i>suŋ</i>	‘house’
<i>moʔod</i>	= <i>mad</i>	‘eye’

(Starosta 1992:85–86, Ramamurti 1931:69ff.)

Another common infix in Sora is *-ən-*. This marks primarily, though not exclusively, instrument nouns (at least synchronically).⁶

(73) *-ən-

Full form	Combining form	Gloss
<i>jəno</i>	= <i>jo</i>	‘broom’
<i>kənuŋ</i>	= <i>kuŋ</i>	‘razor’
<i>pənad</i>	= <i>pad</i>	‘latch’
<i>sənəŋ</i>	= <i>səŋ</i>	‘door’

(Starosta 1992:85–86, Ramamurti 1931:69ff.)

Infix elements of various lexically restricted sorts may be found as well (e.g. *-l-*, *-d-*, *-a-*, or *-s-*) in order to derive free-standing full-forms from corresponding combining root forms of nouns in Sora.

(74)

Full form	Combining form	Gloss
<i>bəled</i>	= <i>bed</i>	‘feathers, plume’
<i>kədib</i>	= <i>kib</i>	‘sword’
<i>ruəŋ</i>	= <i>ruŋ</i>	‘sky’
<i>bəsədlbəsud</i>	= <i>bud</i>	‘salt’
<i>bisiŋ</i>	= <i>biŋ</i>	‘district chief’ (Ramamurti 1931: 69ff.)

True suffixed forms do not appear to be overly common. Some recurrent elements have been noted, which may constitute suffixes used to derive phonotactically acceptable free forms of nouns from their monosyllabic combining forms. These include *-al*, *-eŋ*, and *-en*. For more on this, see below and Anderson and Zide (2002), Anderson (2004b).

(75) (a) $\sqrt{-al}$

Full form	Combining form	Gloss
<i>taŋal</i>	= <i>taŋ</i>	‘crocodile’
<i>aŋəl</i>	= <i>aŋ</i>	‘fuel’

(b) $\sqrt{-eŋ}$

Full form	Combining form	Gloss
<i>dareŋ</i>	= <i>dar</i>	‘horn’
<i>dereŋ</i>	= <i>der</i>	‘horn’

(c) $\sqrt{-en}$

Full form	Combining form	Gloss
<i>raŋen</i>	= <i>raŋ</i>	‘wind’
(~ <i>riŋen</i>)	= <i>riŋ</i>)	‘wind’ (Ramamurti 1931:69ff.)

Much more common is root–root compounding with the dominant root – the one that appears as the combining form – being the first element.

(76) *√-X forms (compounds and suffixed forms)

Full form	Combining form	Gloss
<i>benta</i>	= <i>bēn</i>	'hunting'
<i>boŋtel</i>	= <i>boŋ</i>	'buffalo'
<i>buŋsaŋ</i>	= <i>buŋ</i>	'cattle trough'
<i>daŋgu</i>	= <i>daŋ</i>	'stick'
<i>daŋki</i>	= <i>daŋ</i>	'pot'
<i>darej</i>	= <i>dar</i>	'rice'
<i>sambi</i>	= <i>sam</i>	'buttocks'
<i>tember</i>	= <i>tem</i>	'rat'
<i>saŋka</i>	= <i>saŋ</i>	'neck'
<i>eŋra</i>	= <i>eŋ</i>	'cucumber'

(Starosta 1992:85–86, Ramamurti 1931:69ff.)

Phonological variants of certain nouns may account for some seemingly anomalous full form: combining form correspondence sets. The general lack of **ŋ* in syllable onset position in Sora (with a small number of lexical exceptions) blocks the surface realization of the first combining form below, while the second form preserves the initial-[s] in the combining form that has been lost in the corresponding free form.

Full form	Combining form	Gloss
<i>əŋal</i>	= <i>al</i>	'clearing on field' (< * <i>əŋal</i> -)
<i>ali</i>	= <i>sal</i>	'liquor' (Ramamurti 1931:69ff.)

Suppletive forms are also found, where the combining form and the full form bear no relationship to each other phonologically or morphologically, but rather constitute a set of words that form a synchronic paradigm of heterogeneous diachronic origin (of the *gol/went* sort).

(78) (a) Full form	Combining form	Gloss
<i>rogo</i>	= <i>san</i>	'red-gram'
<i>bati</i>	= <i>pud</i>	'mushroom'
<i>ənselo</i>	= <i>boi</i>	'woman' (Ramamurti 1931:43, 69ff.)

Note that in Sora numeral and adjectives/adjectivals may combine in a stem compound {Num-N}/{Adj-N} rather than appear in a Numeral [Num N] or Adjectival Phrase [Adj N].

(78) (b)	<i>ajaŋid</i>	<i>po</i>	<i>gam-le</i>	<i>anin</i>	<i>suŋa-mər</i>
	true	DOUBT	say-PST	he	big-person
	'“Is that really true?” he, the big guy said.' [Text-1, line 24]				
(c)	<i>dɪ</i>	<i>turkaŋiŋ-ən</i>	<i>anin-adɔʔəŋ</i>	<i>bəndi-lə-ji</i>	
	DISC	jail-N.SFX	he-OBJ	put.in.jail-PST-PL	
	'They put him in jail.' [Text-2, line 12]				

Complex derivational masses can arise through a concatenation of words that typifies many Munda languages (so-called tag and echo words, in addition to the elaborate compounding structures of Sora) each with multiple derivational elements. Note in this regard the following monster noun form: *j-ən-r-om+g-ən-r-a-n-ji* 'food supplies' from *jom+ga* 'eat meal' and two infixes each in the verbal stems *-ən-* and *-r-* (Starosta 1967:66). More than one way of deriving nouns from, for example, a single verb stem with differences in meaning, may also be seen in a limited number of instances

in Sora, for example, from *ga* ‘drink, eat’ we find *g-ən-a-ga-n* ‘potables’ and *ga-ga-n-ən* ‘eating’ (Starosta 1967:215).

3.1.11 Adjectives

Adjective as an independent word class has yet to be justified in Sora. Many adjectival forms in Sora appear in nominal compounds as phrase-words rather than syntactic phrases. Other adjectives in Sora appear pre-nominally in modificational function. They may also be used predicatively with a zero-copula in NPST formations.

- (79) (a) *dijey-antide sua sənna-dud-ən dəku-le-ji*
 how.many 100 little-frog-N.SFX COP-PST-PL
 ‘How many hundreds of little frogs there were!’ [Text-1, line 5]
- (b) *anin suḡa-dud-ən a-ber-n-an etente beren aninji*
 he big-frog-N.SFX PFX-say-n-N.SFX what say-N.SFX they
sənna-dud-ən-ji mane-le-ji
 small-frog-N.SFX-PL obey-PST-PL
 ‘Whatever the big-frog said, the small frogs obeyed, had to obey.’
 [Text-1, line 8]
- (c) *dəjij meḡim etteḡoy bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX/ITR-PL-GEN after
a-sənna-mər sukku-n a-kako-n-adəʔəḡ
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ
sarəba-n er-ti-lə-be anson bara-eten
 paddy.field-N.SFX NEG-give-PST-NF:W/O himself work-T/A:3
 ‘[after] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.’ [Text-2, line 7]

Comparatives are formed by the construction: {Noun *sireḡ* [qualifier] Adjective}, with *sireḡ* being the ablative postposition ‘from’.

- (80) *aninji sənna-dud-ən-ji gam-le-ji amən sireḡ boiboi suḡa*
 they small-frog-N.SFX-PL say-PST-PL you from very big
 ‘The small frogs (they) said, “much bigger than you”.’ [Text-1, line 23]

Superlatives are built off this with the construction: {*kuddib* Noun *sireḡ* [qualifier] Adjective}, where *kuddib* means ‘all’.

- (81) *anin kuddib kəndud-ən-ji sireḡ anin suḡa*
 he all frog-N.SFX-PL from he big
 ‘He is bigger than all the other frogs.’ [Text-1, line 7]

3.1.12 Adverb(ial)s

Adverbials in Sora are uninflecting elements that modify verbal actions and stand in a pre-verbal position preferentially. Many occur in clause-initial position.

- (82) (a) *uan-ji kəniət-le-n-a tiki baḡu-n-ji*
 father-3PL die-PST-N.SFX-GEN after two-N.SFX-PL

- jənnəŋ sərɔba-n-ji mailen bara-le-ji*
 field paddy-N.SFX-PL together work-PST-PL
 ‘After their father died, they both worked in their fields and paddies together.’ [Text-2, line 16]
- (b) *tikki aninji ɔləŋ-le-n-ji dɔ etenasən a-sɔ-le-n*
 afterwards they answer-PST-[ITR]-PL SO why 2PL-hide-PST-ITR
gam-le suʔa-dud-ən
 say-PST big-frog-N.SFX
 ‘After that they answered, and the big-frog said “why did you hide?”’
 [Text-1, line 17]
- (c) *əntəpsəle anlen kən begiʔa a-sɔ-le-n-ai*
 therefore we DEF different.place 1PL-[go.]hide-PST-ITR-1PL
 ‘Therefore we went and hid in (the[se]) different places.’ [Text-1, line 19]
- (d) *bijndɔ a-ubban bɔibɔi barab-le ier-an*
 but OBJ-younger.brother very get.angry-PST go-N.SFX
ier-le anin adɔʔəŋ tuəb-eten
 go-PST he OBJ thrash?-T/A:3
 ‘But the younger brother got very angry went to him and thrashed[?] him.’ [Text-2, line 10]

Nouns in Sora, often in the suffixed *-ən* form, may appear in an adverbial function as well, for example, certain common temporal nouns.

- (83) *aninji tamba-n annəŋ jun-le ga-le tɔgəl-ən annəŋ*
 they noon-N.SFX during eat-PST drink-PST at.night-N.SFX during
aboy mənəŋ bənda-n a-bo ara-leŋ-ən diməd-le-n-ji
 one edge tank-N.SFX OBJ-one tree-LOC-N.SFX sleep-PST-ITR-PL
 ‘They would eat and drink during the day and at night they would sleep in a tree in one corner of a tank.’ [Text-3, line 2]

Some examples of Sora adverbials in our corpus include:

- (84) Wh-words
uan a-ier-re ənn-ɔləŋ-n-ed-ji
 where 2PL-go-PST NEG-answer-n-NEG-PL
 ‘“Where did you go?”; they didn’t answer him.’ [Text-1, line 16]
- (85) Time adverbials
nam ‘now’ (Ramamurti 1931:53)
- (86) Reduplicated time adverbials
ʔaŋaʔaŋa ‘sometimes’
- (87) Deictic day names (Ramamurti 1931:53)
ʔanaŋamam ‘today’
ʔbiyo ‘tomorrow’
ŋeramme, ən(əb)ʔbiyo ‘day after tomorrow’
ʔrbən ‘yesterday’
ʔrubənʔta:ŋ, moyed ‘day before yesterday’
ʔya:gi ʔdinaʔdele ‘three days ago’

Some adverbials formally appear fully reduplicated in Sora. A special class of these is treated in section 3.3.

- (88) *'moyed* *'day before yesterday', 'previous'*
moyed 'moyəd *'recently' (Ramamurti 1931:178)*

3.2 Verbal morphology

The verbal morphology of Sora is complex to be sure. It is fairly straightforward from a pan-Munda comparative perspective, although as a likely early off-shoot of Proto-Munda, it has certain important idiosyncratic and individuating characteristics (see Anderson 2007 for more).

3.2.1 Subject

The subject markers in Sora for one class of verbs are as follows:

- (89) SG PL
 1 *-ay* *a-...-ay*
 2 *-e[y]* *a-e[y]*
 3 *-e[y]* *-ji*
- (90) (a) *giʔy-t-ay* (b) *giʔy-t-ey* (c) *anij giʔy-t-ey*
 see-NPST-1 see-NPST-2/3 see-NPST-2/3
 'I see' 'you see' 'she/he sees'
- (d) *ənlen a-giʔy-t-ay* (e) *amben a-giʔ-t-ey* (f) *aninji giʔ-te-ji*
 we 1/2PL-see-NPST-1 you 1/2PL-see-NPST-2/3 they see-NPST-3PL
 'we see' 'you (PL) see' 'they see' [OG]

A noun may be marked with the plural suffix but the verb lacks it in Sora.

- (91) *əntɛ ɛttele oʃa-leŋ-ən əteŋ kəndud-ən-ji dəku-le*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 'Like that there were many frogs in the ditch.' [Text-1, line 4]

The reverse is also true, where a semantically plural noun (e.g. the complement of a numeral over one if semantically animate) may trigger plural verb agreement even if formally lacking the plural marker itself.

- (92) *dijeŋ-antide sua sənna-dud-ən dəku-le-ji*
 how.many 100 little-frog-N.SFX COP-PST-PL
 'How many hundreds of little frogs there were!' [Text-1, line 5]

Naturally both are possible as well, with both subject noun and verb marked with the plural suffix/enclitic *ji*.

- (93) *aninji sənna-dud-ən-ji gam-le-ji amən sireŋ boiboi suʃa*
 they small-frog-N.SFX-PL say-PST-PL you from very big
 'The small frogs (they) said "much bigger than you".' [Text-1, line 23]

First plural subject marking in Sora is either prefixal or circumfixal, consisting of a prefix *a-* and suffix/enclitic *-ai/ay*, also found with first singular subjects.

- (94) (a) *əntəpsɛle anlen kən begiɾa a-sɔ-le-n-ai*
 therefore we DEF different.place 1PL-[go.]hide-PST-ITR-1PL
 ‘Therefore we went and hid in (the[se]) different places.’ [Text-1, line 19]
- (b) *ənlen daʔa-n a-tiy-t-ay*
 we water-N.SFX 1PL-give-NPST-1
 ‘We give (him/her) water.’ [OG]

First plural subjects with second singular objects drop the *-ai/ay* part of the marker and replace that part of the verbal word-template with the second singular object marker *-am*. They remain one of the few bi-personal verb forms in the language.

- (95) *g-tiy-t-am*
 1PL-give-NPST-[1>]2
 ‘we give you’ [OG]

With first singular subjects and second singular objects, this suppression means the forms agree with their objects only.

- (96) (a) *ɲen daʔa-n tiy-t-am*
 I water-N.SFX give-NPST-[1>]2
 ‘I (will) give you water.’ [OG]
- (b) *rban ɲen daʔa-n tiy-l-am*
 yesterday I water-N.SFX give-PST-[1>]2
 ‘I gave you water yesterday.’ [OG]

Second plural subject has the prefix *a-*; it may have a circumfixal suffix component but the suffix is either zero or *-e[y]*, whatever the second/third singular form is.

- (97) (a) *uan a-ier-re ənn-ɔləŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-n-NEG-PL
 ‘“Where did you go?”; they didn’t answer him.’ [Text-1, line 16]
- (b) *tikki aninji ɔləŋ-le-n-ji dɔ etenasən a-sɔ-le-n*
 after they answer-PST-[ITR]-PL so why 2PL-hide-PST-ITR
gam-le suɾa-dud-ən
 say-PST big-frog-N.SFX
 ‘After that they answered, and the big-frog said “why did you hide?”’
 [Text-1, line 17]

There are at least two other conjugational types (more if you include ones that obligatorily use *-n-* as well) in Sora based on the system of subject inflection. One has the suffixal agreement marker *-be* (*~ -biy*), for example, in the first plural.

- (98) (a) *aninji po pɔisa-n paŋ-ai-ji dɔ gam-be*
 they Q/DOUBT money-N.SFX carry-CLOC-PL DISC say-1PL
 ‘We say “will they bring the money?”’ (Starosta 1967:146)
- (b) *'byoɾiy al'len ar¹-gi[?]-tə-biy* (c) *al'len ar¹-gi[?]-tə-biy*
 tomorrow we RECIP-see-NPST-1PL we RECIP-see-NPST-1PL
 ‘Tomorrow we will see each other.’ ‘We see ourselves.’ [OG]

Another set has what appear to be object markers used as subject markers. These are like undergoer subject markers.

- (99) *anin boiboi barab-l-ij*
 He very get.angry-PST-1
 ‘I got very angry at him.’ (Starosta 1967:109)

Sometimes impersonal formations are encountered and there is no agreement with the logical ‘subject’.

- (100) *jen ete-n-a-gam-ben de-te*
 I what-N.SFX-DEP-say-INF AUX-NPST
 ‘What should I say?’ (Starosta 1967:109)

3.2.2 Object types

Sora belongs to the group of languages that encodes the verbal object properties within the verbal word form itself. The person/number features of a primary object in the Dryer (1986) sense may be encoded suffixally in the Sora verbal word-template following the tense marker. They can appear in forms that lack tense markers as well, for example, negative forms and imperatives.

- (101) (a) *iando tid-t-ij*
 why hit-NPST-1
 ‘Why are (you) hitting me?’ (Starosta p.c.)
- (b) *anlen er-gənij-ji*
 we NEG-see/NEG/see-3PL
 ‘We didn’t see them.’ (Starosta 1967:277)
- (c) *rban jen daʔa-n tiy-l-am*
 yesterday I water-N.SFX give-PST-[1>]2
 ‘I gave you water yesterday.’ [OG]
- (d) *jen daʔa-n tiy-t-am*
 I water-N.SFX give-NPST-[1>]2
 ‘I (will) give you water.’ [OG]
- (e) *jen rban daʔa-n tiy-l-ij*
 I yesterday water-N.SFX give-PST-1
 ‘You gave me water yesterday.’ [OG]
- (f) *daʔa-n tiyʔ-t-ij*
 water-N.SFX give-NPST-1
 ‘You (will) give me water.’ [OG]
- (g) *ubban-adʔəŋ gam-etan amən etenasən*
 younger.brother-OBJ say-T/A you why
səɾɔba-n ət-ti-ŋ
 paddy.field-N.SFX NEG-give-1
 ‘He said to his younger brother “why (do) not (you) give me some paddy field.”’ [Text-2, line 9]
- (h) *sen-ta-ben*
 thank-NPST-2PL
 ‘thank you (PL)’ [LG]
- (i) *aninji sənna-mər-ən-ji əgəndi-le-ji etteden anlen*
 they little-person- N.SFX-PL think-PST-PL therefore we
sɔənja-n ət-təɾe-də jitəŋkə-n
 grain.tribute-N.SFX NEG-measure-NPST tax-N.SFX

ad-ji-e *gam-le* *jore-lə-len* *bij*
 NEG-give-NPST COMP twist-PST-1PL MOD
 ‘The little guys thought “we won’t measure out the grain-tribute and
 won’t pay the tax because he oppressed us.”’ [Text-1, line 32]

The only bi-personal verb forms in Sora are second person (singular/plural) object forms with first plural and third plural subjects, second plural subjects with first person (singular/plural) and third plural objects, and first plural subject with third plural or second person (singular/plural) object in negative (past) conjugations (for some speakers at least).

- (102) (a) *aninji rban daʔa-n a-tiy-l-əm-ji*
 they yesterday water-NN.SFX NEG-give-PST-2-3PL
 ‘Yesterday they didn’t give you water.’ [OG]
 (b) *ənlen aman daʔa-n ʔʔ-tiy-t-am*
 we you water-N.SFX 1PL-give-NPST-2
 ‘We give you water.’ [OG]
 (c) *ənlen daʔa-n a-tiy-t-ay*
 we water-N.SFX 1PL-give-NPST-1
 ‘We give (him/her) water.’ [OG]
 (d) *əʔ-gij-l-ə-be-ji*
 NEG-see-PST-1PL-3PL
 ‘We didn’t see them.’ (Starosta 1967:276)

An incorporated noun in the verbal stem allows the possessor of that verbal stem to be expressed in the verbal form in Sora. This is a relatively common feature in languages worldwide (Anderson 1997).

- (103) (a) *le:m-si:-t-am*
 bow-hand-NPST-2
 ‘I bow to your hand.’ (Ramamurti 1931:43)
 (b) *kuŋ-bəb-t-əm*
 shave-head-NPST-2
 ‘Your head is shaven.’ (Biligiri 1965:240)
 (c) *pokunpoŋ-am-te:n*
 stab:belly:knife-2-3:PST
 ‘(who) stabbed you in the belly?’ (Ramamurti 1931:25)

Note that object encoding in the verbal form is not obligatory in Sora and may be omitted or suppressed.

- (104) *ier-ai-en-a* *tiki* *aninji* *gudeŋ-le*
 go/come-CLOC-N.SFX-GEN after they call-PST
 ‘After he came, he called them.’ [Text-1, line 15]

3.2.3 Tense

Sora has a simple past/non-past opposition in its tense system, the former marked by *-t-*, the latter with *-l-*. Second and third person forms have the vowels *-e[y]/e* and first person forms have *-ay*. There is a class of verbs that require an *-n-* between the tense marker, here always realized as *-le-n/-te-n* in Sora, a class that optionally

may with no apparent meaning change, and a class whose meaning alters and a new (discontinuous) stem is created.

The Sora verbal template looks something like the following:

BOX 7.1: SORA FINITE VERB TEMPLATE

1/2PL:SUBJ//[NEG]-verb.stem-TNS -ITR-OBJ/1[SG/PL:SUBJ]/T/A+SUBJ

that is, with two prefixes or one where they almost always coalesce since they are sometimes homophonous, a verb stem that may be morphologically complex internally itself (e.g. have a causative or reciprocal prefix), a tense marker, then object- or suffixal (parts of) subject-agreement markers or a special tense/aspect form. Sometimes post-verbal uninflecting modal operators (at least one of which may be a LEX-headed auxiliary structure in the Anderson (2006) sense) are also found, see section 3.2.5/3.2.12.

A partial paradigm may be seen in the following sets from the verb *gi[y]ʔj* ‘see’ in the past and non-past forms, positive and negative. Note that the tense marker is lacking in negative past forms in Sora (see section 3.2.9).

- (105) (a) *giʔy-t-ay* (b) *giʔy-t-ey* (c) *aniŋ giʔy-t-ey*
 see-NPST-1 see-NPST-2/3 she/he see-NPST-2/3
 ‘I see’ [OG] ‘you see’ [OG] ‘she/he sees’ [OG]
- (d) *anlen a-giʔy-t-ay* (e) *amben a-giʔ-t-ey*
 we 1/2PL-see-NPST-1 you 1/2PL-see-NPST-2/3
 ‘we see’ [OG] ‘you (pl) see’ [OG]
- (f) *aninji giʔ-te-ji* (g) *nen giʔ-t-ay doŋ-om*
 they see-NPST-3PL I see-NPST-1 OBJ-2
 ‘they see’ [OG] ‘I see you.’ [OG]
- (106) (a) *nen drban giʔ-l-ay*
 I yesterday see-PST-1
 ‘I saw yesterday.’ [OG]
- (b) *aman drban giʔ-l-ey*
 you yesterday see-PST-2/3
 ‘Yesterday she/he saw.’ [OG]
- (c) *aninji řbàn giʔ-le-ji*
 they yesterday see-PST-3PL
 ‘They saw yesterday.’ [OG]
- (d) *anlen drban a-giʔ-l-ay*
 we yesterday 1/2PL-see-PST-1
 ‘We saw yesterday.’ [OG]
- (e) *amben drban a-giʔ-l-ey*
 you (PL) yesterday 1/2PL-see-PST-2/3
 ‘You (PL) saw yesterday.’ [OG]
- (107) (a) *anlen ir-glenliʔ*
 we NEG-see/NEG/see:1
 ‘We didn’t see.’ [OG]

- (b) *nen ir-glʌnli[ʔ]*
I NEG-see/NEG/see:1
'I did not see.' [OG]
- (c) *amən aʔd-gij-aʔd*
you NEG-see:2/NEG-NEG
'You do/did not see.' [OG]
- (d) *aninji a-giʔ-dəʔd-ji* cf. *giʔ-le-ji*
they NEG-see-2/3:NEG-3PL see-PST-3PL
'They didn't see.' [OG] 'They saw.' [OG]
- (108) (a) *ənən git-t-ay*
I see-NPST-1
'I see.' [OG]
- (b) *nen ir-glʌnliy*
I NEG-see/NEG:1/see
'I don't see.' [OG]
- (c) *ənən a-gijj-ay*
I NEG-see[:NEG]-1
'I didn't see (you).' [OG]
- (d) *a-giʔj-ey*
NEG-see-NEG:2/3
'You didn't see (me).' [OG]
- (e) *anin a-giʔj-ey*
s/he NEG-see-NEG:2/3
'She/he didn't see.' [OG]
- (f) *anlen a-giʔj-ay*
we NEG:1PL-see-NEG:1
'We didn't see (you).' [OG]
- (g) *g-giʔj-ay*
NEG:1PL-see-NEG:1
'We didn't see (you).' [OG]
- (h) *āʔ-a-giʔj-ey*
1/2PL-NEG-see-NEG:2/3
'You all didn't see.' [OG]

Some more examples of the non-past in *-t[eʎ]-(n)* and the past in *-l[eʎ]-(n)* in Sora include sentences such as these from the brief texts in section 7 and our field corpus.

- (109) *te-lt-* etc.
- (a) *iando tid-t-ij*
why hit-NPST-1
'Why are (you) hitting me?' (Starosta, p.c.)
- (b) *anin giʔy-t-ey*
she/he see-NPST-2/3
'She/he sees.' [OG]
- (c) *aninji git-te-ji*
they see-NPST-3PL
'they see' [OG]

- (d) *nen daʔa-n tiy-t-am*
 I water-N.SFX give-NPST-[1>]2
 ‘I (will) give you water.’ [OG]
- (e) *opin'o do orun'cu ar'-ge?-ta-jiy*
 Opino and Orunchu RECIP-SEE-NPST-3PL
 ‘Opino and Orunchu see each other.’ [OG]

Note also the following set of immediate future marked by *-t-* and a more remote future encoded by the lack of tense marker. This opposition is localized to a small set of motion verbs and is not characteristic of all Sora speech varieties.

- (110) (a) *paŋ-t-ai-ji*
 carry-NPST-CLOC-3PL
 ‘They(‘ll) bring now.’
- (b) *paŋ-ai-ji*
 carry-CLOC-3PL
 ‘They’ll bring later.’ (Starosta 1967:146)

The forms with *-n-* include reflexive, intransitives, some passives, and other detransitivized stems. Some stems seem to require this, and thus one may speak of an *-n-* conjugation in Sora. This *-n-* element is likely an archaic feature in Sora, shared with North Munda languages (Anderson 2007).

- (111) *teli-n-*
- (a) *giʔgiʔ-ti-n-ay*
 REDPL:scratch-NPST-ITR-1
 ‘I scratch myself.’
- (b) *giʔ-dam-ti-n-ay*
 see-RFLXV-NPST-ITR-1
 ‘I see myself.’
- (c) *jum-te-ti-n-ai* (d) *ijen nam-yo-ti-n-ay*
 eat-banana-NPST-[ITR]-1 I catch-fish-NPST-ITR/RFLXV-1
 ‘I am eating a banana.’ ‘I am fish-catching.’ [OG]

Sometimes alternate forms may be attested with and without the *-n-* inflectional element without any clear difference in meaning. Presumably these differ in at least connotation but how exactly is unknown and has not been tested with speakers in the field.

- (112) (a) *mɔɔŋ-tə-dʒi*
 happy-NPST-PL
 ‘They are very happy.’
- (b) *mɔɔŋ-t-en-dʒi*
 happy-NPST-ITR-PL
 ‘They are very happy.’ (Starosta 1967:106)

Some more examples of past tense forms in *-l-* in Sora are offered below.

- (113) *lel-l-*
- (a) *kan aiəm-leŋ-ən a-ieŋ-ən arsi a-kəndar-leŋ-ən*
 that time-LOC-N.SFX PFX-WHO-N.SFX monkey 3-branch-LOC-N.SFX

- ted-n-eten kun a-kəndar-ən əldiŋ-le*
 hang-n-T/A:3 that 3-branch-N.SFX break-PST
 ‘At that time, the monkey which hung from the branch of the tree, that branch broke.’ [Text-3, line 13]
- (b) *anin suŋa-dud-ən gam-eted əp-puŋpuŋ-le*
 he big-frog-N.SFX say-T/A CAUS-REDPL:puff-PST
 ‘He, the big-frog said puffing himself up.’ [Text-1, line 26]
- (c) *ukij əp-puŋpuŋ-le-n dəkət-ne po*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n annəŋ kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-say-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 ‘He puffed himself up even more while saying “this big”?! and his bloated stomach burst and he died.’ [Text-1, line 31]

Note that the past tense suffix in Sora shows assimilation of its initial consonant to a preceding stem-final *r*:

- (114) *uan a-ier-re ənn-ələŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-n-NEG-PL
 ‘“Where did you go?”; they didn’t answer him.’ [Text-1, line 16]

Some examples of the intransitive inflectional (detransitive) class in the past tense in *-le-n* in Sora includes sentences like these from the texts. As mentioned above, for some stems this *-n-* is obligatory while for others it does not appear to be so.

- (115) (a) *aninji tamba-n annəŋ jum-le ga-le tɔgəl-ən*
 they noon-N.SFX during eat-PST drink-PST at.night-N.SFX
annəŋ aboy mənəŋ bənda-n a-bo ara-leŋ-ən
 during one edge tank-N.SFX OBJ-one tree-LOC-N.SFX
diməd-le-n-ji
 sleep-PST-ITR-PL
 ‘They would eat and drink during the day and at night they would sleep in a tree in one corner of a tank.’ [Text-3, line 2]
- (b) *əntəpsele anlen kən begiŋa a-sɔ-le-n-ai*
 therefore we DEF different.place 1PL-[go.]hide-PST-ITR-1PL
 ‘Therefore we went and hid in (the[se]) different places.’ [Text-1, line 19]
- (c) *pottar-leŋ-ən jɔpba-leŋ-ən tilit-li-ŋ-ji*
 hole-LOC-N.SFX mire-LOC-N.SFX REDPL:bury-PST-ITR-PL
 ‘They buried themselves in holes and mud.’ [Text-1, line 13]
- (d) *tikki aninji ələŋ-le-n-ji dɔ etenasən*
 afterwards they answer-PST-[ITR]-PL so why
a-sɔ-le-n gam-le suŋa-dud-ən
 2PL-hide-PST-ITR say-PST big-frog-N.SFX
 ‘After that they answered, and the big-frog said “why did you hide?”’ [Text-1, line 17]

- (e) *bənlusaj kəniət-le bəŋsa gam-le kun sənna-mər-ən-ji*
 serves.him.right die-PST good say-PST DEF little-person-N.SFX-PL
bəibəi eʔaka-le-n-ji
 very rejoice-PST-ITR-PL
 ‘Serves him right that he died! Good! the little guys said and rejoiced greatly.’ [Text-1, line 33]
- (f) *ukij əp-puŋpuŋ-le-n dəkət-ne po*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n anney kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-say-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 ‘He puffed himself up even more while saying “this big”?! and his bloated stomach burst and he died.’ [Text-1, line 31]
- (g) *aninji aŋgaj-ən-ədəʔəŋ a-ŋidəb-ben-asen əskai-le-n-ji*
 they moon-N.SFX-OBJ DEF-pick.up-INF-FOR prepare-PST-ITR-PL
 ‘They prepared to pick up the moon.’ [Text-3, line 11]

Both the *-t-* non-past and the *-l-* past have parallels in other Munda languages. The former is found in this function throughout the South Munda languages (except Juang and Plains Gta?), while the *-l-* past likely reflects the *-l-* anterior forms of Kherwarian languages (Anderson 2007).

In Sora a range of post-verbal functional elements or post-inflectional operators express different verbal inflectional categories, for example, aspect (+tense), mood, negation, etc. One such element that appears to belong to the system of tense markers synchronically (but probably was historically an aspectual marker of some sort) is the second past in *-eten*. Sometimes it appears as *-eted* (this is seemingly even more true in Juray). It sometimes is used in durative or progressive contexts and other times in perfective/completive ones. How it differs from the ‘regular’ past in *-l-* has not been sufficiently investigated. It may be used with the same verbs and in the same texts and contexts as the *-l-* past, but seems mainly to be restricted to third person (singular?) subjects.

- (116) (a) *kan aiəm-leŋ-ən a-ieŋ-ən arsi a-kəndar-leŋ-ən*
 that time-LOC-N.SFX PFX-who-n.SFX monkey 3-branch-LOC-N.SFX
təd-n-eten kun a-kəndar-ən əldiŋ-le
 hang-n-T/A:3 that 3-branch-N.SFX break-PST
 ‘At that time, the monkey which hung from the branch of the tree, that branch broke.’ [Text-3, line 13]
- (b) *kun arsi-n kun aŋgaj-ən-ədəʔəŋ giŋ-an giŋ-le*
 that monkey-N.SFX that moon-N.SFX-OBJ see-N.SFX see-PST
a-gaŋiŋ-ji-ədəʔəŋ gam-eten
 3-friend-PL-OBJ tell-T/A: 3
 ‘That monkey saw that moon and told her friends.’ [Text-3, line 4]
- (c) *kun arsi-n gam-eten bo-mənra a-kəndar-leŋ-ən*
 that monkey-N.SFX say-T/A: 3 one person OBJ-branch-LOC-N.SFX
a-təd-nə-ba
 1PL-hold/hang-ITR-PL:IMP

- ‘That monkey told them “let one of us guys hold/hang from the tree branch.”’ [Text-3, line 8]
- (d) *ettegɔj diməd-ata diməd-atə aboj arsi-n*
 like.this sleep-NF:DUR/SIMULT sleep-NF:DUR/SIMULT one monkey-N.SFX
kun a-bənda-leŋ-ən aboj aŋgaj-ən-adoʔɔŋ gij-eten
 that OBJ-tank-LOC-N.SFX one moon-N.SFX-OBJ see-T/A:3
 ‘They kept sleeping and sleeping like this and one monkey saw a moon
 in that tank.’ [Text-3, line 3]
- (e) *anin suʔa-dud-ən gam-eted əp-puŋpuŋ-le*
 he big-frog-N.SFX say-T/A CAUS-REDPL:puff-PST
 ‘He, the big-frog said puffing himself up.’ [Text-1, line 26]
- (f) *dəjij meʔim ettegoj bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX/ITR-PL-GEN after
a-sənnə-mər sukku-n a-kako-n-adoʔɔŋ
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ
sarɔba-n er-ti-lə-be anson bara-eten
 paddy.field-N.SFX NEG-give-PST-NF:W/O himself work-T/A:3
 ‘[after] they worked like this for several years, Sukku the younger brother
 cultivated the paddy-field himself without giving his older brother the
 paddy-field.’ [Text-2, line 7]
- (g) *bijɔɔ a-ubban bɔibɔi barab-le ier-an ier-le*
 but OBJ-younger.brother very get.angry-PST go-N.SFX go-PST
anin adoʔɔŋ tuəb-eten
 he OBJ thrashʔ-T/A:3
 ‘But the younger brother got very angry went to him and thrashed(?)
 him.’ [Text-2, line 10]
- (h) *dɔ bɔ dinna kun a-gailɔ-leŋ abɔj raʔa-n iar-eted*
 so one day DEF OBJ-road-LOC one elephant-N.SFX walk-T/A:3
 ‘So one day an elephant was walking along the road.’ [Text-1, line 9]
- (i) *əntannəŋ kuni anin suʔa-dud-ən ier-ai-ted*
 then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
 ‘Then that one, him, the big-frog, came back.’ [Text-1, line 14]

It may appear on a predicate in sentence-medial position, so it has the same syntactic flexibility as *-l-* marked past forms. It may even appear with the *-n-* inflectional element as the second example demonstrates, but its use here seems to be motivated not by the contextual or lexical transitivity specification but rather by some as-yet-not understood factor.

- (117) (a) *bar anin gam-eten jaba anlen kun aŋgaj-ən-adoʔɔŋ*
 and she say-T/A:3 hey we that moon-N.SFX-OBJ
a-təb-n-ai-ba
 1PL-take.out-n-CLOC/1-PL:IMP
 ‘and she told them “let’s go and take out that moon.”’ [Text-3, line 6]
- (b) *do bo-mənra a-kəndar-leŋ-ən ted-n-eted dɔ anin*
 and one person OBJ-branch-LOC-N.SFX hold-n-T/A and he
ale-n bo-mənra təd-n-ete
 tail-N.SFX one man hold-n-3:IMP

‘When that one guy has held onto the branch, let another guy grab his tail.’ [Text-3, line 9]

In negative non-past and negative past, tense markers are often lacking in Sora (cf. similar kinds of negative conjugations in such sister languages as Remo, Gutob, or Gta?).

- (118) *ubban-adɔʔɔŋ* *gam-etan* *amən etenasən sərɔba-n*
 younger.brother-OBJ say-T/A-N.SFX/ITR you why paddy.field-N.SFX
ət-ti-ŋ
 NEG-give-1
 ‘He said to his younger brother “why (do) not (you) give me some paddy field.”’ [Text-2, line 9]

In a small number of instances, there are uses in our corpus of what appears to be *-ə* or *e* as a tense/aspect/mood marker in its own right without a consonantal augment. Its use remains a topic for future investigation.

- (119) Possible suffix *-e/-ə-*
eten aŋgaj jidəb-ə-ji
 what moon pick.up-FUT//MOD//T/A-PL
 ‘What moon will they (be able to) pick up?’ [Text-3, line 15]

3.2.4 Aspect and Aktionsart

Developed systems of morphological aspect as is seen in such Munda languages as Kharia (Peterson this volume) or Juang (Patnaik this volume) or even less developed ones found in such languages as Remo (Anderson and Harrison this volume) are not characteristic of Sora structure. The use of the ‘detransitivizer’ *-n* sometimes appears to have aspectual, rather than valence or voice functions, with certain predicates at least. However, such notions are compatible with scalar concepts of transitivity as discussed the seminal paper by Hopper and Thompson (1980).

- (120) (a) *anin dolba-te*
 he clear.field-NPST
 ‘He will clear the field.’
 (b) *anin dolba-te-n*
 he clear.field-NPST-ITR
 ‘He is clearing the field.’ (Starosta 1976:103)

One aspectual element that appears in post-verbal position, here generally following an inflected verb in the past, seems to have the function of a past habitual in Sora (cf. the *qan* element found in similar function and similar phrasal position in Korku (Zide this volume).

- (121) *aninji sənna-mər-ən-ji* *əgəndi-le-ji* *etteden anlen*
 they little-person- N.SFX-PL think-PST-PL therefore we
səənja-n *ət-təre-de* *jitəŋkə-n* *əd-ji-e*
 grain.tribute-N.SFX NEG-measure-NPST tax-N.SFX NEG-give-NPST
gam-le jure-lə-len *bijn*
 COMP twist-PST-1PL MOD
 ‘The little guys thought “we won’t measure out the grain-tribute and won’t pay the tax because he oppressed us.”’ [Text-1, line 32]

3.2.5 *Mood*

Imperative constructions in Sora may consist of a verb stem and an imperative suffix, or they may have an object marker if they contain a transitive verb with a personal object. As in several other Munda languages, monosyllabic Sora verbs may take an imperative suffix, and are often realized as *-[a]ʔə*, in Gajapati Sora.

- (122) (a) *kinte-n jum-aʔ*
 banana-N.SFX eat-IMP
 ‘Eat the banana.’
- (b) *je:lu-n jum-aʔ*
 meat-N.SFX eat-IMP
 ‘Eat (the) meat!’
- (c) *ʔə¹ -jum-aʔ*
 CAUS-eat-IMP
 ‘Make him eat!’
- (d) *¹jum-ə*
 eat-IMP
 ‘Eat!’
- (e) *gaʔ*
 eat:IMP
 ‘Eat (food w/water)!’
- (f) *ʔag-gaʔ*
 CAUS-eat:IMP
 ‘Feed (him)!’
- (g) *ʔamin ʔa¹g-ga*
 you CAUS-eat
 ‘(you) make him eat!’ [OG]

Object agreement can be suppressed in imperative forms in Sora yielding what appears to be an uninflected stem form of the verb.

- (123) (a) *kulu-n tiy[ɨn]*
 rice-N.SFX give[:1]
 ‘Give (me) the rice.’
- (b) *qarj-en tiy[ɨn]*
 rice-N.SFX give[:1]
 ‘Give (me) rice.’ [OG]

Plural addressee imperatives are mainly formed with the element *-ba*. This may also appear in first plural hortative or other modal forms as well.

- (124) (a) *qarj-en di-n-ba*
 rice-N.SFX give-1-PL:IMP
 ‘Give me rice!’ (PL addressee) [OG]
- (b) *gij-ɨn=ba*
 see-1=2PL
 ‘(y’all) see me!’ (Biligiri 1965:244)
- (c) *bar anin gam-eten jaba anlen kun aŋgaj-ən-adɔʔəŋ*
 and she say-T/A:3 hey we that moon-N.SFX-OBJ

- a-təb-n-ai-ba*
 1_{PL}-take.out-*n*-CLOC/1-_{PL}:IMP
 ‘and she told them “let’s go and take out that moon.”’ [Text-3, line 6]
- (d) *e gəŋŋ-ji a-gij-ba aŋgaj-en daʔa-leŋ-ən*
 hey friend-PL 1_{PL}-see-_{PL}:IMP moon-*n*-SFX water-LOC-*n*-SFX
gəlu-le gam-le ɔppuŋ-eten
 fall-PST QUOT tell-T/A:3
 ‘She told them “Hey friends, let’s go have a look; the moon has fallen in the water.”’ [Text-3, line 5]
- (e) *kun arsi-n gam-eten bo-mənra a-kəndar-leŋ-ən*
 that monkey-*n*-SFX say-T/A: 3 one person OBJ-branch-LOC-*n*-SFX
a-təd-nə-ba
 1_{PL}-hold-*n*-3:IMP
 ‘That monkey told them “let one of us guys hold/hang from the tree branch.”’ [Text-3, line 8]

In this latter function (modal/hortative first plural), it may appear as *-bə* or even *-be*.

- (125) *ettegoy anlen-a alale-n-ji ted-an ted-le-n*
 like.this we-GEN REDPL:tail-*n*-SFX-PL hold-*n*-SFX hold-PST-*n*-SFX/ITR
aŋgaj-ən-adoʔɔŋ a-təb-n-ai-bə
 moon-*n*-SFX-OBJ 1_{PL}-get.out-*n*-CLOC-_{PL}:IMP
 ‘In this way let’s hang by our tails and get out the moon.’ [Text-3, line 10]

Third person imperatives or optatives/hortatives are marked by the suffix or enclitic *-[e]te* in Sora.

- (126) (a) *do bo-mənra a-kəndar-leŋ-ən ted-n-eted do anin*
 and one person OBJ-branch-LOC-*n*-SFX hold-*n*-T/A and he
ale-n bo-mənra ted-n-ete
 tail-*n*-SFX one man 1_{PL}-hold/hang-*n*-_{PL}:IMP
 ‘When that one guy has held onto the branch, let one other guy grab his tail.’ [Text-3, line 9]
- (b) *ək-gɔ-bɔŋ-le set-n-ete təkɔd-aj-te*
 CAUS-sit-buffalo-CV give.up-ITR-3:IMP finish-AUX:COMPL-3:IMP
 ‘Let the making of (him) sit on the buffalo get finished up altogether.’
 (Starosta 1967:232)

The prohibitive or negative imperative formation in Sora is marked by the suffix/enclitic (or original auxiliary) *do[ʊ]ŋ*. As with positive imperatives, object agreement may be suppressed. Note that the object enclitic or suffix attaches to the right of the prohibitive particle in these constructions. The plural addressee element follows this.

- (127) (a) *ʔiy-douŋ-ij* (b) *ʔiy-douŋ-len*
 give-PROHIB-1 give-PROHIB-1_{PL}
 ‘Do not give.’ ‘Don’t give us.’

- (c) *d̪aɽ(a)j-en tiy-douŋ-(iŋ)* (d) *d̪aɽ(a)j-en tiy-douŋ-iŋ-ba*
 rice-N.SFX give-PROHIB-(1) rice-N.SFX give-PROHIB-1-PL:IMP
 ‘Do not give me rice.’ ‘Do not give me rice.’
 (SG addressee) (PL addressee) [OG]

Other modal constructions are also found in Sora. One such form is the evidential or dubitative element *po[ʔ]*. It follows immediately after whatever element it has its scope over.

- (128) (a) *ajaɽid po gam-le anin suɽa-məɽ*
 true DOUBT say-PST he big-person
 ‘“Is that really true?” he, the big guy said.’ [Text-1, line 24]
- (b) *dəkət-ne po gam-le*
 this.much-EMPH DOUBT say-pst
 ‘“This (big)?!” he said.’ [Text-1, line 27]
- (c) *ukij əp-puŋpuŋ-le dəkəd-ne po gam-le*
 AUGM CAUS-REDPL:puff-PST this.much DOUBT say-PST
 ‘So he puffed himself up some more and said “this (big)?!”’ [Text-1, line 29]
- (d) *ukij əp-puŋpuŋ-le-n dəkət-ne po*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n anney kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-SAY-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 ‘He puffed himself up even more while saying “this big”?! and his bloated stomach burst and he died.’ [Text-1, line 31]

One final modal form that bears mention here is the conditional construction in Sora. This is marked by the verb in a dependent-marked past (participle) form in *-le-n-den*. Note that the plural marker comes between the *-n-* and the *-den* in the conditional.

- (129) (a) *amə gil-le-n-den* (b) *gil-l-en-ji-den*
 you see-PST-ITR-COND see-PST-ITR-PL-COND
 ‘if you see’ ‘if they see’ (Ramamurti 1931:28)
 Note: *l-en-den* (always PST)

In Juray, Sora’s closest sister language or a divergent Sora dialect, a kind of strengthening or assimilation of **-n-* to (pre-glottalized) *-d-* is found in certain conditional formations (130); the range of environments triggering this strengthening is still unknown. Note that in Juray, unlike Sora, the verb may appear in forms other than the dependent past participle form. This is also true of the cognate formation in Gorum (Anderson and Rau, this volume).

- (130) (a) Juray (b) Juray
se’d-lə’d-den lakod-en-den
 accompany-PST:ITR-COND carry-ITR-COND
 ‘if you accompany (me)’ ‘if you carry’

- (c) Juray
loŋe'd-tə'd-den
 get.dark-NPST:ITR-COND
 'if it becomes dark' (A. Zide 1983)

3.2.6 Orientation/directionality

In Sora, among the many verbal categories that find some formal expression may be included verbal deixis, orientation, or directionality. This means action moving towards or away from the subject (or topic/discourse locus). The two oppositional categories are translocative/itive (131), or motion away from speaker/discourse locus, and cislocative/ventive, or motion towards it (132). In a small set of motion verbs, the opposition is realized as *-aŋə/ə/Ø* vs. *-ai-ay*.

- (131) (a) *ŋen bajar-ən* 'yer-t-e
 I market-N.SFX go-NPST[: 1/TLOC]
 'I am going to the market.' [OG]
 (b) *ŋem bajar-ən* 'yer-t-e
 I market-N.SFX go-NPST[: 1/TLOC]
 'I am going to the market.' [OG]
 (c) *ʔamm ba'zar-ən* yer-e
 you market-N.SFX go-2:NPST[/TLOC]
 'You go to market.' [OG]
 (d) *ŋem ba'zar-ən yer-er ʔamm ba'zar-ən yer-ə*
 I market-N.SFX go-1:TLOC you market-N.SFX go-2.NPST
 'I am going to market, you are going to market.' [OG]
 (e) *dɔ bɔ dinna kun a-gailɔ-leŋ abɔj raʔa-n iar-eted*
 so one day DEF OBJ-ROAD-LOC one elephant-N.SFX walk-T/A:3
 'So one day an elephant was walking along the road.' [Text-1, line 9]
 (f) *uan a-ier-re ənn-ɔləŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-n-NEG-PL
 '“Where did you go?”; they didn't answer him.' [Text-1, line 16]
 (g) *ə-i-si-a*
 'Make signal with hand to say “go away”.'
 (h) *ə-i-si-ai*
 'Make signal with hand to say “come on”.' (Ramamurti 1933:3)
- (132) (a) *əntannəŋ kuni anin suʔa-dud-ən ier-ai-ted*
 then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
 'Then that one, him, the big-frog, came back.' [Text-1, line 14]
 (b) *ier-ai-en-a tiki aninji guden-le*
 go/come-CLOC-N.SFX-GEN after he-PL call-PST
 'After he came, he called them.' [Text-1, line 15]
 (c) *ettegoy anlen-a alale-n-ji ted-an ted-le-n*
 like.this we-GEN REDPL:tail-N.SFX hold-N.SFX hold-PST-N.SFX/ITR
aŋgaj-ən-advɔŋ a-təb-n-ai-bə
 moon-N.SFX-OBJ 1PL-get.out-n-CLOC-PL:IMP
 'In this way let's hang by our tails and get out the moon.' [Text-3, line 10]

The element *-a* (> *-elə*) in this opposition is simply the regular tense-cum-subject marker while the origin of *-ail-ay* remains obscure. Based on typological and comparative Austroasiatic data, it seems likely that this (and the first person subject marker *-ail-ay*) originally derives from a deictic serial verb construction involving ‘come’ (Anderson 2006, 2007). Note that the cognate element has developed into one of the variant means of marking of version in addition to cislocative meanings in Sora’s sister language Gorum (Parenga) (A. Zide 1972, Anderson 2007, Anderson and Gurevich forthcoming, Anderson and Rau this volume). Developments (of ‘come’) to cislocatives are found, for example, in Turkic languages (Anderson 2004a), to first person markers in Iwaidjan languages (Nicholas Evans, personal communication) and to now-lexicalized version markers in Burushaski (Anderson and Gurevich forthcoming, Bashir 1985).

3.2.7 Voicelversion

Voice categories morphologically realizable in the Sora verbal complex include causative, reciprocal; and reflexive/detransitive/passive/intransitive. Voice categories in Sora deal with argument structure of the predicate and affect the conjugational pattern of the verb stem.

The causative in Sora is marked by the prefix *əb-* (and many variants) which is attached to the rightmost stem syllable. In monosyllabic stems this is realized as a prefix, while the causative in Sora (as a hallmark of Austroasiatic languages generally speaking (Anderson 2007) in stems greater than single syllable, bisyllabic in the case of Sora, the prefix is realized word-internally as an infix.

Some examples with prefixed causative, an assimilation to a geminate consonant can be seen in the following tokens of ‘make him eat’ and some unassimilated examples in (133).

- (133) (a) *ʔamun ʔa¹g-ga?* (b) *ab-gugur-t-ai*
 you CAUS-eat:IMP CAUS-REDPL:call-NPST-1
 ‘Make him eat.’ [OG] ‘I’ll make someone call.’
 (c) *ab-umaɪ-dəm-te-n-ai*
 CAUS-wash-RFLXV-NPST-INTR-1
 ‘I’ll get myself washed’ (Ramamurti 1931:29)

Infixed (or rightmost syllable prefixed) causative forms in Sora include the following:

- (134) (a) *battoŋ* (<*bt) < *baton*
 (b) *kəjjeɔ* (<*bj) ‘kill’ < *kəjeɔ* ‘die’ (Ramamurti 1931:47)
 (c) *əsəntan aninji-dəm tləblubaj-le-ji*
 in.vain they.self drown/CAUS/-PST-PL
 ‘In vain they got themselves drowned.’ [Text-3, line 16]
 (d) *də a-ieŋ-ən a-mənra səɔbə-n ət-ti-əd anin*
 so OBJ-WHO-N.SFX OBJ-man paddy.field-N.SFX NEG-give-NEG he
bəndij ier-le də a-ieŋ-ən ə-mənra səɔb-ən
 jail:N.SFX go-PST and OBJ-WHO-N.SFX OBJ-man paddy.field-N.SFX
əd-ŋəŋ-əd anin su?uŋ-leŋ-ən dəku-le-n də tiki
 NEG-get-NEG he house-LOC-N.SFX COP-PST-N.SFX and afterwards

- (c) *er-dzum-le-be* ‘without having eaten’
vs.
(d) *er-uma-le-n-be* ‘without having bathed’ (Ramamurti 1931:29)
(e) *paŋ-n* ‘take back/home for oneself’ (Starosta 1967, 1976)
(f) *puttar-leŋ-ən jɔpba-leŋ-ən tilit-li-ŋ-ji*
hole-LOC-N.SFX mire-LOC-N.SFX REDPL:bury-PST-ITR/RFLXV-PL
‘They buried themselves in holes and mud.’ [Text-1, line 13]
(g) *nen tʰaʀa[giʔna-leŋ giʔ-t-əy ... giʔ-dam-ti-n-ay*
I mirror-LOC see-NPST-1 see-RFLXV-NPST-[ITR]-1
‘I (can) see myself in a mirror.’ [OG]
(h) *nen tʰaʀa[giʔna-leŋ giʔ-giʔ-ti-n-ay*
I mirror-LOC REDPL:see-NPST-ITR-1
‘I (can) see myself in a mirror.’
(i) *so:-te-n-ai*
hide-NPST-ITR-1
‘I’ll hide myself.’ (Ramamurti 1931:26)
(j) *ran-n-e:ten*
crush-ITR-3.PST
‘It was crushed.’ (Starosta 1967, Ramamurti 1931)
(k) *mo:-te-n*
swallow-NPST-ITR
‘It can be swallowed.’ (Starosta 1967, 1976)

Another reflexive marker is the suffix *-dəm/dam-* which appears after the verb root and before the tense marker. This (always?) co-occurs with *-n-* inflectional forms.

- (137) (a) *giʔ-dam-ti-n-ay* (b) *tid-dəm-te-n*
see-RFLXV-NPST-[ITR]-1 beat-RFLXV-NPST-ITR
‘I see myself.’ ‘He beats himself.’ (Ramamurti 1931:26)
(c) *anin pɔ-dəm-n-eten*
he stab-RFLXV-ITR-T/A
‘He stabbed himself.’ (Starosta 1967:136, Ramamurti 1931:23)

Sometimes the semantics are unclear and general detransitive or low transitivity marking seems to be involved.

- (138) Juray
(a) *bulbule-dəm-ləl-lən* (b) *paŋ-dəm-lən-ti*
REDPL:wander-RFLXV-PROG-PST:ITR take-RFLXV-PST.ITR-DEP
‘She kept wandering.’ ‘took with her’ (A. Zide 1983)

Another option for creating passive-like semantics in Sora is through use of object markers with transitive stems with suppressed agents.

- (139) Sora
gij-l-iŋ
see-PST-1
‘I was seen.’ (Biligiri 1965:233)

3.2.8 Finiteness

Finiteness in Sora is not formally marked *per se* as it is Kherwarian languages. A verb in Sora in complex clause and complex predicate structures may appear in a range of formal guises or non-final forms. The finite verb template was given in Box 7.1.

Non-finite forms show a range of other realizations. One dependent or nominalized verb form takes the shape of *alə...-ənlan*. This appears sentence medially in clause-final position in complex sentences of various types.

- (140) (a) *anin suʔa-dud-ən a-ber-n-an etente beren aninji*
 he big-frog-N.SFX DEP-say-n-N.SFX what say-N.SFX they
sənnə-dud-ən-ji māne-le-ji
 small-frog-N.SFX-PL obey-PST-PL
 ‘Whatever the big-frog said, the small frogs obeyed, had to obey.’ [Text-1, line 8]
- (b) *də a-ieŋ-ən a-mənra səɾəbə-n ət-ti-əd anin*
 SO OBJ-who-N.SFX OBJ-man paddy.field-N.SFX NEG-give-NEG he
bəndij ɪer-le də a-ieŋ-ən ə-mənra səɾəb-ən
 jail:N.SFX go-PST and OBJ-who-N.SFX OBJ-man paddy.field-N.SFX
əd-ɲaŋ-əd anin suʔuŋ-leŋ-ən dəku-le-n də tiki
 NEG-get-NEG he house-LOC-N.SFX COP-PST-N.SFX and afterwards
kəlkəl-ən batte anin kləb/ied-dəm-le-n
 REDPL:worry-N.SFX with he kill/CAUS/kill-RFLXV-PST-ITR
 ‘The one who didn’t give the fields, he went to jail and the one who didn’t get the fields, he was at home, with worry, and afterwards he killed himself.’ [Text-2, line 14]
- (c) *ukij əp-puŋpuŋ-le-n dəkət-ne pə*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n amneŋ kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-say-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 ‘He puffed himself up even more while saying “this big”?! and his bloated stomach burst and he died.’ [Text-1, line 13]

A subtype of this had been grammaticalized as a kind of purposive infinitive in *alə...be-n [asən]*

- (141) (a) *aninji aŋgaj-ən-ədɔŋ a-ɲidəb-ben-asen ɔskai-le-n-ji*
 they moon-N.SFX-OBJ DEP-pick.up-INF-FOR prepare-PST-ITR-PL
 ‘They prepared to pick up the moon.’ [Text-3, line 11]
- (b) *aninji kən gam-ən gam-le aɲaŋij-nə a-kəndar-leŋ-ən*
 they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX
səreŋ alan-ji təd-le-n təd-le-n aninji
 from [3-]tail-PL hang-PST-N.SFX hold-PST-N.SFX they

aŋgaj-ən-ədɔʔəŋ a-nidəb-ben-asən ɔskai-le-n-ji
 moon-N.SFX-OBJ PFX-pick.up-INF-FOR prepare-PST-[ITR/RFLXV]-PL
 ‘They thus discussed and then really began to pick up the moon hang-
 ing from the branch from each other’s tails.’ [Text-3, line 12]

A kind of same subject structure called ‘perfect participle’ by Starosta (1967) but with a range of functions (including causal subordinate clause formation), is commonly found in narratives which consists of a kind of verbal reduplication of the shape Verb-*an* Verb-*le*.

- (142) (a) *kun asən kun sənna-dəd-ən-ji raʔan-ədɔʔəŋ*
 DEF for DEF small-frog-N.SFX-PL elephant:N.SFX-OBJ
gij-an gij-le bətəŋ-le iersed-le-ji
 see-N.SFX see-PST be.frightened-PST run.away-PST-PL
 ‘Because of seeing the elephant, the small frogs were frightened and
 ran away.’ [Text-1, line 12]
- (b) *ettegoy anlen-a alale-n-ji ted-an ted-le-n*
 like.this we-GEN REDPL:tail-N.SFX hold-N.SFX hold-PST-N.SFX/ITR
aŋgaj-ən-ədɔʔəŋ a-təb-n-ai-bə
 moon-N.SFX-OBJ 1PL-get.out-n-CLOC-PL:IMP
 ‘In this way let’s hang by our tails and get out the moon.’ [Text-3, line 10]

Note that the reduplicated infinitive found in various Munda languages is largely confined to various grammaticalized uses with (often fused) TAM auxiliaries or ‘explicator’ verbs. It probably is a historical relic of this system even in tense-inflected forms, since these elements themselves are likely to derive from auxiliary structures historically (Anderson 2007).

nen ʔaʔa[giʔna-leŋ giʔ-giʔ-ti-n-ay giʔ-giʔ-ti-n-ay
 I mirror-LOC REDPL:SEE-NPST-[ITR]-1 REDPL:SEE-NPST-[ITR]-1
 ‘I see myself in a mirror.’ [OG]

3.2.9 Negation

There are a number of different negative constructions in Sora. Most involve the use of the prefix, variably realized as *aʔ-*, *əd-*, *ət-*, *aʔn-*, *ənn-*, *er-*, etc. These form various combinations with other elements to yield a range of negative conjugations. The presence of negative conjugations that differ from corresponding ones is a feature of South Munda languages, for example, Gutob, Remo, or Gtaʔ.

The negative non-past is usually marked by the combination of the negative prefix *a-/ə-* with the tense-cum-subject-cum-polarity suffix *-e[y]*. This corresponds to the *t-* non-past declension in the positive.

- (143) *a-/ə-...-e[y]*
- (a) *nen baʔzar-m əʔ-ye:r-ey*
 I market-N.SFX NEG-go-1
 ‘I don’t go to the market.’
- (b) *byəne baʔzar-ən [ə-]ʔ-ye:r-ey*
 tomorrow market-N.SFX [NEG:]go-1
 ‘Tomorrow I won’t go to the market.’

- (c) *man'yeyley 'guna dako'-lenden nen ba'zar-ən a?n-iy-e*
 money find:DEP AUX-COND I market-N.SFX NEG-go-[FUT:1]
 'If I do not find my money, I cannot go to the market.'
- (d) *man'yeyley 'guna dako'-lenden nen ba'zar-ən 'a?n-iy-e*
 money find AUX-COND I market-N.SFX NEG-go-[FUT:1]
 'If I do not find my money, I cannot go to the market.' [OG]
- (e) *aninji sanna-mər-ən-ji əgəndi-le-ji etteden anlen*
 they little-person-N.SFX-PL think-PST-PL therefore we
səənja-n ət-təre-də jitənkə-n əd-ji-e
 grain.tribute-N.SFX NEG-measure-NPST tax-N.SFX NEG-give-NPST
gam-le jure-lə-len bin
 COMP twist-PST-1PL MOD
 'The little guys thought "we won't measure out the grain-tribute and
 won't pay the tax because he oppressed us".' [Text-1, line 32]

Some hints of object demotion in negative structures may be found in the speech of certain Sora, from objective to allative adessive, at least with certain verbs.

- (144) (a) *nen maŋ-nam ?a'-dirdjir-n-e... ?a'-dirdjir-n-e*
 I ALL-2 NEG-believe-n-1 NEG-believe-n-1
 'I do not believe you.'
- (b) *nen maŋ'-nam ?a-dirdjir-n-e...*
 I ALL-2 NEG-believe-n-1
 'I do not believe you.'
- (c) *nen do' ?oŋ-nam ?a?-'gəlam-ai*
 I OBLQ/OBJ-2 NEG-know-1
 'I do not know you.' [OG]

Object encoding in the verb is permitted however in negative forms in Sora.

- (145) (a) *g-tiy-l-am* (b) *nemi da?a-n a-tiy-l-əm*
 NEG-give-PST-[1>]2 today water-N.SFX NEG-give-PST-[1>]2
 'I didn't give you.' 'I didn't give (it to) you today.'
- (c) *g?-til-l-iŋ*
 NEG-give-PST-1
 'You didn't give me.'
- (d) *aninji rban da?a-n a-tiy-l-əm-ji*
 they yesterday water-N.SFX NEG-give-PST-2-3PL
 'Yesterday they didn't give you water.'
- (e) *r'ban 'aninjij a?n-ar-gi?-lə-ji*
 yesterday they NEG-RECIP-see-PST-3PL
 'Yesterday they did not see each other.' [OG]
- (f) *ubban-ado?əŋ gam-etan amən etenasən*
 younger.brother-OBJ say-T/A-N.SFX/ITR you why
səɾɔba-n ət-ti-ŋ
 paddy.field-N.SFX NEG-give-1
 'He said to his younger brother "why (do) not (you) give me some
 paddy field."'

Sometimes present tense reciprocal forms may keep the *-t-* non-past form in negative forms. The use of this remains uninvestigated.

aʎʎm aʎʎn-ar-giʎ-t-biʎ
 we NEG-RECIP-see-NPST-1PL
 ‘We don’t see each other.’ [OG]

Past tense negative forms with second and third person subjects tend to have a circumfix negative+co-negative construction. Unlike many other languages where this kind of construction may be found (e.g. French *ne* and *pas*), the negative and co-negative elements in Sora are identical (subject of course to morphophonological and idiosyncractic idiolectal alter[*n*]ation). Also, like the negative non-past, there is no tense marker *-l-* in negative past forms in Sora.

(146) *ədlt...-əd*

- (a) *dɔ a-ieŋ-ən a-mənra sərɔbən ət-ti-əd anin*
 so OBJ-WHO-N.SFX OBJ-man paddy.field-N.SFX NEG-give-NEG he
bəndiŋ ier-le dɔ a-ieŋ-ən ə-mənra sərɔb-ən
 jail:N.SFX go-PST and OBJ-WHO-N.SFX OBJ-man paddy.field-N.SFX
əd-ŋaŋ-əd anin suʎuŋ-leŋ-ən dəku-le-n dɔ tiki
 NEG-get-NEG he house-LOC-N.SFX COP-PST-N.SFX and afterwards
kəl-kəl-ən batte anin kləbʎied-dəm-le-n
 REDPL:WORRY-N.SFX with he kill/CAUS/kill-RFLXV-PST-ITR
 ‘The one who didn’t give the fields, he went to jail and the one who didn’t get the fields, he was at home, with worry, and afterwards he killed himself.’ [Text-2, line 14]
- (b) *ənt aiəm-leŋ-ən anin suʎa-dud-ən ət-dəku-əd*
 that time-LOC-N.SFX he big-frog-N.SFX NEG-COP-NEG[PST: 3]
 ‘At that time he, the big-frog, wasn’t there.’ [Text-1, line 10]
- (c) *uan a-ier-re ənn-ɔləŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-n-NEG-PL
 ‘“Where did you go?”; they didn’t answer him.’ [Text-1, line 16]
- (d) *amən aʎd-gij-aʎd* (e) *giyʎ-le-ji*
 you NEG-see: 2/3NEG-NEG see-PST-3PL
 ‘You do/did not see.’ [OG] ‘They saw.’ [OG]
- (f) *aninji a-giʎ-dəʎd-ji* (g) *a-giʎ-jeŋ*
 they NEG-see-2/3:NEG-3PL NEG-see-NEG: 2/3
 ‘They didn’t see.’ [OG] ‘You didn’t see (me).’ [OG]
- (h) *anin a-giʎ-jeŋ* (i) *giʎ-a-giʎ-jeŋ*
 she/he NEG-see-NEG: 2/3 NEG-1/2PL-see-NEG: 2/3
 ‘She/he didn’t see.’ [OG] ‘You all didn’t see.’ [OG]

First person subject forms in the negative past tense on the other hand are overtly nominalized signalled both by the use of the nominal negator prefix *er-* (*er-lir-*) in combination with the nominalizing infix *-n-*. This is used with first singular and first plural forms.

- (147) (a) *anlen a-giʎ-l-ay* (b) *anlen ir-gənliʎʎ*
 we 1/2PL-see-PST-1 we NEG-see/NEG:1/see
 ‘We saw.’ [OG] ‘We don’t/didn’t see.’ [OG]

- (c) *ir-glənli?*
NEG-see/NEG:1/see
'I/we didn't see.'
- (d) *ɲen ir-glənli?*
I NEG-see/NEG:1/see
'I did not see.'
- (e) *anlen a-giʔ-l-ay*
we 1/2PL-see-PST-1
'We saw'
- (f) *anlen ər-glənliy*
we NEG-see/NEG:1/see
'We didn't see.'
- (g) *anlen a-giʔ-l-ay*
we 1/2PL-see-PST-1
'We saw.'
- (h) *ər-glenliy*
NEG-see/NEG:1/see
'We didn't see.'
- (i) *ɲnen giʔ-t-ay*
I see-NPST-1
'I see.'
- (j) *ɲen ir-glənliy*
I NEG-see/NEG:1/see
'I didn't see. [OG]
- (k) *ɲem ba'jar-in ɛrgelir... ɛrni... ɲem ba'jar-ən ɛrni*
I market-N.SFX NEG-n-go:[1]
'I didn't go to the market.' [OG]

Note that the *-n-* infix is lacking in first person subject negative forms in the non-past in Sora.

- (148) (a) *ɲnen a-gijj-ay*
I NEG-see[:NEG]-1
'I don't see (you).'
- (b) *a ɲen a-giʔd[j]-ay*
DISC I NEG-see-NEG-1
'I don't see (you).'
- (c) *anlen a-giʔj-jay*
we NEG:1PL-see-NEG:1
'We don't see (you).'
- (d) *a-giʔj-jay*
NEG:1PL-see-NEG:1
'We don't see (you).' [OG]

Negative nominalized verbs are marked by this same construction (*er-...-ən-*) in Sora as well. Compare the following two examples in this regard, with a nominalized and non-finite negative verb, respectively.

- (149) (a) *ɛr-t-ən-i-ən asən*
NEG-give-NEG-give-N.SFX SUBORD
'because of not giving...'
- (b) *əʔ-ti-lə-n asən*
NEG-give-PST-DEP SUBORD
'because he didn't give'
(Starosta 1967:216)

A negative subordinate clause is made by combining the negative prefix *er-* with the subordinator or case/post-positional element *-be* in a circumfixal combination of *er-...-be*.

- (150) *dəjiŋ meŋim ettegoy bara-le-n-ji-a tiki*
several year like.this work-PST-N.SFX/ITR-PL-GEN after
a-səŋna-mər sukku-n a-kako-n-adɔʔŋ sarɔba-n
3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ paddy.field-N.SFX
er-ti-lə-be anson bara-eten
NEG-give-PST-NF:W/O himself work-T/A:3
'[after] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.' [Text-2, line 7]

Negative imperatives (prohibitives) are formed by the suffix *do[ɔ]/[ŋ]*-. Object suffixes attach to this element so it occupies the same slot in the template as do the tense markers.

- (151) =do[ʊ]ŋ-[1/2] ‘don’t’
- (a) *ʔamin ga-dŋ...* *jum¹-dʊ[ŋ]*
 you eat/drink-PROHIB eat-PROHIB
 ‘Do not eat.’ [OG]
- (b) *ʔamin jum¹-dʊ[ŋ]* (c) *jum¹-dʊŋ*
 you eat-PROHIB eat-PROHIB
 ‘Do not eat.’ [OG] ‘Do not eat.’ [OG]
- (d) *ʔag-ga-dʊŋ* (e) *ʔa^ʔ-jum-dʊ[ŋ]*
 CAUS-eat-PROHIB CAUS-eat-PROHIB
 ‘Don’t make him eat.’ [OG] ‘Don’t make him eat.’ [OG]
- (f) *ʔa^ʔ-jum-dʊ[ŋ]*
 CAUS-eat-PROHIB
 ‘Don’t make him eat.’ [OG]
- (g) *tʔy-dʊŋ-ij* (h) *dʔ(a)j-en tʔy-dʊŋ-(ij)*
 give-PROHIB-1 rice-N.SFX give-PROHIB-1
 ‘do not give’ [OG] ‘do not give me rice’ [OG]
 (SG addressee)
- (i) *dʔ(a)j-en tʔy-dʊŋ-ij-ba*
 rice-N.SFX give-PROHIB-1-PL:IMP
 ‘Do not give me rice.’ (PL addressee) [OG]

Note that in causative prohibitive forms, the ‘object’ marking on *dŋ* is the causee and semantically the subject of the verb, not its object, thus *ʔa^ʔ-jum-dŋ-ij* ‘don’t make me eat’ not ‘don’t make (him) eat me’.

- (152) (a) *ʔa^ʔ-jum-dŋ-ij* (b) *ʔa^ʔ-jum-dŋ-len*
 CAUS-eat-PROHIB-1 CAUS-eat-PROHIB-1PL
 ‘Don’t make me eat.’ ‘Don’t make us eat.’ [OG]

There are two negative nominal copular forms in Sora. The non-equational or non-identity copula is *tet* and the existential/locational negative copula is *agasa*. Compare the contrastive use of the two of these with and without the possessed kin noun below.

- (153) *tet*
- (a) *anin bəŋsa tet* (b) *keke-mər tet*
 he good NEG.COP crazy-man NEG.COP
 ‘He is not good.’ ‘He is not a nut.’
- (c) *anin oaa-n tet*
 he 3-father-N.SFX NEG.COP
 ‘He isn’t a father.’ (Starosta 1967:107, 114)

- (154) *agasa*
- (a) *anin suʔuŋ-ən agasa*
 she house-N.SFX NEG.COP.EX
 ‘She isn’t in the house/at home.’ (Starosta 1967:113)
- (b) *anin a-oaa-n agasa* cf. (c) *anin a-oaa-n dəko*
 he 3-father-N.SFX NEG.COP.EX he 3-father-N.SFX COP
 ‘He does not have a father.’ ‘He has a father.’
 (Starosta 1967:114)

3.2.10 Derivation

Verbal derivation in Sora is properly subsumed under different sub-headings in this chapter. Thus, the major derivational processes are ones of voice covered in section 3.2.7. The extensive and elaborate compounding of verbal and nominal stems that typifies Sora verb structure is addressed in section 3.2.11.

3.2.11 Noun incorporation and combining forms

Of all the Munda languages, Sora has far and away the most elaborate and developed system of verb–noun stem combining of a type that is generally known in the literature as noun incorporation. Formally speaking, the nominal component that combines with the verb root is the so-called combining form, the often monosyllabic ‘root’ form of the noun. Like many languages with incorporation, the most common use of this structure in Sora is to incorporate the object/patient of a transitive (2-argument) verb. Most typically, the incorporation of this element into the verb reduces the valency of the verb in Sora, and it appears with intransitive inflection (marked by the suffix *-n-*).

- (155) (a) *jen jum-t-ai* (b) *ijen kante-n jum-t-ai*
 I eat-NPST-1 I banana-N.SFX eat-NPST-1
 ‘I am eating.’ ‘I am eating a banana.’
- (c) *jen jum-te-ti-n-ai*
 I eat-banana-NPST-ITR-1
 ‘I am eating a banana.’
- (d) *ijen jaʔat nem-t-ay* (e) *jen nam-jaʔt-[t]I-n-ay*
 I snake catch-NPST-1 I catch-snake-NPST-ITR-1
 ‘I am catching a snake.’ ‘I am snake-catching.’
- (f) *ajen nam-jaʔt-ti-n-ay*
 I catch-snake-NPST-ITR-1
 ‘I am snake-catching.’ [OG]

This is a highly productive feature of the Sora verbal system with virtually every noun having a combining form, and most combinations being acceptable, given an appropriate discourse context for the creation of the form. Full paradigms in the past and non-past can be generated easily by speakers.

- (156) (a) *ijen nam-yo-ti-n-ay* (b) *amən nam-yo-ti-ŋ*
 I catch-fish-NPST-ITR-1 you catch-fish-NPST-ITR:2/3
 ‘I am fish-catching.’ ‘You are catching fish.’
- (c) *anin nam-yo-ti-ŋ*
 she/he catch-fish-NPST-ITR:2/3
 ‘She/he is catching fish.’
- (d) *anlen a-ŋəm-yo-ti-n-ay*
 we 1/2PL-catch-fish-NPST-ITR-1
 ‘We are catching fish.’
- (e) *ambin á-ŋəm-yó-ti-ŋ*
 you (PL) 1/2PL-catch-fish-NPST-ITR:2/3
 ‘You (PL) are catching fish.’

- (f) *aninji nam-yo-tr-n-ji*
they CATCH-fish-NPST-ITR-3PL
'They are catching fish.'
- (g) *nam-yim-ti-n-ai* (h) *nam-jat-ti-n-ai*
catch-chicken-NPST-[ITR]-1 catch-snake-NPST-[ITR]-1
'I catch a chicken.' 'I catch a snake.'
- (i) *nam-yoo-ti-n-ai*
catch-fish-NPST-[ITR]-1
'I catch a fish.' [OG]
- (157) (a) *drban nen nam-ja?t-li-n-ay*
yesterday I catch-snake-PST-ITR-1
'Yesterday I snake-caught.'
- (b) *amən drban nem-ja?t-li-n*
you yesterday catch-snake-PST-ITR: 2/3
'Yesterday you snake-caught.'
- (c) *a-nam-ja?t-li-n-ay*
1PL-catch-snake-PST-RFLXV/ITR-1
'We snake-caught.'
- (d) *nem-ja?t-li-n-ji*
catch-snake-PST-RFLXV/ITR-3PL
'They snake-caught.'
- (e) *aninji nemi nam-ja?t-li-n-ji*
they today catch-snake-PST-RFLXV/ITR-3PL
'They were snake-catching today.'
- (f) *nen drban nem-jat-li-n-ay*
I yesterday catch-snake-PST-[ITR]-1
'Yesterday I caught a snake.' [OG]

Forms cannot just be created without using a proper combining form. Anderson used the following two sentences with Orunchu Gomango. The first one was accepted readily and repeated. The second one was said, 'productively' (or so was being tested) formed from the noun by dropping the derivational prefix *kən-* (see section 3.1.10), he mused over it for a while before rejecting it and supplying the third form with the proper but semi-suppletive combining form for 'chicken'.

- (158) (a) *nen kansim-ən nem-t-ay*
I chicken-N.SFX catch-NPST-1
'I am catching a/the chicken.'
- (b) **nen nam-sim-te-n-ay ??*
I catch-?chicken-NPST-RFLXV/ITR-1
'I am catching a/the chicken.'
- (c) *nem-yim-tr-n-ay*
catch-chicken-NPST-RFLXV/ITR-1
'I am catching a/the chicken.' [OG]

In incorporated complexes the use of the intransitive marker *-n-* encodes reflexive/detransitive or actor-oriented action and contrasts with the transitive, undergoer-oriented form lacking this element in incorporated complexes.

Object noun incorporation is apparently also permitted in non-finite, infinitive and nominalized forms in Sora.

- (159) (a) *ɲam-jo:-le-n*
 catch-fish-NF/SS/PST-ITR
 ‘having caught fish’ (Ramamurti 1931:142)
- (b) *ə-gik-kid-ben*
 2PL-see-tiger-INF
 ‘(for you) to see the tiger’ (Ramamurti 1931:44)
- (c) *ɲəram-jo:n* (d) *an-ɲənam-jo:-n*
 catch.NMLZ.catch-fish-NOUN PFX-catch.NMLZ.catch-fish-NOUN
 ‘means of catching fish’ ‘fish that have been caught’
 (Ramamurti 1931:46) (Ramamurti 1931:44)

Note that not all incorporated object forms require the intransitive suffix *-n* in Sora.

- (160) (a) *gad-bo:ŋ-t-e-ji*
 cut-buffalo-NPST-3-PL
 ‘They are cutting the buffalo.’ (Ramamurti 1931:49)
- bagu-n-ji ɲaŋ=boj-lə-ji*
 two-N.SFX-PL get/take=woman-PST-PL
 ‘They both got married.’ [Text-2, line 4]

Contrast the following forms in this regard:

- (161) (a) *kuŋ-bəb-t-ε*
 shave-head-NPST
 ‘You shave (s.o.’s) head.’
- (b) *kuŋ-bəb-te-n*
 shave-head-NPST-ITR
 ‘You shave your head.’ (Biligiri 1965:240)

Here the opposition appears to be between reflexive object with *-n* and action directed at another which lacks this element.

Pronominal object marking may be found with verb forms also exhibiting noun incorporation, so these incorporated complexes are not obligatorily intransitive in Sora, that is, incorporation is *not* an inherently valence-reducing process in this language.

- (162) (a) *lem-jiŋ-t-am* (b) *lem-si-t-am*
 bow-feet-NPST-2 bow-hand-NPST-2
 ‘I bow to your feet.’ (ɔwɫajŋ foot) [OG] ‘I shake (bow to) your hand.’ [OG]
- (c) *lem-jeŋ-te-ben-ji*
 bow-foot-NPST-2PL-3PL
 ‘They bow to your feet.’ [OG]
- (d) *aninji paŋ-sal-l-iŋ-ji*
 3PRON-PL bring-liquor-PST-1-PL
 ‘They brought me liquor.’ (Ramamurti 1931:142)

Thus, in addition to the stem change (causativization) with the incorporated noun forms below, there is alternation between presence and absence of *-n* corresponding to reflexive vs. non-reflexive possessor of the hand being washed. First person subjects remain in

an uninflected form but third person pronouns are in the oblique/dative object form, and the verb is marked transitive (at least non-intransitive). This shows that, as are certain Kherwarian languages, Sora is sensitive to degrees of transitivity in its verbal morphology. This likely is an archaic feature in these two Munda subgroups.

- (163) (a) *nen* ¹*ʔa-si-ti-n-ai*
 I wash-hand-NPST-RFLXV/ITR-1
 'I am washing my hand.' [OG]
- (b) *nen anim aɔŋ aba-si-t-ai*
 I he OBJ wash-hand-NPST-1
 'I am washing his hand.' [OG]
- (c) *nen annji aɔŋ aba-si-t-ai aba-jiŋ-t-ai*
 I they OBJ wash-hand-NPST-1 wash-foot-NPST-1
 'I am washing their hands.' [OG] 'I am washing their feet.' [OG]

Apparently the *-si-* incorporated in this verb can be an instrumental as well as a patient in Gajapati Sora.

- (164) *nen a-jiŋ-m-ji aɔŋ aba-si-t-ai*
 I 3-foot-N.SFX-PL OBJ wash-hand-NPST-1
 'I am washing their feet.' [OG]

'Double marking' with an external noun and an incorporated noun together may also be possible in this Sora variety.

- (165) *a-jiŋ-m-ji aba-jiŋ-t-ai*
 3-foot-N.SFX-PL wash-foot-NPST-1
 'I am washing their feet.' [OG]

This is certainly possible with words meaning 'right hand', 'left hand', and 'hand' incorporated into the verb. As already pointed out by Sadock (1991) (with regard to Gta?), Munda languages have an unusual but not unheard of transparency between incorporated elements and referents in the external phrasal syntax.

- (166) (a) *nen jandʒumsi-nej aba-si-t-ai*
 I right.hand-1 wash-hand-NPST-1
 'I am washing my right hand.' [OG]
- (b) *kap'dʒabuʒisi-nej aba-si-t-ai*
 left.hand-1 wash-hand-NPST-1
 'I am washing my left hand.' [OG]

Under questioning, a speaker was able to produce a form with an external modifier 'big' modifying an incorporated noun 'snake'; he then replaced it and was happy with a nominal form of this type instead. Other forms he outright rejected so there may be some cline of acceptability of external modifiers in Sora. This needs to be investigated further.

- (167) (a) *nen suʒa jaʔad-an jam-t-ay*
 I big snake-N.SFX catch-NPST-1
 'I am catching a big snake.' [OG]

- (b) *nen suʒa nam-jaʔt-ti-n-ay*
I big catch-snake-NPST-RFLXV/ITR-1
'I am big snake catching.' [OG]
- (c) *nen suʒa nam-jaʔt-mar*
I big catch-snake-man
'I am a big snake catcher.' [OG]
- (d) *ijen suʒa nam-jaʔp-mar*
I big catch-snake-man
'I am a big snake catcher.' [OG]
- (e) *ijen kuluʔ jaʔr-an jem-t-ay*
I green snake-N.SFX catch-NPST-1
'I am catching the green snake.' [OG]
- (f) **nen kuluʔ nam-jat-ti-n-ay*
I green catch-snake-NPST-RFLXV/ITR-1
'*I am green snake catching.' [OG]

As in many languages with noun incorporation (Anderson 1997), the non-subject agreement position in Sora is frequently used to mark affected animate possessors. Possessor raising is frequently found in such incorporated constructions.

- (168) (a) *soi-tam-t-am* burn-mouth-NPST-2
'I will burn your mouth.'
- (b) *nen ag-gaʔ-siʔ-am*
I NEG-drink-hand-2
'I will not drink from your hand.'
(Ramamurti 1931:142)
- (c) *ji-loʔ-siʔ-t-am* stick-earth-hand-NPST-2
'mud will stick to your leg.'
- (d) *nen aj-jaʔ-dar-siʔ-am*
I NEG-receive-cooked.rice-hand-2
'I won't receive rice from your hand.'
(Ramamurti 1931:44)

Note that forms (c-e) have both possessor raising and double noun incorporation. Subject possessor raising is also possible in Sora incorporated forms.

- (169) (a) *kuy-bəb-t-əm*
shave-head-NPST-2
'Your head is shaven.' (Biligiri 1965:240)
- (b) *asuʔ-boʔb-t-ijn*
hurt-head-NPST-1
'My head hurts.' (Ramamurti 1931:143)

Another of the noteworthy characteristics of the system of noun incorporation in Sora is the presence of multiply incorporated constructions. Two such examples are found in (168 c and d) and one more is offered as follows.

- (170) *jo-me-boʔb-dem-te-n-ai*
smear-oil-head-RFLXV-NPST-ITR-1
'I will anoint myself with oil.' (Ramamurti 1931:143)

There are also forms with an original serialized verb construction and a single incorporated noun in Sora as well.

- (171) *paŋ-ti-dar-ijn-tem*
bring-give-cooked.rice-1-3.PST
'He brought and gave me cooked rice.' (Ramamurti 1931:43)

One of the most, if not the most, noteworthy aspects of Sora noun incorporation is that a transitive verb may incorporate its *agent* argument. While many languages allow for incorporation of intransitive (unaccusative) subjects, incorporated agent forms of transitive verbs are not frequently attested cross-linguistically, if at all, in other languages. Because of the nature of Sora verb morphology with its distinction between free forms and combining forms of nouns, it is clear that these are indeed incorporated complexes. These incorporated stems remain transitive in Sora: they allow for the formal indexing of objects as well within the incorporated complex.

- (172) (a) *nam-kit-t-am*
 seize-tiger-NPST-2
 ‘Tiger will seize you.’ (Ramamurti 1931:40)
- (b) *sa-bud-t-am*
 mangle-bear- NPST-2
 ‘Bear will mangle you.’ (Ramamurti 1931:142)
- (c) *mo-kul-t-am*
 swallow-ghost- NPST-2
 ‘Ghost will swallow you’
- (d) *paj-sum-t-am*
 carry-spirit-NPST-2
 ‘Spirit will carry you away’ (Ramamurti 1931:142)

The forms with incorporated agent differ from corresponding forms lacking incorporation both formally and functionally in Sora. From a formal perspective, the free-standing full form noun subject appears with the absolutive or nominal suffix *-n* and stands in pre-verbal position. In incorporated subject forms, the noun appears in a monosyllabic combining form after the verb root within the verb complex before the tense marker. Compare the following synonymous variants.

- (173) (a) *kina-n nam-t-am*
 tiger-N.SFX seize-NPST-2
 ‘The tiger will seize you.’
- (b) *nam-kid-t-am*
 seize-tiger-NPST-2
 ‘Tiger will seize you.’ (you will be tiger-seized) (Ramamurti 1931:40)

These incorporated agent forms contrast formally with (most) incorporated object forms by never allowing the intransitive suffix *-n*, while incorporated object forms as mentioned and exemplified above frequently (but not obligatorily) have this element. Unfortunately, the details of this usage have to date not been adequately worked out and must remain a subject for future research.

- (174) *nam-kid-te-n-ai*
 seize-tiger-NPST-ITR-1
 ‘I will seize the tiger.’ (Ramamurti 1931:40)

Like incorporated object forms, a full paradigm of incorporated subject forms may be found in Sora. Note not only the usual lack of a tense marker in the negative non-past forms but also its use of negative past forms (they are highly marked forms and do not have ‘regular-irregular’ inflection).

(175) ‘tiger’ = subject/agent, pronominal suffix = object/patient

Past	Non-past	Imperative	Prohibitive
<i>ɲam-kil-l-ɲn</i>	<i>ɲam-kit-t-ɲn</i>	<i>ɲam-kid-ɲn-te</i>	<i>ɲam-kid-doŋ-ɲn-te</i>
<i>ɲam-ki-l-am</i>	<i>ɲam-kit-t-am</i>	<i>ɲam-kid-am-te</i>	<i>ɲam-kid-doŋ-am-te</i>
<i>ɲam-kil-l-e</i>	<i>ɲam-kit-t-e</i>	<i>ɲam-kid-e-te</i>	<i>ɲam-kid-doŋ-e-te</i>
<i>ɲam-kil-lə-len</i>	<i>ɲam-kit-tə-ten</i>	<i>ɲam-kid-len-te</i>	<i>ɲam-kid-doŋ-len-te</i>
<i>ɲam-kil-l-ai</i>	<i>ɲam-kit-t-ai</i>	<i>ɲam-kid-ai-te</i>	<i>ɲam-kid-doŋ-ai-te</i>
<i>ɲam-kil-lə-ben</i>	<i>ɲam-kit-tə-ben</i>	<i>ɲam-kid-ben-te</i>	<i>ɲam-kid-doŋ-ben-te</i>
<i>ɲam-kil-l-e-ji</i>	<i>ɲam-kit-t-e-ji</i>	<i>ɲam-kid-e-te-ji</i>	<i>ɲam-kid-doŋ-e-te-ji</i>
<i>Neg.past</i>	<i>Neg.non-past</i>		
<i>ɲn-ɲam-kil-l-ɲn</i>	<i>ɲn-ɲam-kid-ɲn</i>	(Ramamurti 1931:40–41)	

Agent incorporation may be found in the imperative as well in Sora. Thus, both of the following sentences are grammatical in Sora.

- (176) (a) *kina-n ɲam-ɲn-te* vs. (b) *ɲam-kid-ɲn-te*
 tiger-N.SFX seize-1-3.IMP seize-tiger-1-3.IMP
 ‘May the tiger seize me!’ ‘May I betiger-seized.’ (Ramamurti 1931:41)

3.2.12 Auxiliary verb constructions

In Sora there is also considerable evidence that speaks to a formally active system of complex predicate formation in addition to a synchronically active use of deictic serial verb constructions. Many South Asian languages make use of complex predicates to mark a wide range of tense, aspect, and mood categories, or show morphologically complex verb forms which originated in such formations; Munda languages generally being no exception to this tendency (Hook 1991, Anderson 2007). In Sora, many of these former auxiliaries or explicator verbs function as semi-affixal elements that encode primarily aspectual (e.g. habitual, frequentative, completive), less commonly modal categories. As is the case in complex predicate structures across the South Munda languages, monosyllabic lexical verbs in such formations may be reduplicated, but bisyllabic stems are not.

- (177) *-lo:*
gugu-lo:te-n
 RDPL:call-FREQ-NPST-ITR
 ‘He calls (me) frequently.’ (Ramamurti 1931:28)

- (178) *-laŋ*
 (a) *kəŋ-kəŋ-laŋ-te-n*
 RDPL:abuse-HAB-NPST-ITR
 ‘He abuses (all people)’ (Ramamurti 1931:28)
 (b) *gʷər-ləŋ-te-n*
 sacrifice-HAB-NPST-ITR
 ‘he sacrifices’ (Starosta 1967)

Not all of these now-fused original auxiliary verb constructions take reduplicated lexical stems in Sora. Some rather require a bare, unmarked (or Ø-marked) lexical verb stem. Again, this patterning is also familiar from the synchronically bi-partite

auxiliary verb constructions in a number of Sora's sister Munda languages (e.g. Remo or Gta?). One such element is the completive 'auxiliary' or aspectual suffix *-a[ɔ]j*.

- (179) *-a:ɔj*
 (a) *ək-gɔ-bɔŋ-le* *set-n-ete* *təkɔd-aj-te*
 CAUS-sit-buffalo-CV give.up-ITR-3:IMP finish-AUX:COMPL-3:IMP
 'let the making of (him) sit on the buffalo get finished up altogether'
 (Starosta 1967:232)
 (b) *jum-a:ɔj-a*
 eat-COMPL-IMP
 'eat it (all) up!' (Ramamurti 1931:27)

The functions of such elements appear to be in line with areal norms, insofar as these can be gleaned *per se* in the Sora structures. Thus a now-fused verb originally meaning 'throw' seems to form perfective action (Hook 1991, Anderson 2003) in both Sora and Juray, at least with motion verbs. Note that there is another homophonous functional element in Sora that means 'leave, give up' that appears to be used as a marker of durativity, for example, 'keep Verb-ing, continue to Verb'.

- (180) Juray
jir-(e)-sed-en
 go-AUX-ITR
 'went away' (A. Zide 1983)

- (181) Sora
kun asən kun sənna-dud-ən-ji *raʔan-adɔʔɔŋ* *gij-an*
 DEF for DEF small-frog-N.SFX-PL elephant:N.SFX-OBJ see-N.SFX
gij-le *bətɔŋ-le* *iersed-le-ji*
 see-PST be.frightened-PST run.away-PST-PL
 'Because of seeing the elephant, the small frogs were frightened and ran away.' [Text-1, line 12]

The tense markers of Sora are likely to have originated in auxiliary structures. In careful speech, the tense markers can still be given independent stress, speaking to the original phrasal nature of the construction.

- (182) *jen nam-yim* *ti-n-ay*
 I catch-chicken NPST-RFLXV/ITR-1
 'I am chicken-catching.' [OG]

There are a range of verbal inflectional elements that synchronically appear to be uninflecting functional elements but which have the position and function typically found with auxiliaries in Sora. One of these is the second past tense form in *-[e]tem*. These may have been original LEX-headed or split inflectional structures in the Anderson (2006) sense, with at least object marking on the lexical verb.

- (183) *paŋ-ti-dar-ijn-te:n*
 bring-give-cooked.rice-1-3.PST
 'He brought and gave me cooked rice.' (Ramamurti 1931:43)

The capabilitive construction in Sora is one that bears special mention. Unlike other verbal operators which generally follow the lexical head verb, this element, which can

be used predicatively when no complement is expressed and take tense and subject marking, the capability verb *rapti*, precedes the lexical verb. The lexical verb remains the inflectional head and takes subject and tense marking (and cislocative, intransitive) as appropriate. A representative set of forms (Starosta 1967: 187–188) in the positive and negative of this construction is offered in (184).

- (184) *rapti ə-i-tɛ* *rapti ə-paŋ-t-ai/ə-paŋ-te-n-ai*
 CAP 1PL-GO-NPST CAP 1PL-CARRY-NPST-1PL/CLOC//1PL-CARRY-NPST-ITR-1PL/CLOC
rapti ə-i-tɛ paŋ-n-ai
 CAP 1PL-GO-NPST CARRY-ITR-CLOC/1PL
 ‘We can go.’ ‘We can bring’ // ‘we can bring back.’ ‘We can go and bring back.’
 (Starosta 1967:187)

- rapti əʔ-i-ɛ* *rapti əʔ-paŋ-[n-]ai*
 CAP NEG-GO-NPST CAP NEG-CARRY-NPST-[ITR]-1PL/CLOC
rapti əʔ-i-ɛ paŋ-n-ai
 CAP NEG-GO-NPST CARRY-ITR-CLOC/1PL
 ‘We can go.’ ‘We can bring’ // ‘we can bring back.’ ‘We can go and bring back.’
 (Starosta 1967:188)

Serial verb structures or (fused) semi-dependent clause-chaining structures are also characteristic of the Sora verb system. These come in two basic formal types. The first type is akin to nuclear serialization and is a close compounding of predicates.

Note, for example, the following forms of the shape verb₁-verb₂-NPST, with single non-past inflection.

- (185) (a) *paŋ-ti-t-am*
 bring-give-NPST-2
 ‘I’ll bring and give (it to) you.’ (Ramamurti 1931:44)
 (b) *ti-jum-t-am*
 give-eat-NPST-2
 ‘I’ll give you to eat.’ (Ramamurti 1931:44)

Such forms can also appear with incorporated nouns as well, yielding rather frightful complexes of the following sort:

- (186) *mal-jum-pu-da-tam-t-əm* *poʔŋ*
 wish-eat-cake-AUX-mouth-NPST-2 Q/DOUBT
 ‘Do you long to eat cake?’ (Ramamurti 1931:143)

Other ‘serialization’ patterns are more like a set of fused clause-chained structures, with a semi-inflected form of verb₁ that obligatorily appears in the past tense and functionally the completion of the action of which is necessary and prior to the action described by verb₂.

- (187) (a) *gil-le-jir-ij*
 see-PST-leave-1
 ‘See me before you go.’ (Ramamurti 1931:44)
 (b) *ə-berna op(p)uŋ-le-jir-ij=te:n*
 ACC-word tell-PST-go-1-3.PST
 ‘Having told me the word, he left.’ (Ramamurti 1931:44)

Sometimes this yields structures that are ambiguous between a non-finite verb and a finite verb or a sequence of juxtaposed coordination of two finite verbs.

- (188) *ənsələ-n oroŋ-le ier-le*
 woman-N.SFX carry-CV/PST go-PST
 ‘(He) went carrying the woman’ or ‘having carried the woman, (he) went’ or
 ‘he carried the woman and then left.’ (Starosta 1967:180)

Non-past marking is permitted in such sequences as well.

- (189) *anin ijai-te-n-gu-am*
 he come-NPST-ITR-call-2
 ‘He came and called you.’ (Ramamurti 1931:44)

Serialized forms of this latter type can appear with (an) incorporated object(s) as well in Sora. In these structures, however, the order of elements in serialized and incorporated sequences is mostly set in Sora, following a morphological, not semantic ordering, for example:

BOX 7.2: SORA INCORPORATION AND COMPLEX VERB TEMPLATE

(PERS-NUMB/NEG)-verb₁-PST/CNCTV-(-ITR-)-verb₂-N:CF-PRS/ TENSE-ITR-PERS/NUMB

Note that the intransitive marker may appear following either verb₁ or verb₂. The element *-le-* frequently appears in serialized forms. Subject/object markers of person and number on the other hand, always appear in final position.

- (190) *bagun-ben ə-il-le-ga-sal-n-e*
 both-2PL 1/2.PL-go-PST-drink-liquor-ITR-1PL
 ‘Both of you went and drank liquor.’ (Ramamurti 1931:44)

3.3 Expressives

In Sora, as in other Munda languages, there is a set of tag or echo word forms which are identical or are sometimes phonological variants of each other. These may be typologized as follows:

- (191) Full reduplication (type A)
- | | |
|-------------------------|--------------------------------------|
| <i>duŋ' duŋ</i> | ‘thud, thump’ |
| <i>yeŋ' yeŋ</i> | ‘to blow gently, as breeze’ |
| <i>duŋ' duŋyeŋ' yeŋ</i> | (Adv.) ‘like a storm’ |
| <i>rumrum</i> | ‘rapidity of movement in dancing’ |
| <i>meder: meder:</i> | ‘dim, dusky’ |
| <i>'takar takar</i> | ‘in a tottering or trembling manner’ |
| <i>yub yub</i> | ‘sounds of footsteps’ |
| <i>kejem kejem-</i> | ‘to be fond of’ |
| <i>pada: pada:</i> | ‘smack’ (sound) |

- (192) Full reduplication with consonant overwriting (type B)
- | | |
|----------------------------|--|
| <i>sɔːra moːra</i> | ‘the Soras and the Oriyas, etc.’ |
| <i>ɲaguːɲaguː</i> | ‘sound of pounding, blowing or thumping’ |
| <i>mandiːn tadiːn</i> | ‘plates, dishes’ (< <i>mandiːn</i> ‘a dish’) |
| <i>ɲaːjeːɲte paːjeːɲte</i> | ‘toddle, walk’ |
- (193) Full reduplication with vowel(s) overwriting (type C)
- | | |
|------------------------------|--|
| <i>padum padam</i> | ‘the sound produced by blows’ |
| <i>sai le mui le</i> | ‘having searched’ |
| <i>rum ram</i> | ‘to trot or prance, as horses’ |
| <i>kasulaː kaselaː</i> | ‘falling into the well’ |
| <i>kiki koko</i> | ‘to produce the click sound heard in kissing’ |
| <i>kermoiloge karmoiloge</i> | ‘smilingly, cheerfully’ (< <i>kermoi</i> ‘to smile’) |
| <i>meːɲ maːɲ</i> | ‘humming of bees and insects’ |
| <i>medoːr madoːr</i> | ‘glitteringly’ |
| <i>sukaːl sakaːl-ən</i> | ‘early’ (< Marathi <i>sakaːl</i>) |
| <i>brəiɲ bəraːɲ-</i> | ‘to dazzle’ |
| <i>pəˈseiloge pəˈsuiloge</i> | ‘loudly with a hissing sound’ |
| <i>ˈkokede ˈkakode</i> | ‘crooked, curved’ |
| <i>abuɲle abuɲle</i> | ‘turning from side to side in bed’ |
| <i>ˈpadum ˈpadam</i> | ‘sound produced by blows’ |
| <i>ˈrədəmai ˈrademai</i> | ‘sound of crackling’ |
- (194) Full reduplication with both vowel and consonant overwriting (type D)
- | | |
|-------------------------|----------|
| <i>modɛːte paddeːte</i> | ‘twists’ |
|-------------------------|----------|

Some verb roots appear only in reduplicated forms:

- (195)
- | | |
|---------------------------|----------------------------------|
| <i>mel mel</i> | ‘to inspect’ |
| <i>me me</i> | ‘to nod, to shake’ |
| <i>ɲaɲ ɲaɲ</i> | ‘to teach, admonish’ |
| <i>yey yey</i> | ‘to blow gently, as breeze’ |
| <i>piɲ piɲ (~peɲ peɲ)</i> | ‘to be cracked’ ‘to have chinks’ |
| <i>se se</i> | ‘to choose’ |
| <i>sib sib</i> | ‘to feel pinched’ |

These inherently reduplicated verbs may sometimes themselves undergo type C reduplication:

- (196) *piɲ piɲ paɲ paɲ* ‘to be sunken (as the eyes of an old person)’

Other verbs have semantically distinct unreduplicated and reduplicated forms:

- (197)
- | | |
|-----------------|--|
| <i>guɲ-</i> | ‘to fall, to tumble’ |
| <i>guɲ guɲ-</i> | ‘to strike, to knock’ |
| <i>saɲ-</i> | ‘to crush, to bruise’ |
| <i>saɲ saɲ-</i> | ‘to shell or break open (cashew nuts, etc.)’ |
| <i>poi-</i> | ‘to hoe around a plant’ |
| <i>poi poi-</i> | ‘to curl, as a tendril’ |
| <i>jun-</i> | ‘to escort’ |
| <i>jun jun-</i> | ‘to take back, return, restore’ |

Similarly, some nouns and adjectives exist only in reduplicated forms:

- (198) *niɫ'niɫən* 'masturbation'
saŋsaŋən 'turmeric paste'

Morphological retriPLICATION is infrequently encountered in Ramamurti's materials; the authors have not confirmed these forms with contemporary speakers.

- (199) *siɫ'siɫsiɫ'lamge* 'commotion'

We also find restricted co-lexicalized collocations that duplicate certain elements.

- (200) (a) *əsu-bəb* *rəŋa-bəb*
 sick-head cold-head
 'sick as a dog' (Starosta 1967:263)
 (b) *ənselɔ* *tet* *kimbɔj tet*
 woman NEG.COP woman:TAG NEG.COP
 'there is no woman here' (Starosta 1967:264)

4 SYNTAX

Much remains to be investigated in the syntax of the Sora language. A first but now rather dated attempt at Sora syntactic analysis may be found in Starosta's (1967) dissertation to which the interested reader is referred. We offer only the briefest of comments below.

4.1 Syntax of the simple sentence

The basic clausal constituent order of Sora is Subject-Object-Verb which is typical of the languages of South Asia. Intransitive sentences have the structure Subject-Verb/Copula.

- (201) (a) *allen* *si¹ʔiŋ-m* *a-sab¹ja-le*
 we house-N.SFX 1PL-build-PST:1PL
 'We built the house.'
 (b) *abɔj* *gailɔ-n* *dəku-le*
 one road-N.SFX COP-PST
 'There (once) was a road.' [Text-1, line 1]

'Have' possession constructions in Sora are marked by a copular verb and a possessed nominal (if permitted), as is commonly found in Eurasian SOV languages.

- (202) *yagi* *on-ŋen* *di dəku*
 three child-1 DISC COP
 'I have three kids, ya know.' [OG]

The relative order between two objects in a ditransitive is not set and subject to as yet unexplored discourse considerations.

- (203) *ənlen* *aman* *da?a-n* *q²-tiy-t-am*
 we you water-N.SFX 1PL-give-NPST-[1>]2
 'We give you water.' [OG]

While dominantly verb final, some post-verbal nominal elements are permitted, under conditions that remain opaque and a subject for future research.

- (204) *jen giʔ-t-ay doŋ-om*
 I see-NPST-1 OBJ-2
 ‘I see you.’

As an SOV language of Eurasia, certain other phrasal features of Sora are areally typical. This includes within the domain of nominal syntax a predominance of head-final order, so in addition to OV we find Gen N, Num N, Dem N, N Postpos, Adv Adj and Adj N. Within the verbal phrase, one typically encounters the order Adverb–Verb and Verb–Auxiliary (often in Sora, as mentioned in section 3.2.12, in a synchronically fused order of Verb–Auxiliary). Some examples of phrasal structures illustrating these are offered below.

- (205) (a) *migelturu minyun* (b) *migelyagi minyu* (c) *abɔj gailɔ-n*
 eighteen year:N.SFX fifteen year one road-N.SFX
 ‘eighteen years old’ ‘fifteen years old’ ‘one road’
 (d) *bo-mənra-n ɔnim sukku bar bo-mənra*
 one person-N.SFX 3:name Sukku and one person:GEN
ɔnim mənɟiɾa
 3:name Mangaɾa
 ‘One was named Sukku one was named Mangaɾa.’ [Text-2, line 3]
 (e) *ijen kulu? jaʔr-an jem-t-ay* (f) *ejə bɔibɔi suɾa*
 I green snake-N.SFX catch-NPST-1 No very big
 ‘I am catching the green snake.’ ‘No, very big’ [Text-1,
 line 22]
 (g) *kun a-gailɔ-lej*
 that OBJ-ROAD-LOC
 ‘on that road’

Note that while (pro)nominal possessors precede nouns, first and second person pronominal possessives are enclitic to the noun.

- (206) (a) *ɔnim-jen opino gomango*
 name-1 Opino Gomango
 ‘My name is O.G.’
 (b) *jen-a si-jen* (c) *jen si-jen* (d) *amən-a si-nam*
 I-GEN hand-1 I hand-1 YOU-GEN hand-2
 ‘my hand’ ‘my hand’ ‘your hand’
 (e) *anlen si-len*
 we hand-1PL
 ‘our hand(s)’ [OG]

Adverbs are not infrequently found in clause-initial position, but this is in no sense obligatory.

- (207) (a) *jen rban daʔa-n tiy-l-əm*
 I yesterday water-N.SFX give-PST-[1>]2
 ‘I gave you water yesterday.’ [OG]

- (b) *rban jən daʔa-n tiy-l-am*
 yesterday I water-N.SFX give-PST-[1>]2
 ‘I gave you water yesterday.’ [OG]

Thus, in intransitive clauses the order of Adverb-Oblique-Subject-Verb is not uncommon in Sora narratives.

- (208) (a) *də bə dinna kun a-gailə-lej abəj raʔa-n iar-eted*
 so one day DEF OBJ-road-LOC one elephant-N.SFX walk-T/A:3
 ‘So one day an elephant was walking along the road.’ [Text-1, line 9]
- (b) *dijəj-antide sua səmma-dud-ən dəku-le-ji*
 how.many 100 little-frog-N.SFX COP-PST-PL
 ‘How many hundreds of little frogs there were!’ [Text-1, line 5]
- (c) *a-gam-le-n anlen kən təŋət-ən abəj*
 1PL-say-PST-ITR we this road-N.SFX INDEF
suxa-mər-ən ier-re
 really.big-person- N.SFX go-PST
 ‘We said “one really big guy travelled this road.”’ [Text-1, line 18]

4.2 Complex sentence structure

Very little can be said here about complement clause structure or the range of use of sentential or clausal arguments in Sora. Some predicates which take clausal complements require a purposive-marked complement (with *asən* taking either an unmarked or dependent for the verb stem) or a purposive infinitive in (*a...*) *-ben asən*.

- (209) (a) *jən bajar-ən ar'yu yɪr-asən ʔi'siʔm-t-ai*
 I market towards?? go-PURP DES-NPST-1
 ‘I want to go to the market.’ [OG]
- (b) *jən 'jum-bin 'asen ʔisim-t-ai*
 I eat-INF PURP DES-NPST-1
 ‘I want to eat.’ [OG]
- (c) *je'nuxm-m 'jum-bin 'ʔasen ʔisim-t-ai*
 something-N.SFX eat-INF PURP DES-NPST-1
 ‘I want to eat something.’ [OG]

Other verbs require their complement clause to be simply marked as non-final or non-finite (or perhaps nominalized) with the ‘noun’ suffix *-ən*.

- (210) (a) *a'min a-git'-[e-n a-man'dɪra jən gə'lam-t-ai*
 you DEP-see-PST-N.SFX OBJ-man I know-NPST-1
 ‘I know that you saw the man.’ [OG]
- (b) *a-tit-le-ŋ 'a-mandɪra jəŋ galam-t-ai*
 DEP-hit-PST-N.SFX OBJ-man I know-NPST-1
 ‘I know that you hit the man.’ [OG]

Complements usually precede the verbs that subcategorize for them in Sora but may also follow them under appropriate conditions (the details of which remain to be worked out).

- (211) *jən ʔlanʔisimai amin ba'zar-ən 'yer-gi-n-asim*
 I want/NEG:1/NEG.NPST you market-N.SFX go-ECHO-N.SFX-PURP
 ‘I do not want you to go to market.’ [OG]

Like many SOV languages of Eurasia, a non-finite or participle/converb form of the verb ‘say’, in Sora *gamle*, may function to introduce direct discourse as a quotative, and with certain predicates of mental action to also serve as a complementizer.

- (212) (a) *e gəŋŋi-ji a-gij-ba aŋgaj-en daʔa-leŋ-ən*
 hey friend-PL 1PL-see-PL:IMP moon-N.SFX water-LOC-N.SFX
gəlu-le gam-le ɔppəŋ-eten
 fall-PST QUOT tell-T/A:3
 ‘She told them, “Hey friends, let’s go have a look; the moon has fallen in the water.”’ [Text-3, line 5]
- (b) *amən ian gam-le gənur-te gam-le gələm-le*
 you how rain-NPST COMP know-PST
 ‘How did you know that it would rain?’ (Starosta 1967:197)

4.2.1 Relative-type clauses

Like most other topics in Sora syntax, only the briefest of comments can be offered here on the nature of relative clauses in Sora. For more detail see Starosta (1967:242–254).

Relative clauses in Sora appear to reflect areal norms to a large degree. They are usually restrictive clauses consisting of a relative pronoun often marked by the focus marker *-te*, in a clause preceding the main clause and linked to it through a correlative element. Thus, in the following examples, the marked dependent relative marker *a-ien-ən* appears in the preposed relative clause often followed by a noun and co-indexed with the pronoun *anin* or deictic *kun* in the main clause.

- (213) (a) *dɔ a-ieŋ-ən a-mənra sərɔbə-n ət-ti-əd anin*
 so OBJ-who-N.SFX OBJ-man paddy.field-N.SFX NEG-give-NEG he
bəndiŋ ier-le dɔ a-ieŋ-ən ə-mənra sərɔb-ən
 jail:N.SFX go-PST and OBJ-who-N.SFX OBJ-man paddy.field-N.SFX
əd-ŋaŋ-əd anin suʔuŋ-leŋ-ən dəku-le-n dɔ tiki
 NEG-get-NEG he house-LOC-N.SFX COP-PST-N.SFX and afterwards
kəlkəl-ən batte anin kləblied-dəm-le-n
 REDPL:worry-N.SFX with he kill/CAUS/kill-RFLXV-PST-ITR
 ‘The one who didn’t give the fields, he went to jail and the one who didn’t get the fields, he was at home, with worry, and afterwards he killed himself.’ [Text-2, line 14]
- (b) *kan aiəm-leŋ-ən a-ieŋ-ən arsi a-kəndar-leŋ-ən*
 that time-LOC-N.SFX PFX-who-N.SFX monkey 3-branch-LOC-N.SFX
ted-n-eten kun a-kəndar-ən əldiŋ-le
 hang-n-T/A:3 that 3-branch-N.SFX break-PST
 ‘At that time, the monkey which hung from the branch of the tree, that branch broke.’ [Text-3, line 13]
- (c) *dɔ ete-n-te j-ən-om-jom-ən ŋaŋ-te-ji kun batte*
 so what-N.SFX-FOC eat-NMLZ-eat-N.SFX get-NPST-3PL that with
aninji mɛeŋ-te-ji
 they live-NPST-3PL
 ‘so whatever food they get that’s what they live on’ (Starosta 1967)

4.2.2 Other subordinate clauses (time, manner, cause, purpose)

Sora makes use of a range of subordinate clause types. One such formation is a kind of temporally subordinate clause of simultaneous ongoing or repeated action. It is formed by the suffix *-atalatə* and appears in a doubled or fully reduplicated form, that is, *X-ata X-ata*, etc.

- (214) *ettegɔj diməd-ata diməd-atə aboj arsi-n*
 like.this sleep-NF:DUR/SIMULT sleep-NF:DUR/SIMULT one monkey-N.SFX
kun a-bānda-leŋ-ən aboj aŋgaj-ən-adoʔɔŋ gi-j-eten
 that OBJ-tank-LOC-N.SFX one moon-N.SFX-OBJ see-T/A:3
 ‘They kept sleeping and sleeping like this and one monkey saw a moon in that tank.’ [Text-3, line 3]

Another common formation is made with a non-finite verb form in *-ən-* in the archaic genitive case form *-a* followed by the subordinating adverb *tiki* ‘after[wards]’.

- (215) (a) *bar bəndi-lə-n-ji-na tiki dəjin dinna de-le anin*
 and put.in.jail-PST-N.SFX-PL-GEN after few day INS he
a-dukəŋɪ-n dɔ a-on-ən-ji jinaŋ doləjən batte kəniət-lə-ji
 3-wife-N.SFX and 3-child-N.SFX-PL also hunger-N.SFX SOC/INS die-PST-PL
 ‘And a few days after they put him in jail, his wife and children also died of starvation.’ [Text-2, line 13]
- (b) *dəjiŋ meŋim ettegoj bara-le-n-ji-a tiki*
 several year like.this work-PST-ITR-PL-GEN after
a-sənna-mər sukku-n a-kako-n-adoʔɔŋ
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ
sarɔba-n er-ti-lə-be anson bara-eten
 paddy.field-N.SFX NEG-give-PST-NF:W/O himself work-T/A: 3
 ‘[After] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.’ [Text-2, line 7]
- (c) *uan-ji kəniət-le-n-a tiki bagu-n-ji jənnəŋ*
 father-3PL die-PST-N.SFX-GEN after two-N.SFX-PL field
sərɔba-n-ji mailen bara-le-ji
 paddy-N.SFX-PL together work-PST-PL
 ‘After their father died, they both worked in their fields and paddies together.’ [Text-2, line 6]
- (d) *ier-ai-ən-a tiki aninji gudeŋ-le*
 go/come-CLOC-N.SFX-GEN after they call-PST
 ‘After he came, he called them.’ [Text-1, line 15]
- (e) *amun a'-ga:t-ti-na' tiki 'anji 'yeŋ-tə-jiy*
 you PFX-eat-NPST-N.SFX:GEN after they go-NPST-3PL
 ‘They went away after you ate.’ [OG]

Sometimes the same subject construction in *X-an X-le[-n]* (see section 4.2.3) can be used in such ‘after’ clauses (temporally subordinate clauses where the action of the first clause precedes that of the second).

- (216) (a) *'anji ga'ga-n gaga'-li-n tm-se'ɽiy 'yer-jiy*
 they REDPL:eat-N.SFX REDPL:eat-PST/CV-ITR here-from go:PST-3PL
 'They went away after eating it.' [OG]
- (b) *'anji ga'ga-n gaga'-li-n tɪnə-se'ɽiy 'yer-ə-jiy*
 they REDPL:eat-N.SFX REDPL:eat-PST/CV-ITR here-from go-PST-3PL
 'They went away after eating it.' [OG] (slow repetition)

Formal variants are attested for seemingly isofunctional complex sentence structures, semantically speaking. Thus, one type of negative subordinate clause in Sora with a first person subject 'since...I not...' may be expressed in either of the following ways.

- (217) (a) *əʔ-naŋ-a jən əpsələ*
 NEG-get-NEG:CV I SUBORD
 'since I didn't get ...'
- (b) *er-i-ən-aŋ-ən asən*
 NEG-take-NEG-take-N.SFX SUBORD
 'since I didn't get ...' [OG]

How these differ remains a subject for the future investigation of the syntax of complex sentences in Sora.

Causally subordinate clauses are formed with the complex (often clause-*initial*) subordinator *ɽitə'nasʔangamle'den* 'because'. Note also the use of the non-finite 'same subject' verbal construction as well in *X-an X-le[-n]*.

- (218) (a) *ɽitə'nasʔangamle'den jən bazar-an 'yɪɽ-an ye'r-[r]i*
 Because I market-N.SFX go-N.SFX go-CV
gaga'na-n 'yi-t-ay
 food-N.SFX buy-NPST-1
 'Because I went to the market, I bought some food.'
- (b) *tə'nasʔangamle'dən jən bazar-an 'yɪɽ-an ye'r-[r]i*
 Because I market-N.SFX go-N.SFX go-CV
gaga'na-n 'yi-t-ay gana'gaŋ 'yi-t-ay
 food-N.SFX buy-NPST-1 food-N.SFX buy-NPST-1
 'Because I went to the market, I bought some food.' [OG]

As mentioned in section 3.2.5, conditionals in Sora are formed by the complex sequence *-le-n-de[:]n*

- (219) *man'yen[ey 'guna dako'-lenden jən ba'zar-ən aʔn-iy-e*
 money find:DEP AUX-COND I market-N.SFX NEG-go-[FUT:]1
 'If I do not find my money, I cannot go to the market.' [OG]

As mentioned in section 3.2.9 a negative or privative subordinate clause in *er-X-le-be* is found meaning 'without having Xed'.

- (220) *dəjiŋ meŋim ettegoj bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX-PL-GEN after
a-sənna-mər sukku-n a-kako-n-adoʔəŋ sarəba-n
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ paddy.field-N.SFX
er-ti-lə-be anson bara-eten
 NEG-give-PST-NF:W/O himself work-T/A:3

'[After] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.' [Text-2, line 7]

4.2.3 Coordination

One of the most common ways of stringing together conjoined clauses in Sora is a structure in which there is either a reduplicated verb with a nominalizing suffix/infixed after the first verb stem and in a past 'participle' form or, perhaps more likely a full copy of the verb stem in the attributive/dependent/noun suffix form *-an*. Subject marking is found only on the final, (fully) finite verb.

- (221) (a) *ukij əp-puŋpuŋ-le-n dəkət-ne po*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n anney kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-say-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 'He puffed himself up even more while saying "this big?"! and his bloated stomach burst and he died.' [Text-1, line 31]
- (b) *bjndə a-ubban bəibəi barab-le ier-an ier-le*
 but 3-younger.brother very get.angry-PST go-N.SFX go-PST
anin adəʔəŋ tuəb-eten
 he OBJ thrash?-T/A: 3
 'But the younger brother got very angry went to him and thrashed him.' [Text-2, line 10]
- (c) *kun arsi-n kun aŋgaj-ən-adəʔəŋ gij-an gij-le*
 that monkey-N.SFX that moon-N.SFX-OBJ see-N.SFX see-PST
a-gaŋij-ji-adəʔəŋ gam-eten
 3-friend-PL-OBJ tell-T/A: 3
 'That monkey saw that moon and told her friends.' [Text-3, line 4]
- (d) *ettegoy anlen-a alale-n-ji ted-an ted-le-n*
 like.this we-GEN REDPL:tail-N.SFX-PL hold-N.SFX hold-PST-N.SFX/ITR
aŋgaj-ən-adəʔəŋ a-təb-n-ai-bə
 moon-N.SFX-OBJ 1PL-get.out-n-CLOC-PL:IMP
 'In this way let's hang by our tails and get out the moon.' [Text-3, line 10]
- (e) *aninji kən gam-ən gam-le ajaŋij-nə a-kəndar-leŋ-ən*
 they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX
səreŋ alan-ji ted-le-n ted-le-n aninji
 from [3-]tail-PL hang-PST-N.SFX hold-PST-N.SFX they
aŋgaj-ən-adəʔəŋ a-nidəb-ben-asən əskai-le-n-ji
 moon-N.SFX-OBJ PFX-pick.up-INF-FOR prepare-PST-ITR-PL
 'They thus discussed and then really began to pick up the moon hanging from the branch from each other's tails.' [Text-3 line 12]

Such formations can also have a causal (dependent clause) sense in Sora under certain conditions.

- (222) *kun asən kun sənna-dəd-ən-ji raʔan-adəʔəŋ giʒ-an giʒ-le*
 DEF for DEF small-frog-N.SFX-PL elephant:N.SFX-OBJ see-N.SFX see-PST
bətəŋ-le iersed-le-ji
 be.frightened-PST run.away-PST-PL
 ‘Because of seeing the elephant, the small frogs were frightened and ran away.’ [Text-1, line 12]

In some instances coordinating particles or conjunctions can be found, often in leftmost clause position.

- (223) (a) *bar bəndi-lə-n-ji-na tiki dəʒin dinna de-le anin*
 and put.in.jail-PST-N.SFX-PL-GEN after few day INS he
a-dukəŋi-ŋ də a-on-ən-ji jinaŋ doləʒən batte kəniət-lə-ji
 3-wife-N.SFX and 3-child-N.SFX-PL also hunger-N.SFX SOC/INS die-PST-PL
 ‘And a few days after they put him in jail, his wife and children also died of starvation.’ [Text-2, line 13]
- (b) *tikki aninji ələŋ-le-n-ji də etenasən a-sə-le-n*
 after they answer-PST-ITR-PL so why 2PL-hide-PST-ITR
gam-le suʔa-dəd-ən
 say-PST big-frog-N.SFX
 ‘After that they answered, and the big-frog said “why did you hide?”’ [Text-1, line 17]

Simple juxtaposition of clauses with no coordinating conjunction or particle is also possible in Sora.

- (224) *aninji təmba-n annəŋ juŋ-le ga-le təgəl-ən annəŋ*
 they noon-N.SFX during eat-PST drink-PST at.night-N.SFX during
aboy məneŋ bənda-n a-bo ara-leŋ-ən
 one edge tank-N.SFX OBJ-ONE tree-LOC-N.SFX
diməd-le-n-ji
 sleep-PST-ITR-PL
 ‘They would eat and drink during the day and at night they would sleep in a tree in one corner of a tank.’ [Text-3, line 2]

Disjunctive sentences can be introduced in Sora with the complex element *bjndə* used in clause-initial position.

- (225) (a) *bjndə a-kako-n kan-ate əmdəŋ-ən əmdəŋ-le*
 but 3-elder.brother-N.SFX that-PRTCL hear-N.SFX hear-PST
də giʒ-an giʒ-le bəibəi barab-le
 and see-N.SFX see-PST very get.angry-PST
 ‘But the elder brother got very angry when he heard and saw (all) that (his brother was doing).’ [Text-2, line 8]
- (b) *bjndə a-ubban bəibəi barab-le ier-an ier-le*
 but 3-younger.brother very get.angry-PST go-N.SFX go-PST
anin a-dəʔəŋ tuəb-eten
 he OBJ thrash?-T/A:3

‘But the younger brother got very angry went to him and thrashed?
him.’ [Text-2, line 10]

5 SEMANTICS/DISCOURSE

5.1 Semantics

The study of the semantics of Sora remains in its infancy and nothing will be offered here. However, one noteworthy aspect of Sora narrative syntax with regards to definiteness that merits mentions is that in addition to demonstrative elements as might be expected, third singular pronouns in Sora are not otherwise used as demonstrative elements as found in a range of languages, but may augment the definite and known qualities of a salient referent.

- (226) (a) *əntannəŋ kuni anin suʔa-dəd-ən ier-ai-ted*
then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
‘then that one, him, the big frog, came back’ [Text-1, line 14]
- (b) *anin suʔa-dəd-ən gam-le suʔə uʔu*
he big-frog-N.SFX say-PST big yes
‘the big frog (, he) said “(he was) big?” “yes!”’ [Text-1, line 20]

5.2 Discourse

As in any language, Sora makes liberal use of various discourse particles, the nuances of which are nearly impossible to render in English translation. Examples of a small number of these are offered below. Formal study of Sora discourse as a whole has not been undertaken. An anthropological study of shamanic and public discourses with the dead among the Sora may be found in Vitebsky (1993).

- (227) (a) *dɪ turkasiŋ-ən anin-ədɔʔəŋ bəndi-lə-ji*
DISC jail-N.SFX he-OBJ put.in.jail-PST-PL
‘They put him in jail.’ [Text-2, line 12]
- (b) *aninji kən gam-ən gam-le aʔaʔij-nə a-kəndar-leŋ-ən*
they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX
səreŋ alan-ji təd-le-n təd-le-n aninji
from [3-]tail-PL hang-PST-N.SFX hold-PST-N.SFX they
aŋgaj-ən-ədɔʔəŋ a-nidəb-ben-asən ɔskai-le-n-ji
moon-N.SFX-OBJ PFX-pick.up-INF-FOR prepare-PST-ITR-PL
‘They thus discussed and then really began to pick up the moon hanging
from the branch from each other’s tails.’ [Text-3, line 12]

The dubitative/interrogative particle *polpa* and so on is especially characteristic of Sora conversational genres.

- (228) (a) *dəkət-ne po gam-le*
this.much-EMPH DOUBT say-PST
‘“This (big)?!” he said’ [Text-1, line 27]
- (b) *anin sukku-n pa a-dukəʔi-n*
she Sukku-N.SFX Q 3-wife-N.SFX
‘Is she Sukku’s wife?’ (Starosta 1967:291)

6 LEXICON

6.1 Austroasiatic/Munda components

The Sora lexicon shows fewer loans than many other Munda languages. Austroasiatic elements and derivational processes abound and the language has an archaic feel about it especially in the make-up and structure of its lexical component (and indeed Sora is archaic in many features of its verbal morphology as well). The following forms were collected in Gajapati district, Orissa in early March 2007.

(229)	<i>siʔi</i> ‘hand’, ‘arm’	<i>boʔb</i> ‘head’
	<i>mû:</i> ‘nose’	<i>amɔr</i> ~ <i>amɔʔd</i> ~ <i>mɔr</i> ‘eye’
	<i>ʃi:</i> ~ <i>jiʔi</i> ~ <i>ʃiʔi</i> ‘tooth’	<i>aluʔd</i> ~ <i>áluʔd</i> ~ <i>áluʔd</i> ‘ear’
	<i>alaj</i> ‘tongue’	<i>uʔu</i> ‘hair’
	<i>tɔʔd</i> ~ <i>tɔʔr</i> ~ <i>tɔʔdʔd</i> ‘mouth’	<i>jiʔiŋ</i> ~ <i>jeʔiŋ</i> ~ <i>jiʔiŋ</i> ‘leg’
	<i>ɔndersi</i> ‘finger’	<i>abobsi</i> ‘thumb’
	<i>senka</i> ‘neck’	<i>kampuj</i> ‘belly’
	<i>kəndɔuj</i> ‘back’	<i>dîàʔ</i> ‘butt’
	<i>kap[ə]ra</i> ‘shoulder’	<i>sanaŋan</i> ~ <i>sanaŋ</i> ‘door’
	<i>sajsaŋ</i> ‘yellow’	<i>aʔsi</i> ‘monkey’
	<i>kambun</i> ‘pig’	<i>kinte</i> ‘banana’
	<i>uʔa</i> ‘mango’	<i>aʔej</i> ‘stone’
	<i>uyuj</i> ‘sun’	<i>aŋgai</i> ‘moon’
	<i>kənsim</i> ~ <i>kansim[a]</i> ‘chicken’	<i>ʔadresim</i> ‘egg’
	<i>aɔaŋ bu:[ʔ]</i> ‘honey bee’	
	<i>jaʔat</i> ‘snake’	<i>jaʔaranji</i> ‘snakes’
	<i>samaiʔ</i> ‘mosquito’	<i>kandɔʔ</i> <i>kandɔd</i> ‘frog’
	<i>gimɔŋlori[ʔ]</i> ‘cashew nut’	<i>sarou</i> ‘rice (growing in paddy)’
	<i>tiʔtin</i> ‘tamarind’	<i>dʔuŋku</i> ‘rice (eaten form)’
	<i>jelou</i> ‘meat’	<i>bɔŋtel</i> ‘buffalo’
	<i>ayou</i> ‘fish’	

6.2 Loan strata

The contacts Sora speakers have had with Indo-Aryan and Dravidian speakers have left their mark. Loans can be seen from Telugu, Oriya varieties, and even English.

(230)	<i>bæg</i>	‘bag, sack’
	<i>saikəl</i>	‘bicycle’
	<i>gouʔi</i>	‘watch’
	<i>pen</i>	‘pen(cil)’
	<i>pənka</i>	‘fan’
	<i>sopa</i>	‘bamboo mat’
	<i>nily</i>	‘blue, purple, violet’
	<i>kuʔap kuʔap</i>	‘light blue’
	<i>kuʔsi</i>	‘chair’

7 BRIEF ANALYSED TEXTS

The following texts were collected in the 1960s by Stanley Starosta from a speaker of the Serango variety of Sora. They are used here with permission from this deceased

linguist, secured orally before he died (he was slated to be the original author of this chapter, but his untimely death obviously prevented this). Translations are the result of both Starosta's work and that of Anderson and Harrison, the authors of the present contribution. Interlinear glossing is solely the responsibility of Anderson and Harrison.

7.1 Text 1: The leader

- (i) *abɔj gailɔ-n dəkɔ-le*
 one road-N.SFX COP-PST
 'There (once) was a road.'
- (ii) *kun a-gailɔ-n-a jattə abɔj ɔʁa-n dəkɔ-le*
 DEF PFX-road-N.SFX-GEN below one ditch/drain-N.SFX COP-PST
 'There was a ditch/drain below that road.'
- (iii) *eɣale kun gəʁan-ə məndraʔ-ji ə-lo-n-a məneɣ ɔʁe*
 how DEF TOWN-GEN person-PL PFX-road-N.SFX-GEN beside or
gailun-a məneɣ ɔʁa-n daku
 river-N.SFX:EN beside ditch/drain-N.SFX be[come]
 'How it is is that the town people have a drain/ditch either beside a river or by a road.'
- (iv) *ənte ettele ɔʁa-leɣ-ən əteɣ kəndud-ən-ji dəkɔ-le*
 thus like.that drain-LOC-N.SFX many frog-N.SFX-PL COP-PST
 'Like that there were many frogs in the ditch.'
- (v) *dijeɣ-antide sua sənna-dud-ən dəkɔ-le-ji*
 how.many 100 little-frog-N.SFX COP-PST-PL
 'How many hundreds of little frogs there were!'
- (vi) *ənte a-sua-leɣ əbɔj suʁa-dud-ən dəkɔ-le*
 thus PFX-100-LOC one/INDEF big-frog-N.SFX COP-PST
 'Thus among (= in) the hundreds, there was one big frog.'
- (vii) *anin kuddib kəndud-ən-ji sireɣ anin suʁa*
 he all frog-N.SFX-PL from he big
 'He is bigger than all the other frogs.'
- (viii) *anin suʁa-dud-ən a-ber-n-an etente beren aninji*
 he big-frog-N.SFX DEP-say-n-N.SFX what say-N.SFX they
sənna-dud-ən-ji mane-le-ji
 small-frog-N.SFX-PL obey-PST-PL
 'Whatever the big frog said, the small frogs obeyed, had to obey.' < manniba
 'obey' Oriya
- (ix) *dɔ bɔ dinna kun a-gailɔ-leɣ abɔj raʔa-n iar-eted*
 so one day DEF OBJ-road-LOC one elephant-N.SFX walk-T/A:3
 'So one day an elephant was walking along the road.'
- (x) *ənt aiəm-leɣ-ən anin suʁa-dud-ən ət-dəkɔ-əd*
 that time-LOC-N.SFX he big-frog-N.SFX NEG-COP-NEG[PST:3]
 'At that time he, the big-frog, wasn't there.'

- (xi) *anin-a suʔuŋ-ban iar-re[...iar-le]*
 he-GEN house-ALL go-PST
 ‘He (elephant) went to his (big-frog’s) house.’
- (xii) *kun asən kun sənna-dud-ən-ji raʔan-adoʔŋ gi-an*
 DEF for DEF small-frog-N.SFX-PL elephant.N.SFX-OBJ see-N.SFX
gij-le bətŋ-le iersed-le-ji
 see-PST be.frightened-PST run.away-PST-PL
 ‘Because of seeing the elephant, the small frogs were frightened and ran away.’
- (xiii) *puttar-leŋ-ən jɔpba-leŋ-ən tiltil-li-ŋ-ji*
 hole-LOC-N.SFX mire-LOC-N.SFX REDPL:bury-PST-ITR/RFLXV-PL
 ‘They buried themselves in holes and mud.’
- (xiv) *əntannəŋ kuni anin suʔa-dud-ən ier-ai-ted*
 then DEF he big-frog-N.SFX go/come-CLOC-T/A:3
 ‘Then that one, him, the big frog, came back.’
- (xv) *ier-ai-en-a tiki aninji gudeŋ-le*
 go/come-CLOC-N.SFX-GEN after he-PL call-PST
 ‘After he came, he called them.’
- (xvi) *uan a-ier-re ənn-ɔləŋ-n-ed-ji*
 where 2PL-go-PST NEG-answer-PST-NEG-PL
 ‘“Where did you go?”; They didn’t answer him.’
- (xvii) *tikki aninji ɔləŋ-le-n-ji dɔ etenasən a-sɔ-le-n*
 after they answer-PST-[ITR]-PL so why 2PL-hide-PST-ITR
gam-le suʔa-dud-ən
 say-PST big-frog-N.SFX
 ‘After that they answered, and the big-frog said “why did you hide?”’
- (xviii) *a-gam-le-n anlen kən təŋɔʔ-ən abɔj sɔ:ʔa-mər-ən*
 1PL-say-PST-ITR we this road-N.SFX INDEF really.big-person-N.SFX
ier-re
 go-PST
 ‘We said “one really big guy travelled this road.”’
- (xix) *əntəpsɛle anlen kən begiʔa a-sɔ-le-n-ai*
 therefore we DEF different.place 1PL-[go.]hide-PST-ITR-1PL
 ‘Therefore we went and hid in (the[se]) different places.’
- (xx) *anin suʔa-dud-ən gam-le suʔə ʊʔu*
 he big-frog-N.SFX say-PST big yes
 ‘The big frog (he) said “(he was) big?” “yes!”’
- (xxi) *dəkəŋən-a suʔa*
 so/yay/how-N.SFX-GEN big
 ‘Yay big??’
- (xxii) *ejə bɔibɔi suʔa*
 No very big
 ‘No, very big.’

- (xxiii) *aninji sənna-dəd-ən-ji gam-le-ji amən sirej boiboi suʔa*
 they small-frog-N.SFX-PL say-PST-PL you from very big
 ‘The small frogs (they) said “much bigger than you.”’
- (xxiv) *ajaʔid po gam-le anin suʔa-mər*
 true DOUBT say-PST he big-person
 ‘“Is that really true?” he, the big guy said.’
- (xxv) *ʊʔʊ amən sirej kərɾa suʔə*
 yes you from much big
 ‘Yes, much bigger than you.’
- (xxvi) *anin suʔa-dəd-ən gam-eted əp-puŋpuŋ-le*
 he big-frog-N.SFX say-T/A CAUS-REDPL:puff-PST
 ‘He, the big frog said puffing himself up.’
- (xxvii) *dəkət-ne po gam-le*
 this.much-EMPH DOUBT say-PST
 ‘“This (big)?!” he said.’
- (xxviii) *ejə ukij suʔe*
 No AUGM big
 ‘No bigger still.’
- (xxix) *ukij əp-puŋpuŋ-le dəkəd-ne po gam-le*
 AUGM CAUS-REDPL:puff-PST this.much DOUBT say-PST
 ‘So he puffed himself up some more and said “this (big)?!”’
- (xxx) *ejə suʔə*
 No big
 ‘Even bigger.’
- (xxxi) *ukij əp-puŋpuŋ-le-n dəkət-ne po*
 AUGM CAUS-REDPL:puff-PST-N.SFX this.much-EMPH DOUBT
a-gam-le-n anney kun ə-puŋ-ən-ə kəmpuŋ-ən
 DEP-say-PST-N.SFX while DEF PFX-bloat-N.SFX-GEN stomach-N.SFX
pətaj-ən pətaj-le kəniət-le
 burst-N.SFX burst-PST die-PST
 ‘He puffed himself up even more while saying “this big?!” and his bloated stomach burst and he died.’
- (xxxii) *aninji sənna-mər-ən-ji əgəndi-le-ji etteden anlen*
 they little-person-N.SFX-PL think-PST-PL therefore we
səənja-n ət-təre-de jitəŋkə-n
 grain.tribute-N.SFX NEG-measure-NPST tax-N.SFX
əd-ji-e gam-le jore-lə-len bij
 NEG-give-NPST COMP twist-PST-1PL MOD
 ‘The little guys thought “we won’t measure out the grain-tribute and won’t pay the tax because he oppressed us.”’
- (xxxiii) *bənlusaj kəniət-le bəŋsa gam-le kun sənna-mər-ən-ji*
 serves.him.right die-PST good say-PST DEF little-person-N.SFX-PL

bɔibɔi eʔaka-le-n-ji
 very rejoice-PST-ITR-PL
 ‘Serves him right that he died! Good! the little guys said and rejoiced greatly.’

7.2 Text 2: The two brothers

- (i) *bagu-mər-an-ji əmele kata-n-a+ber*
 two person-N.SFX-PL about talk-N.SFX-GEN-WORD
 ‘The story of two people’
- (ii) *aninji bagu-n-ji bʌnʌnɔŋ dəku-le-ji*
 they two-N.SFX-PL clan/NMLZ/clan COP-PST-PL
 ‘There were two brothers (of one clan).’
- (iii) *bo-mənra-n ɔnim sukku bar bo-mənra ɔnim mənɔɔra*
 one person-N.SFX 3:name Sukku and one person:GEN 3:name Mangaɔra
 ‘One was named Sukku one was named Mangaɔra.’
- (iv) *bagu-n-ji ɲaŋ-boj-lə-ji*
 two-N.SFX-PL get/take-woman-PST-PL
 ‘They both got married.’
- (v) *bo dinna aninji-a-uan kəniet-le*
 one day they-GEN-[3:]father die-PST
 ‘One day their father died.’
- (vi) *uan-ji kəniet-le-n-a tiki bagu-n-ji jənnəŋ*
 father-3PL die-PST-N.SFX-GEN after two-N.SFX-PL field
səɔba-n-ji mailen bara-le-ji
 paddy-N.SFX-PL together work-PST-PL
 ‘After their father died, they both worked in their paddy-fields together.’
- (vii) *dəjiŋ meɲim ettegoɔy bara-le-n-ji-a tiki*
 several year like.this work-PST-N.SFX-PL-GEN after
a-sənna-mər sukku-n a-kako-n-adoʔɔŋ sarɔba-n
 3-young-man Sukku-N.SFX 3-older.brother-N.SFX-OBJ paddy.field-N.SFX
er-ti-lə-be anson bara-eten
 NEG-give-PST-NF:W/O himself work-T/A: 3
 ‘[After] they worked like this for several years, Sukku the younger brother cultivated the paddy-field himself without giving his older brother the paddy-field.’
- (viii) *bijɔɔ a-kako-n kan-ate əmdəŋ-ən əmdəŋ-le dɔ*
 but 3-elder.brother-N.SFX that-PRTCL hear-N.SFX hear-PST and
gij-an gij-le bɔibɔi barab-le
 see-N.SFX see-PST very get.angry:EMPH-PST
 ‘But the elder brother got very angry when he heard and saw (all) that (his brother was doing).’
- (ix) *ubban-adoʔɔŋ gam-etan amən etenasən*
 younger.brother-OBJ say-T/A-N.SFX/ITR you why

- sərɔbɑ-n* *ət-ti-ŋ*
 paddy.field-N.SFX NEG-give-1
 ‘He said to his younger brother “why (do) not (you) give me some paddy field?”’
- (x) *bijndɔ* *a-ubban* *bɔibɔi* *barab-le* *ier-an* *ier-le*
 but 3-younger.brother very get.angry-PST go-N.SFX go-PST
anin *adɔʔɔŋ* *tuəb-eten*
 he OBJ thrash?-T/A:3
 ‘But the younger brother got very angry went to him and thrashed him.’
- (xi) *kən-ate* *əmdən-le* *turkan-ji* *anin-adɔʔɔŋ* *ŋəm-lə-ji*
 That-PRTCL hear-PST police-N.SFX-PL he-OBJ take-PST-PL
 ‘The police heard about it and seized him.’
- (xii) *dɪ* *turkasij-ən* *anin-adɔʔɔŋ* *bɔndi-lə-ji*
 DISC jail-N.SFX he-OBJ put.in.jail-PST-PL
 ‘They put him in jail.’
- (xiii) *bar* *bɔndi-lə-n-ji-na* *tiki* *dəjin* *dinna* *de-le* *anin*
 and put.in.jail-PST-N.SFX-PL-GEN after few day INS he
a-dukəŋɪ-ŋ *dɔ* *a-on-ən-ji* *jinəŋ* *doləjən* *batte* *kəniət-lə-ji*
 3-wife-N.SFX and 3-child-N.SFX-PL also hunger-N.SFX soc/INS die-PST-PL
 ‘And a few days after they put him in jail, his wife and children also died of starvation.’
- (xiv) *dɔ* *a-iej-ən* *a-mənra* *sərɔbɑ-n* *ət-ti-əd* *anin* *bəndij*
 SO OBJ-who-N.SFX OBJ-man paddy.field-N.SFX NEG-give-NEG he jail:N.SFX
ier-le *dɔ* *a-iej-ən* *ə-mənra* *sərɔb-ən* *əd-ŋəŋ-əd*
 go-PST and OBJ-who-N.SFX OBJ-man paddy.field-N.SFX NEG-get-NEG
anin *suʔuŋ-leŋ-ən* *dəku-le-n* *dɔ* *tiki* *kəlkəl-ən*
 he house-LOC-N.SFX COP-PST-N.SFX and afterwards REDPL:WORRY-N.SFX
batte *anin* *kʌbʌliəd-dəm-le-n*
 with he kill/CAUS/kill-RFLXV-PST-ITR/N.SFX
 ‘The one who didn’t give the fields, he went to jail and the one who didn’t get the fields, he was at home, with worry, and afterwards he killed himself.’
- (xv) *təkod-le*
 finish/end-PST
 ‘The End’

7.3 Text 3: Monkey (moon) shines

- (i) *aboj* *tuləb-leŋ-ən* *dəjeŋ* *arsi-n-ji* *dəku-le-ji*
 one forest-LOC-N.SFX several monkey-N.SFX-PL be-PST-PL
 ‘A number of monkeys were in a certain forest.’
- (ii) *aninji* *təm̩bɑ-n* *ənnəŋ* *jum-le* *gɑ-le* *tɔgəl-ən* *ənnəŋ*
 they noon-N.SFX during eat-PST drink-PST at.night-N.SFX during
aboy *mənəŋ* *bənda-n* *a-bo* *ara-leŋ-ən* *diməd-le-n-ji*
 one edge tank-N.SFX OBJ-one tree-LOC-N.SFX sleep-PST-ITR-PL
 ‘They would eat and drink during the day and at night they would sleep in a tree in one corner of a tank.’

- (iii) *ettegəj diməd-ata diməd-atə aboj arsi-n*
 like.this sleep-NF:DUR/SIMULT sleep-NF:DUR/SIMULT one monkey-N.SFX
kun a-bənda-leŋ-ən aboj aŋgaj-ən-adoʔəŋ gij-eten
 that OBJ-tank-LOC-N.SFX one moon-N.SFX-OBJ see-T/A:3
 ‘They kept sleeping and sleeping like this and one monkey saw a moon in that tank.’
- (iv) *kun arsi-n kun aŋgaj-ən-adoʔəŋ gij-an gij-le*
 that monkey-N.SFX that moon-N.SFX-OBJ see-N.SFX see-PST
a-gaʔin-ji-adoʔəŋ gam-eten
 3-friend-PL-OBJ tell-T/A:3
 ‘That monkey saw the moon and told her friends.’
- (v) *e gəʔin-ji a-gij-ba aŋgaj-en daʔa-leŋ-ən*
 hey friend-PL 1PL-see-PL:IMP moon-N.SFX water-LOC-N.SFX
gəlu-le gam-le ɔppuŋ-eten
 fall-PST QUOT tell-T/A:3
 ‘She told them “Hey friends, let’s go have a look; the moon has fallen in the water.”’
- (vi) *bar anin gam-eten jaba anlen kun aŋgaj-ən-adoʔəŋ*
 and she say-T/A:3 hey we that moon-N.SFX-OBJ
a-təb-n-ai-ba
 1PL-take.out-n-CLOC-PL:IMP
 ‘And she told them “let’s go and take out that moon”’ (CLOC?)
- (vii) *bində a-gaʔin-ji gam-le-ji ian-gamle aŋgaj-ən-adoʔəŋ təb-be*
 but 3-friend-PL say-PST-PL how! moon-N.SFX-OBJ take.out-1PL
 ‘“But how can we get the moon out?” her friends said.’
- (viii) *kun arsi-n gam-eten bo-mənra a-kəndar-leŋ-ən*
 that monkey-N.SFX say-T/A:3 one person OBJ-branch-LOC-N.SFX
a-təd-nə-ba
 1PL-hold/hang-n-PL:IMP
 ‘That monkey told them ‘let one of us guys hold/hang from the tree branch.’
- (ix) *do bo-mənra a-kəndar-leŋ-ən təd-n-eted do anin*
 and one person OBJ-branch-LOC-N.SFX hold-n-T/A and he
ale-n bo-mənra təd-n-ete
 tail-N.SFX one man hold-n-3:IMP
 ‘When that one guy has held onto the branch, let one other guy grab his tail.’
- (x) *ettegoy anlen-a alale-n-ji təd-an təd-le-n*
 like.this we-GEN REDPL:tail-N.SFX hold-N.SFX hold-PST-N.SFX/ITR
aŋgaj-ən-adoʔəŋ a-təb-n-ai-bə
 moon-N.SFX-OBJ 1PL-get.out-n-CLOC-PL:IMP
 ‘In this way let’s hang by our tails and get out the moon.’
- (xi) *aninji aŋgaj-ən-adoʔəŋ a-ŋidəb-ben-asen ɔskai-le-n-ji*
 they moon-N.SFX-OBJ PFX-pick.up-INF-FOR prepare-PST-ITR-PL
 ‘They prepared to pick up the moon.’

- (xii) *aninji kən gam-ən gam-le ajaɽij-nə a-kəndar-leŋ-ən səreŋ*
 they that say-N.SFX say-PST really-EMPH OBJ-branch-LOC-N.SFX from
alan-ji tɛd-le-n tɛd-le-n aninji aŋgaj-ən-adɔʔɔŋ
 [3-]tail-PL hang-PST-N.SFX hold-PST-N.SFX they moon-N.SFX-OBJ
a-ŋidəb-ben-asən ɔskai-le-n-ji
 PFX-pick.up-INF-PURP prepare-PST-ITR-PL
 ‘They thus discussed and then really began to pick up the moon hanging from
 the branch from each other’s tails.’
- (xiii) *kan aiəm-leŋ-ən a-ieŋ-ən arsi a-kəndar-leŋ-ən*
 that time-LOC-N.SFX PFX-who-N.SFX monkey 3/OBJ-branch-LOC-N.SFX
tɛd-n-eten kun a-kəndar-ən əldiŋ-le
 hold-n-T/A:3 that 3-branch-N.SFX break-PST
 ‘At that time, the monkey which hung from the branch of the tree, that branch
 broke.’
- (xiv) *kudub-ən-ji daʔa-leŋ-ən gəlu-le-ji*
 all-N.SFX-PL water-LOC-N.SFX fall-PST-PL
 ‘They all fell into the water.’
- (xv) *eten aŋgaj ɲidəb-ə-ji*
 what moon pick.up-FUT//MOD//T/A-PL
 ‘What moon will they (be able to) pick up.’
- (xvi) *ɔsəntan aninji-dəm tləbhubaj-le-ji*
 in.vain they-self drown/CAUS/-PST-PL
 ‘In vain they got themselves drowned.’
- (xvii) *təkod-le*
 end-PST
 ‘The End’

NOTES

* EDITOR’S NOTE: This chapter more than any other in this volume has a long and storied history. It was originally assigned to the late Stanley Starosta, whose untimely death precluded him from completing it. His family was unable to provide his draft or notes to the editor before his demise. The chapter was then offered to others who reluctantly agreed but whose previous commitments unfortunately prevented them from doing this chapter as well. Therefore, it fell upon the present authors of this chapter who have put this sketch grammar together using unpublished texts from Starosta’s collection given to the editor via Norman Zide before Starosta’s death, along with the results of a brief exploratory field visit to the Sora during the winter of 2007. Much remains to be worked out in the grammar of this Munda language.

- 1 G.V. Ramamurti (1986) also uses vowel symbol [ü] in a few of his dictionary entries, for example, *mūra- ~ mira-* ‘to be frightened’. This may be an archaic feature of the language as Mayurbhanj Ho seems to have two *u* sounds, one phonologically back and one that appears to be phonologically front (Anderson, Osada, and Harrison, this volume). Ramamurti also uses [ɪ] but notes that this sound is one that speakers are ‘apt to confuse with [i] and [e], so that it has not been possible to accurately distinguish between [i] and [ɪ] in all the words in which they occur’ (Ramamurti 1986:107). Clearly, the phonemic status of [ɪ] needs further study.

- 2 The page numbers in this section refer to Ramamurti 1986.
- 3 This table is based on a thorough analysis of Ramamurti (1986) and of our field corpus. We do not exclude the possibility that additional clusters may occur.
- 4 Consonants followed by a glottal stop (in Ramamurti's notation) may in fact be pre-glottalized, for example, {bʔ} = [bʔ̚]. Further analysis of the phonetics is required.
- 5 In Ramamurti's transcription system, syllable coda [y] is transcribed as [i]; for example, *penai'po:d* 'tobacco hoe', *osoimarən* 'flatterer', *sattoi'ber-* 'to declare'.
- 6 Note that this is a pan-Munda feature (dating from the Proto-Munda era) and indeed an archaic derivational process attested throughout the Austroasiatic languages.

REFERENCES

- Anderson, G.D.S. (1997) 'On "animacy maximization" in Fox (Mesquakie)', *International Journal of American Linguistics*, 63(2): 227–247.
- (2003) 'Dravidian influence on Munda', *International Journal of Dravidian Linguistics*, 32(1): 27–48.
- (2004a) *Auxiliary Verb Constructions in Altai-Sayan Turkic*, Turcologica 51, Wiesbaden: Harrassowitz Verlag.
- (2004b) 'Advances in Proto-Munda reconstruction', *Mon-Khmer Studies*, 34: 159–184.
- (2006) *Auxiliary Verb Constructions*, Oxford: Oxford University Press.
- (2007) *The Munda Verb*, Trends in Linguistics Studies and Monographs 174, Berlin: Mouton de Gruyter.
- (this volume, Chapter 13) 'Gta?'.
- Anderson, G.D.S. and O. Gurevich (forthcoming) 'Towards a typology of version: Formal and functional perspectives', *Chicago Linguistic Society*, 41.
- Anderson, G.D.S. and K.D. Harrison (this volume, Chapter 11) 'Remo (Bonda)'.
- Anderson, G.D.S. and F. Rau (this volume, Chapter 8) 'Gorum'.
- Anderson, G.D.S. and N.H. Zide (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The Bimoraic constraint', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Anderson, G.D.S., T. Osada and K.D. Harrison (this volume, Chapter 5) 'Ho and the other Kherwarian languages'.
- Bashir, E. (1985) 'Towards a semantics of the Burushaski Verb', in A. Zide, David Magier and Eric Schiller (eds) *Proceedings of the Conference on Participant Roles: South Asia and Adjacent Areas*, Bloomington: Indiana University Linguistics Club, 1–32.
- Bible Society of India (no date) *Sora Bible*, Bangalore: Bible Society of India.
- Billigiri, Hemmige Shrinivasarangachar (1965) 'The Sora verb: A restricted study', *Lingua*, 15: 231–250.
- Das Gupta, D. (1978) *Linguistic Studies in Juang, Kharia-Thar, Lodha, Mal-Paharia, Ghatoali, Pahariya*, Calcutta: Anthropological Survey of India.
- Hook, P.E. (1991) 'The compound verb in Munda: An areal overview', *Language Sciences*, 13(2): 181–195.
- Hopper, P. and S. Thompson (1980) 'Transitivity in grammar and discourse', *Language*, 56: 251–300.
- Patnaik, M. (this volume, Chapter 9) 'Juang'.
- Peterson, J. (this volume, Chapter 10) 'Kharia'.
- Ramamurti, G.V. (1931) *A Manual of the So:ra: (or Savara) Language*, Madras: Government Press.

- Ramamurti, G.V. (1933/1986) *Sora–English Dictionary*, Delhi: Mittal Publications.
- Sadock, J. (1991) *Autolexical Syntax*, Chicago: University of Chicago Press.
- Starosta, Stanley (1967) *Sora Syntax: A Generative Approach to a Munda Language*, PhD dissertation, University of Wisconsin.
- (1976) ‘Case forms and case relations in Sora’, in P. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austroasiatic Studies, Vol. 2*, Honolulu: University of Hawaii, 1069–1108.
- (1981) ‘Sora noun inflection’, *University of Hawaii Working Papers in Linguistics*, 13(3): 78–117.
- (1992) ‘Sora combining forms and pseudo-compounding’, *Mon-Khmer Studies*, XVIII–XIX: 78–105.
- Vitebsky, Piers (1993) *Dialogues with the Dead: The Discussion of Mortality among the Sora of Eastern India*, Cambridge: Cambridge University Press.
- Zide, Arlene R.K. (1972) ‘Transitive and causative in Gorum’, *Journal of Linguistics*, 8: 201–216.
- (1976) ‘Nominal combining forms in Sora and Gorum’, in P. Jenner, L. Thompson and S. Starosta (eds) *Austroasiatic Studies, Vol. 2*, Honolulu: University of Hawaii, 1259–1293.
- (1983) ‘The story of two girls (excerpt from a Juray text)’, *International Journal of Dravidian Linguistics*, 12(1): 112–125.
- Zide, N.H. (1996) ‘Scripts for Munda languages’, in Peter Daniel and William Bright (eds) *The World’s Writing Systems*, New York: Oxford University Press, 612–618.
- (this volume, Chapter 6) ‘Korku’.

GORUM

*Gregory D.S. Anderson and Felix Rau**

1 INTRODUCTION

1.1 Name of language

For this language (*Ethnologue* code [pcj]) a wide variety of names is in use: Gorum, Parengi/Parenga and Porja (Basa)/Poroja (Basa). Its speakers mostly refer to the language as Porja Basa (Por[o]ja < Oriya *prɔja* ‘people’; Basa < Oriya *bhasa* ‘language’) a name of obvious Indo-Aryan etymology.

The name now most widely used in linguistic and anthropological literature is Gorum. The etymology of this word is unclear, it is most probably of Munda origin and can be segmented into two parts *go-* and *rum*. The second part *rum* may be (related to) the root that occurs in several Munda words for ‘(hu)man’ or ‘cat’ (cf. Gutob *remol* ‘man’, *girem* ‘cat’, the language name *Remo*, or the Gorum word for ‘cat’ *rumaŋ*). The prefix *gV-* very frequently occurs in lexemes denoting animal species or people; it has been suggested to be a fossilized classifier for animate nouns (Anderson and Zide 2002, Smith 1975), but in some languages it is found in inanimate nouns as well. The use of prefixes (among other morphological processes) in free forms of nouns in Munda is a complex problem; for more, see Anderson and Zide (2002). The origin of the name Parengi/Parenga is unclear. In this chapter the tribe and the language will be referred to as Gorum.

1.2 Classification

Since the work of the Chicago Munda Language Project the position of Gorum in the Sora–Juray–Gorum (SJG) branch of South Munda is well established. Its wider connections in South Munda are disputed. Some (Zide and Stampe 1968) believe this group to form a larger subgroup called Koraput Munda, while others (Anderson 2001a) view this to be an areal grouping, and SJG to have split off from the rest of South Munda rather early.

1.3 Number of speakers and location of speakers

The Gorum, classified as a scheduled tribe by the Indian administration, are scattered over Koraput district of southern Orissa, specifically the former Nandapur and Pottangi taluks, and adjacent parts of Munchingput block of Visakhapatnam district, of northern Andhra Pradesh. Within southwestern Orissa, Gutob is spoken to the north of Gorum, with *Gta?* to the west. Estimates of total numbers of speakers range from a handful to a few hundred to a few thousand. This results among other things from a confusing variety of ethnonyms and subgroupings of the tribe. The Gorum apply the name *Gorum* not only to themselves and their inherited language

but also to the neighboring Gutob (Gadaba). This may be the reason why Thurston (1909) treats the Parenga as a section of the Gadaba and also why the language documented under the name Gadaba (Vizagapatnam) in the *Linguistic Survey of India* (Grierson 1906) is actually Gorum. The name Parenga is even more complex. While it was the most common ethnonym for the tribe during the colonial period till recently, it is now sometimes assumed to be one section of the tribe, the Kolai being the other inferior or junior one (Parkin 1991). But the (non-Kolai) informants of Felix Rau insist that the Kolai are also Parenga, at the same time they can offer no name for the non-Kolai section except Parenga. Results of surveys are additionally blurred by the fact that the Gorum prefer the Indo-Aryan Por[o]lja for self-designation, this being a term also used by other tribes and castes speaking different Dravidian or Indo-Aryan languages.

All Gorum speakers are bilingual in Desia Oriya. Among the Gorum the speakers of the language are now a decreasing minority. The authors have met no speaker younger than 20–30 years old. It is apparently not used in village meetings or in rituals any more but is restricted to occasional private communication among elders. It is severely endangered and possibly moribund, at least in Orissa.

1.4 Dialect divisions

There are two communalects, reported by Aze (1971, 1973) and A. Zide (1982), Gorum proper and the less linguistically Aryanized Kolai; the latter differs from the first variety in a number of mainly trivial ways. This division seems to have nearly vanished since that time and could not be verified by Rau during his fieldtrips.

1.5 Literary status

Gorum is unwritten. The second language literacy rate among the Gorum is very low, the *Ethnologue* (Gordon 2005) reports a rate of 11.5% which is probably higher in Andhra Pradesh and lower in Orissa.

Gorum has been studied in English by Sitapati (1933), Bhattacharya (1954), A. Zide (1963, 1966, 1972, 1976, 1982, 1997), Aze (1971, 1973), Aze and Aze (1973) and in Oriya by Mahapatra (1995). It has recently been studied by Senkuttuvan (2002) and Rau (2004).

2 PHONOLOGY

Although phonology is the only part of the grammar of Gorum to which several publications are devoted, a lot of questions remain unsolved. For example Aze (1971, 1973) and A. Zide (1963, 1982) both analyzed the phonological system of Gorum and came to different conclusions. Further research has to be dedicated especially to prosodic and syllabic aspects in particular. The following presentation should therefore be considered no more than preliminary.

2.1 Vowel inventory

Like many languages of the Munda family, Gorum has a triangular five-vowel system, shown in (1).

- (3) *al* ‘hole (for pounding rice)’ [ɑl]
al-u ‘to thatch’
- (4) *da-u* ‘to do’
da? ‘water’ [dɑ^{2a}]
da²d ‘for’ [dɑ²dⁿ]/[dɑ²d^m]

As is well-known, the pre-glottalized or ‘checked’ release of voiced obstruents in coda position is a family level characteristic of Munda, found in both North and South Munda, and indeed is one of the characteristic features of Munda phonology as a whole, distinguishing these languages in this way from Dravidian and Indo-Aryan languages of South Asia. Similarly, an echo or release vowel often accompanies syllables with glottalized nuclei and/or glottal stop in coda position, even in North Munda (cf. the description of Ho (Anderson *et al.*) this volume). Creaky voice, on the other hand, is characteristically Gorum, being absent, for example, even in (most varieties of) the closely related Sora (although Gajapati Sora may well have something cognate with this, see Anderson and Harrison, this volume-b), but may be historically related to various other kinds of suprasegmental phenomena in Munda, for example, Korcu low tone or certain kinds of aspiration in Kharia seen at particular morpheme juncture loci.

In the current stage of research, it remains unclear whether there are one, two, or three contrasting glottal phenomena in Gorum and what their place in the phonological system of the language might be. So whether Gorum should be regarded a register language and thus probably preserves the distinction reconstructed for Proto-Munda (Zide 1976) and Proto-Austroasiatic (Diffloth 1989), or not, is still an unresolved question.

2.3 Consonant inventory

From a South Asian perspective, Gorum consonantism is characterized by a lack of phonemic aspiration and a peripheral dental/retroflex distinction. The status of the dental and retroflex articulation is not as prominent as in other South Asian languages and has a very limited distinctive function in Gorum. All of these Gorum features are typical of Munda vis-à-vis Dravidian and Indo-Aryan; native vocabulary has only /d/ except the word *dumom* ‘cattle’. On the other hand the phoneme /t/ occurs frequently. The retroflex nasal and lateral are missing completely and the difference between dental and retroflex articulation is lost in coda position. Therefore Aze (1971) did not assume a systematic separation of retroflex and dental sounds. Another characteristic of the Gorum sound system is the absence of palatal stops which have been replaced by the alveolar fricatives /s/ [s] and /z/ [z] varying with [dʒ].¹ The palatal nasal is remarkably rare and its phonemic status is disputable. Similar arguments have been made for other Munda languages as well, for example, Ho (Anderson *et al.* this volume).

- (5) *p* *t* *s* (t) *k* ?
b (d) *z* ḍ *g*
m *n* (ɲ) ŋ
r, l ɽ
 y

The series of glottalized unreleased stops mentioned in the preceding section are allophones of their voiced counterparts and occur in the coda or in morpheme-final position. The only question is how to treat [-³j] (graphically represented as *j*, following Mundological tradition) whose counterpart could only be the alveolar fricative /z/ or the palatal approximant /y/; both would be exceptions. But the whole palatal series is exceptional since the palatal nasal /ɲ/ occurs rarely and often sounds less like [ɲ] and more like a slightly nasalized version of the palatal approximant [j̃].

There are a few words with non-glottalized unreleased voiced stops in the coda: [bag] ‘two’ and [yag] ‘three’ but phonologically they might best be represented as *bagu* and *yagulyagi* (cf. so-called Dravidianized Gadba (Gutob) pronunciation (Subba Rao and Patnaik (1992)). Unvoiced obstruents in the coda are rare but there are instances of /p/, /t/ and /k/ all clearly distinguished from the unreleased voiced obstruents by a very audible though slightly delayed aspiration in careful pronunciation. Note that Aze (1971, 1973) does not allow syllable-final unvoiced stops. More experimental work will have to be performed to resolve this issue.

Loanwords (mostly from Desia Oriya) possess a different phonology. Here dental and retroflex sounds in the coda are distinguished and voiced stops are allowed in the coda. Since most speakers of Gorum are primary users of Desia Oriya, the picture drawn above might nowadays not be psychologically adequate anymore. We therefore represent the glottalized obstruents by the combination of a small superscript glottal stop and the corresponding consonant (ʔC). However this glottal stop has to be distinguished from the phoneme /ʔ/ which has been discussed together with the creaky voice in the preceding section.

2.4 Syllable structure and phonotactics

With very few exceptions, Gorum syllables have the structure C₁VC₂ in which C₁, C₂ or both can be empty. There are no consonant clusters in the Gorum syllable, with two exceptions:

There are consonants in the syllable coda which are preceded by their homo-organic nasal alternating with a nasalization of the preceding vowel as in the following examples:

- (6) [rūk^h]~[ruŋk^h] ‘husked rice’
[tāk^h]~[taŋk^h] ‘cooked rice’

In native Gorum words the phenomenon seems to be restricted to [ŋk^h]. Thus there are three possibilities: we can either assume a maximal syllable structure C₁VC₂C₃ with /ŋk/ being the only allowed cluster in the coda, or treat [ŋk^h] as a single segment /ŋk^h/, or regard the vowel as phonologically nasalized. Similar arguments for phonologically pre-nasalized unit segments have been made in the discussion of syllable phonotactics of other Munda languages as well, for example, Santali (Anderson 2001b).

The other exception to the proposed rule for syllable structure is the cluster [dr]/[dʀ] which occurs in some words in syllable onset:

- (7) *druka* [dʀuk^ha] ‘tiger’
indranj [in.draŋ] ‘medicine’

While *druka* clearly violates the assumed syllable structure, *indranj* may actually be *inranj* and thus conform to the assumed maximal syllable like another candidate for

an exceptional syllable structure *mɔnlɔy* ‘five’, which varies between [mɔnlɔj] and [mɔndlɔj].

Additionally there are a few restrictions on single phonemes. The retroflex flap /ɭ/ is restricted to intervocalic position except for the word *kiɑɭ* ‘when (temporal)’ and maybe a few other words but never occurs in the onset. The glottal stop does also not occur in the onset and this same restriction holds true for the velar nasal /ŋ/; two apparent exceptions are discussed in the following paragraph in connection with syllabic properties of segments.

An uninvestigated part of the Gorum syllable is the nucleus. While mostly occupied by one simple vowel, there are instances where the whole syllable consists of an /ŋ/ like in [ŋ.gai] ‘which’ and [ŋ.ɔŋ] ‘darkness’. Another aspect of the nucleus are words with two vowels like *liɔŋ* ‘irrigated field’, *biel* ‘non-irrigated field’ or *aɔ²b* ‘vegetables’. It is unknown whether they should be regarded as disyllabic or monosyllabic. Possible vowel combinations are shown in Box 8.1:

BOX 8.1: POSSIBLE VOWEL COMBINATIONS IN GORUM

i	e	u	ɔ	a
i	<i>biel</i> ‘non-irrigated field’	<i>liɔŋ</i> ‘irrigated field’	<i>miam</i> ‘blood’	
e	<i>erei-nu</i> ‘to nod’			
u	<i>ui</i> ‘to go’		<i>adua</i> ‘sunset’	
ɔ	<i>mɔnlɔi</i> ‘five’		<i>ɔa²</i> ‘to dance’	
a	<i>gai</i> ‘to dig’	<i>aɔ²b</i> ‘vegetables’		

2.5 Intonation/stress

According to Aze (1971:35), in Gorum isolated words, that is, those produced in isolation or not embedded within a larger prosodic structure, stress is generally found in word-final position, if the final syllable is closed; otherwise, primary stress falls on the penultimate syllable in Gorum nominal forms. Gorum verb forms exhibit a range of morpholexical complications that deviate from this idealized system.

2.6 Morphophonology

No more than two sequences of vowels are found in any Gorum stem. In morphemically complex forms, three vowels may arise morphotactically. Generally, the first one is lost (and the means of encoding a grammatical category in some instances). This coalescence may occur in similar morphotactic locations with disallowed or dis-preferred two-vowel sequences as well.

- (8) *le-ɔa²-bu* ‘let’s dance’ > [le.ɔa²bu] ~ [lɔa²bu]
ne-ɔŋ-tu ‘we will sell it’, often [nɔŋtu] (Aze 1971:33)

Consonant clusters in morphologically determined contexts exceed those typically found stem-internally in Gorum, that is, the morphotactics may differ not insignificantly from the phonotactics of uninflected stem forms. For example (Aze 1971:30), clusters of consonant followed by alveolar stop or nasal are restricted to the following morphological contexts: *-Ct-* is limited to verbs with stem-final consonants followed by the *-tu* NPST element, *-Cq-* clusters are basically found in nouns alone with the *-qi* ‘noun auxiliary’, *-Cn-* clusters arise through the use of the common *-nu* GEN/modifier or the *-n(u)* subject affected (usually intransitive) element; lastly clusters in which the initial element is an [r], that is *-rC-*, are limited to (verb) forms preceded by the negative prefix *ar-* (or *ɔr-*).

3 MORPHOLOGY

3.1 Nominal morphology

As is the case with many Munda languages, the real complexity of the morphology of Gorum lies in the verb, but nominal forms nevertheless encode a range of functional categories, including number, a limited system of adnominal relations marked by case-like elements, person, etc.

3.1.1 Number

Gorum distinguishes two number categories in the noun: singular and plural.² While singular is unmarked, plural is marked by the suffix *-gi*. This suffix nearly always takes the last affix slot on the noun.

- (9) (a) *ana-nɔm-gi*
 elder.brother-2SG.POSS-PL
 ‘your elder brothers’
 (b) *gursu²d-nen-gi*
 mosquito-DEM-PL
 ‘These mosquitoes’ (Aze and Aze 1973:300)

The first and second person pronouns are an exception in that they show separate plural stems *bileŋ* and *maiŋ*, respectively, whereas the third person pronoun *no²d* takes the regular plural marker.

The plural suffix does not seem to be used to form honorific forms of nouns as in other South Asian languages. But occasionally another non-standard use is found in juxtaposed singular nouns which together form the subject or object of the clause and are then as a group of singular terms collectively marked with the plural marker:

- (10) *aya-niŋ* *aba-niŋ-gi* *zel-ey*
 mother-1SG.POSS father-1SG.POSS-PL tell-3PL
 ‘My father and mother told’ (Aze and Aze 1973:250)

There are some nouns which possess no plural form, but number agreement in Gorum may also be semantic, as in the case of the Desia Oriya loan *lɔk*, which

triggers plural agreement, although the noun lacks the plural suffix itself, as in the following example:

- (11) *enuŋ-nu lɔk uqubun bay-j=gi*³
 Enung-ATTR people yesterday come-PST-PL
 ‘The people from Enung came yesterday.’

Some expressions are marked by the marker *digin*⁴ which can be analyzed as consisting of the *-di* which will be discussed in section 3.1.4, the plural marker *-gi*, and an element *-n* probably connected with either the locative suffix *-n* or the attributive marker *-nu*. They sometimes are generic expressions but always denote some sort of collective. In accordance with the supposed plural marker component of *digin*, these expressions trigger plural agreement.

- (12) *sedaj-digin bay=gi=ni mij ba?*
 constable-COLL come=3PL.SUBJ=PROG 1SG place
 ‘Officials come to me’

Note also that there is a clitic *=gi* as in the two previous examples which marks third person plural on verbs replacing, or reinforcing, the regular plural suffix *-ey*. Whether the nominal *-gi* and the verbal *=gi* are actually historically identical remains to be investigated. A cognate element is found in numerous other Munda languages.

3.1.2 Case

There are two case forms in Gorum: object/oblique marking and locational/genitive marking.

The formal expression of grammatical case relations is restricted to object forms of pronominals in Gorum. Like in Gutob-Remo, but unlike most other South Asian (or other) languages, this objective case relation is encoded by what might be analyzed as the prefix *e-*. Nouns are combined with the postpositional element *etur*; these nouns can be picked up anaphorically as case-marked pronominals. The postposition *etur* is part of a series of (locational/directional) postpositions: *etur* ‘OBJ’, *ɔtur* ‘from’, *entur* ‘to, towards’, and *ba?tur* ‘with’. The elements *e-*, *ɔ-*, and *en-* are directly connected with the deictic part of the demonstrative system which possesses an *e-*, *ɔ-*, as well as an *en-* series (see section 3.1.7). The remaining *ba?* is also used without the element *-tur* and can be related to the noun *ba?* ‘place’ (cf. *bo[?]* locatives in various other Munda languages; also cf. the Juang comitative/instrumental).

For third singular pronouns the *e-* element is combined with the third person subject pronoun stem *nɔ²d(-gi)* forming *enɔ²d(-gi)*. The first and second person pronouns however are a combination of the *e-* element and the affix form of the respective pronoun either from the series used for object marking on the verb or from the series used for indicating the possessor on the possessed.

From a comparative Munda or even Austroasiatic perspective (cf. Tao-ih, Solntseva 1996), case prefixes attached to pronominal stems are expected. Only the vocalism is unexpected in Gorum. It has perhaps been partially reformed on the basis of the third singular pronoun form, which is deictic in origin.

The objective form also marks experiencer and other ‘dative’ subject-type constructions. In a ditransitive argument structure the recipient/goal participant

is marked as object and so is treated like the object in a monotransitive argument structure. Thus like other South Munda languages Gorum follows the primary object pattern (Dryer 1986).

While the use of *etur* with animate nouns is obligatory, its use with object pronouns and inanimate nouns is optional. The conditions of variation in both cases and whether they are meaningful or not remains unclear.

The adnominal marker *-n*, which may be connected with the modifier *-nu* since the loss of word-final */ul* is very frequent, is used to mark possessors as well as goal or location functions. In its former use, it attaches to any (pro-)nominal stem.

- (13) (a) *mam-nu miam kalay bəbə*
 you-GEN blood mark dear
 ‘your blood stain, my dear’ (Aze and Aze 1973:323)
- (b) *aya-dɔy zel-u dʌʔd mɔdʌ-n gɔʔzɔŋgi-dɔy*
 mother-3.POSS tell-PST for Modu-MODFR trousers-3.POSS
taʔ-ru du pheɖi baʔ tɔʔb-u taʔj
 come.out-PST and trunk place put-PST AUX
 ‘because his mother told him this Modu took off his trousers and put them in the trunk’ (Aze and Aze 1973:229)

The same case element also has a locative or allative meaning (or general ‘oblique’) in certain formations, for example, with certain demonstratives and inherently locational elements.

- (14) (a) *en-leŋ etur begi zel-i-leŋ-ay inɖi-n nay-ta ki*
 OBJ-WE DO quickly tell-EPEN-1PL-CLOC this-LOC what-EMPH Q
mɔ-laʔ-ru mɔ-k-ɥ
 2-hit-PST 2-AUX-AFF
 ‘tell us quickly, what is it that you have put in this?’ (Aze and Aze 1973:323)
- (b) *tɔl-ɥ du ziʔg-dɔy zin-ey du lɔbɔʔ-n*
 fasten-AFF and leg-3.POSS pass-LV and ground-LOC
ɔrgɔʔd-ru ɥi
 trail-PST.AFF AUX:AFF
 ‘when he fastened them, they went past his legs and trailed onto the ground.’ (Aze and Aze 1973:228)

3.1.3 Person

There is one kind of person-marking on nouns: The possessor may be marked for person and number on the possessed noun, mostly on body parts and kin terms (i.e. canonically inalienably possessed nouns), but sometimes on other nouns as well. But this is probably restricted to relations that are conceptualized as inalienable possession or at least socioculturally determined inalienability as in the second example below. For a Gorum family the possession of fields, especially irrigated ones, is socially highly important, since religious rituals concerning the household and the village are connected with the fields and the harvest grown on them. Thus, although a household can *de facto* sell a *liɔŋ* ‘irrigated field’, it remains connected to the former possessor even after generations.

- (15) (a) *nay aba-nɔm-nu aya-nɔm-nu za²g le-tay-ɔm-ay*
 what father-2-GEN mother-2-GEN bone 1PL-give-2OBJ-CLOC
 ‘Shall we give you your parent’s bones?’ (Aze and Aze 1973:318)
- (b) *bileŋ liɔŋ-leŋ pu² bay=ni.*
 We irrigated.field-1PL.POSS rain.water come-PROG
 ‘Rain water is coming into our fields.’

The set of possessor markers shows close resemblance in the first and second person to the respective object markers on the verb. Both sets are given below:

(16)	possessor marker	object marker
1 SG	-niŋ	-iŋ
2 SG	-nɔm	-ɔm
3 SG	-ɖɔy	Ø
1 PL	-leŋ	-(i)leŋ
2 PL	-beŋ	-(i)beŋ
3 PL	-ɖɔy(-gi)	Ø

In the third person plural, A. Zide (1997:255) gives *-ɖɔygi*. Our informants do not distinguish between third person singular and plural possessor and use the plural marker *-gi* only for marking the plural of the possessed. Thus while *ana-ɖɔy-gi* (elder brother-3POSS-PL) with A. Zide’s information can mean ‘his elder brothers’, ‘their elder brothers’, and ‘their elder brother’ our information rules out the last meaning, but then *aba-ɖɔy* (father-3POSS) is ambiguous between ‘his father’ and ‘their father’. The latter meaning is ruled out in the situation described in A. Zide (1997).

There is one phenomenon of person marking on nouns, which occurs in the preamble of some stories and sometimes in the story texts themselves. A noun used to address a second person group – in the case of a story preamble example, like the one below, mostly the audience – is combined with the second person singular possessor marker *-nɔm*. There are no instances of second person plural marking even though the referent is clearly a plural noun. Neither the mechanisms allowing this marking nor its semantics and pragmatics are known.

- (17) *o bubɔŋ-nɔm bay bay bɔ²j sama zel-u taj-u*
 EXCL child-2POSS come come one story tell-INF give-INF
 ‘Oh you children come come I will tell a story.’

3.1.4 Definiteness

Gorum possesses no article or any other device uniquely used for marking definiteness or indefiniteness. Beside demonstratives, pronouns, and proper nouns, which are inherently definite, there are some affixes and clitics signaling definiteness of nouns though not encoding it. All these words have in common that their semantics are only roughly known.

The nominal affix *-ɖi* glossed by Aze and Aze (1973:214) as ‘focus’ and explained to be an equivalent of the definite article seems indeed to be a strong indicator of definiteness. Note however, that it also combines with words that are inherently definite like proper nouns. Aze and Aze (1973:214) have recognized the suffix *-n* to be an allomorph of *-ɖi* used after vowels. We gloss it simply here as *-ɖi*; in forms cited from Aze, we retain his use of *-FOC*.

- (18) *Ramdas-qi ayte?ne, Sobas-qi ayte?ne, Liti-n ayte?ne*
 Ramdas-*qi* like.that Sobas-*qi* like.that Liti-*qi* like.that
 ‘Ramdas is like that, Sobas is like that, Liti is like that’

Two demonstratives *in?i* ‘this’, *a?i* ‘that’ are frequently used adnominally as demonstrative determiners. Both convey a component of definiteness and at least the first probably contains the formative *-qi* discussed above.

Another proximal demonstrative =*nen* appears as a clitic in Gorum and though its semantic and pragmatic effects are far from understood, it functions as a topic marking device, with effects often like those in conditional sentences. It also seems to convey a component of definiteness, which is not surprising when its origin in the demonstrative system is taken into account.

- (19) *gɔʔzɔŋgi-nen ar-man-ey-ij-ay*
 trouser-this NEG-suit-LV-1OBJ-CLOC
 ‘These trousers do not fit me.’ (Aze and Aze 1973:228)

Like in other South Asian languages, object marking on inanimate nouns, where it is optional, may also be connected with definite arguments.

3.1.5 Class/gender

In the grammar of Gorum no morphologically marked class or gender exists. Nevertheless inanimate and animate nouns are in some respects treated differently. As mentioned above animate nouns in object position are obligatorily marked while inanimate nouns are not. Inanimate nouns do not seem to be picked up with anaphoric pronouns. Historically, class may once have been a semi-productively encoded feature of Austroasiatic morphology, possibly seen frozen in a lexicalized way in the Gorum nominal derivation patterns, discussed in section 3.1.10 below. Additionally there are pairs of loanwords from Desia Oriya showing a lexical male/female distinction, for example, *banza* ‘nephew’ and *banzi* ‘niece’ as well as *bu?ka* ‘dwarf’ and *bu?ki* ‘dwarf-girl’. Such couplets form a minor system of masculine/feminine gender oppositions across the Munda languages.

3.1.6 Pronouns

Gorum has a set of subject pronouns and a set of object pronouns. There is no dual or inclusive–exclusive distinction. The possible source of the object pronoun set has been discussed in section 3.1.1. Note that the first person subject pronoun appears to be partially analogically re-formed perhaps on the basis of such forms as the second singular pronoun (with initial *m-* not *n-*). The object forms reflect the earlier forms, based on comparison with other Munda languages.

(20)	Subject	Object
1s	<i>miŋ</i>	<i>eniŋ</i>
2s	<i>maŋ</i>	<i>enɔm</i>
3s	<i>nɔʔd</i>	<i>enɔʔd</i>
1pl	<i>bileŋ</i>	<i>enleŋ</i>
2pl	<i>maiŋ</i>	<i>enbeŋ</i>
3pl	<i>nɔʔdgi</i>	<i>enɔʔdgi</i>

Aze (1973) has an alternative second form *baiŋ* beside *maiŋ* for the second person plural subject, while A. Zide (1997:255) lists only *babiŋ* as the second person plural subject. In the texts published by Aze and Aze (1973) an alternative form *mam* for the second person singular subject also appears.

- (21) *bileŋ da-u le-ru da-u leʔ-ru*
 we DO-INF 1PL.AFF-PST DO-INF 1PL-PST
 'we are doing it.' 'we are doing it.' (Aze and Aze 1973:314)

Zero anaphora is frequent but its pragmatics are not known. The interrogative pronouns of Gorum are *mɔy* 'who' and its objective form *emɔy* 'to whom' for animate referents and *nay* 'what', *enay* 'to what' for inanimate referents and the interrogative pronoun for the possessor *monay*.

3.1.7 Demonstratives

As in other Munda languages the demonstrative system of Gorum is rather complex. It has not been studied until now and on account of the rapid loss of proficient speakers it might be already too late to study it in all its pragmatic nuances. The tables below are a tentative list of the different demonstratives and other words showing a systematic resemblance with the demonstrative system and trying to order them in a sensible manner. They are not supposed to be exhaustive nor are the translations supposed to be minute paraphrases of their semantic or pragmatic content.

- | | | | | |
|------|--------------------|----------------------|-----------------------|----------------------------|
| (22) | <i>e</i> 'this' | <i>en</i> 'this one' | <i>et</i> 'that one' | <i>etur</i> 'OBJECT' |
| | | | <i>ent</i> 'that one' | <i>entur</i> 'to, towards' |
| | <i>inɟi</i> 'this' | <i>aŋi</i> 'that' | | |

The demonstrative pronouns given above do show very little regularity: there is a series of pronouns all having an element *e* combined with an element *n* PROXIMATE or an element *t* DISTAL; *e* 'this' may also stand alone, but how this form contrasts with the other two forms is unknown. The object marking postposition *etur* is probably grammaticalized from this series, Aze (1973b) gives for *etur* the meaning 'to, towards' in addition to the object marking function.⁵ Due to incompleteness of the data, it is unclear whether there is a separate *en*-series, but the existence of the postposition *entur* 'to, towards' is a strong hint that it is indeed a separate series. There are two other demonstratives, which have a DISTAL/PROXIMATE contrast: *inɟi* 'this' and *aŋi* 'that'. They are the most frequent demonstratives of Gorum applied for local and non-local referents which can be animate as well as inanimate and they nearly always appear adnominally.

The *e*-series shows strong resemblance with the locational deictics, which possess a *se*-series and an *ɔ*-series parallel to the *e*-series with *-n* in its proximate and *-t* in its distal form. The *ɔ*-form also forms a corresponding postposition with the formative *-tur*. Both *se*- and *ɔ*-series have a directional component as part of their meaning. The other locational demonstratives *tin/tun* 'here' and *till/tut* 'there' are not directed. Whether or not the vowel alternation is meaningful, as in the positional postpositions *aliŋ* 'on, above' and *aluŋ* 'in, into', is unknown. There is an interrogative *tigay/tugay* 'where' in the locational deixis series. There is also an interrogative for the temporal deixis *agay* 'when', but no corresponding demonstratives seem to exist. There is also a series of manner demonstratives and one for quantity, contrasting PROXIMATE and DISTAL, and corresponding to an interrogative series containing the formative *-ay*.

(23)	<i>se</i> 'this side' ɔ 'yonder'	<i>sen</i> 'this side, in front of' ɔn 'hither'	<i>set</i> 'that side, behind' ɔt 'yonder'	<i>ɔtur</i> 'from'
		<i>tin</i> 'here' <i>tun</i> 'here'	<i>tit</i> 'there' <i>tut</i> 'there'	<i>tigay</i> 'where' <i>tugay</i> 'where' <i>agay</i> 'when' <i>amnay</i> 'how' <i>ijay</i> 'how much'
		<i>ayne?</i> 'like this' <i>ine?</i> 'this much'	<i>ayte?</i> 'like that' <i>ite?</i> 'that much'	

3.1.8 Numerals

The inherited numerals of Gorum have been replaced to various degrees by their Desia Oriya equivalents. While the numerals from 'one' to 'three' are still used frequently, the numerals 'four' and 'five' are not actually used any more but are still remembered today and can be elicited. The numerals for 'six' and 'seven' could not be verified during recent fieldwork on Gorum but are uniformly reported in the older literature.⁶ These numerals are given below:

(24)	<i>bo²j</i>	<i>bag</i>	<i>yag</i>	<i>ungi</i>	<i>mɔnlɔy</i>	<i>turgi</i>	<i>gulgi</i>
		<i>bagu</i>	<i>yagu</i>				
			<i>yagi</i>				
	'one'	'two'	'three'	'four'	'five'	'six'	'seven'

For the numerals beyond 'seven' our knowledge is restricted, and the information is limited and contradictory. N. Zide (1978:61) lists all the variants from different older sources and field notes in the case of Aze and Zide. There are never less than three alternatives, and their forms differ widely. These and the ones from Mahapatra (1995) are given below for each numeral.

(25)	'eight'	<i>tamgi</i> Sitapati (1933)	<i>galgi</i> Zide	<i>gul bo²j</i> Aze	<i>galga</i> Mahapatra (1994)
	'nine'	<i>timgi</i> Sitapati (1933)	<i>algab</i> Zide	<i>galgab</i> Zide	<i>talgi galba</i> Aze Mahapatra (1994)
	'ten'	<i>galgi</i> Aze, Bhattacharya (1954), Sitapati (1933), Zide		<i>galgab</i> Zide	<i>algab alga</i> Zide Mahapatra (1994)

Numbers above 'ten' are always borrowed from Desia Oriya.

3.1.9 Adpositions

Gorum makes use of a range of postpositions in various grammaticalized ways. They can be grouped according to their lexical component. The first group consists of postpositions formed with the formative *-tur*. They mostly encode directional meaning, but the most frequent member of this group is the object marker *etur* that has already been discussed in section 3.1.2 above.

- (26) *abi-nij etur mita²d zum-t-ey-gi dɔm*
 sister-1 DO today eat-NPST-PL-PL must
 ‘Oh today they will eat my sister.’ (Aze and Aze 1973:325)

Other members of that group are: *ɔtur* ‘from’, *etur* ‘to, towards’, and *baʔtur* ‘with’. The last postposition is a combination of *-tur* with *baʔ* which can be used separately as a postposition with a locational/directional meaning and is most likely identical with the noun *baʔ* ‘place’.

- (27) (a) *ɔt ɔtur bay=gi=ni*
 there from come=3PL=PROG
 ‘from there they are coming.’
 (b) *ɔsuŋ-qi baʔ sat ɔsuŋ ne-ɖa²j ne-gi² kɔnɖek kuma²b iŋku²*
 house-FOC place seven house 1-climb 1-see little powder NEG
 ‘I climbed into the lofts of seven houses and saw that there was not even a little powder.’ (Aze and Aze 1973:320)

Another group of postpositions is formed with *-iŋ/-uŋ* and encode stationary topological relations. Members of this group are *aliŋ* ‘on, above’ and *aluŋ* ‘in, into’ and *ɔtiŋ* ‘onto, upon’

- (28) *lu²g aluŋ bɔ²j tsuŋɕia kuntur ɖuku-ry²*
 hole inside one shrew rat be-PST.AFF
 ‘Inside the hole there was a shrew.’ NB: *-ry*. (Aze and Aze 1973:270)

3.1.10 Derivation

As in many South Munda languages (Anderson and Zide 2002), a variety of pan-Munda root forms may be identified in the makeup of a wide range of free forms of nouns in Gorum. The pattern, seen over and over again throughout not just Munda but also Austroasiatic languages as a whole (and the numerous Austroasiatic subgroups) is for certain prefixes, infixes, processes of reduplication, compounding, etc. (which may reflect formerly active processes) to be used in the derivation of acceptable free forms of nouns derived from a common pool of monosyllabic noun stems. That is, most free forms of nouns have some kind of internal morphemic structure. Some of the more salient of these derivational devices are discussed briefly below. It should be noted however, that these processes are not productive anymore and that the internal structure of noun forms are not transparent for the speakers in any sense.

Prefixes used in nominal free forms in Gorum include *pi-*, *bu-/boʔ-*, *u-*, *a-* (<?**n-*), *k/gV[N]-*, *sul/sV-*, infixes *-n-*, *-ʔ-*, and suffixes (?) *-om-*, *-li* (could be compound elements), and the compound prefix elements *auŋ-* and *taŋ-*.

- (29) *pi-*
pitɔm ‘wooden pillow, rolling plate’ *pisap* ‘chicken’
bu- (*could be reduplication)
budi ‘ash’ *bunɔl* ‘soft’ *bubɔŋ* ‘child’* *bubuʔ* ‘snake’*
burɔb ‘lung’ ?? *bɔsuʔ* ‘salt’
- (30) *u-*
usal ‘skin, leather’ *uruŋ* ‘bamboo’

- (31) *a-* (<?*n-)
akaŋ ‘father’s mother’ *akan* ‘elder brother’s wife’
asu ‘sick’ *asu?* ‘younger’
aru²b ‘crab’ *aŋu²b* ‘breast’
alu²b ‘hare’ *asəŋ* ‘shit’ (cf. *asəŋlu?* ‘ear-wax’)
au ‘mother’ *arəy* ‘fly’ *ara?* ‘tree’ (cf. *ara?²sali* ‘tree bark’)
anu ‘elder brother’ *aŋu* ‘younger brother’, ‘brother’s son’
aril ‘hailstone’ *areŋ* ‘stone’ *alaŋ* ‘straw’
- (32) *kV-*
kuril ‘snake’ *kula?* ‘tiger’
kusə? ‘dog’ *kibəŋ* ‘cow buffalo’
kibu? ‘bear’ *kisə* ‘antelope’ *kuŋəŋ* ‘fire’
kikəy ‘hen’ *kuŋur* ‘camel’
- (33) *gV-*
gutə ‘small boy’ *giyei* ‘lizard’ *gusi?* ‘ghost’
gurgi ‘shame’ *gūŋub* ‘fist’ *gəlia* ‘black drongo’
gəmə ‘shoulder’ *gəŋən* ‘throat, neck’
- (34) *kVn-*
kanmun ‘pig’
 ?? *kənun* ‘daughter-in-law’ may also be -n-
kinme²d ‘goat’
kinda? ‘river’
- (35) *su-??*
sua ‘parrot’ *suŋu²b* ‘leaf cap’ *sumə* ‘deer’
səŋdə ‘elephant trunk’ ??

There are at least two distinct reduplication patterns attested in Gorum nominal lexemes. The first type is CVC copy, or CV when there is no C₂ or if C₂ is a stop when there is no copy of the coda consonant.

- (36) *sisi* ‘meat, flesh’ *lulu* ‘shadow’ *səŋsəŋ* ‘turmeric, yellow’
samsam ‘riddle’ *gungum* ‘cheek’ *jəŋjəŋ* ‘storm’
saisai ‘to stir’ *gaga?* ‘rice’ (cf. *ga?a* ‘eat’)
balbalu ‘hot’

The second type is a full copying, compare *balbalu* and *samsam* with *jimijimi*.

- (37) *jimijimi* ‘drizzling’

Sometimes it is hard to know if a prefix or a reduplication is involved in the derivation of the free form of a given noun.

- (38) *səsən* ‘testicle’

Other forms appear to reflect old compounds, some transparent, some not.

- (39) *tə?* ‘mouth’ *tə?rə* ‘snore’

- (40) *auŋ-*
auŋkibu? ‘she-bear’ *auŋkusə?* ‘bitch’ *auŋkula* ‘tigress’

- (41) *taŋ-*
taŋli ‘bullock’ *taŋkur* ‘horse’
mu? ‘nose’ *mu? -lup* ‘nostril’ *mu-ŋe?* ‘snot’

Some forms with final vowels in Gorum may reflect a vocalization of a previous glottal/laryngeal or nasal segment.

- (42) *bulu* ‘thigh’ *pəŋi* ‘bird’

Compound elements or derivational suffixes found in Gorum may include *-li* and *-om*.

- (43) (a) *-li* ??
taŋli ‘bullock’
 (b) *-om* ??
dunom ‘cow’ *pitom* ‘wooden pillow, rolling plate’

Combining forms of nouns found in noun incorporation (section 3.2.11) also serve as the combining form for noun+noun formations in Gorum: *gu?urda?* ‘pus’ (ripe-water)’ Deverbal nouns derived by infixing *-Vn-* are not infrequent in Gorum. If a stem is reduplicated (like ‘sweep’) or has an incorporated noun (‘drink’), the *-n-* form is attached to the stem form lacking these.

- (44) *raj* ‘to comb’ *rinaj* ‘comb’
reŋ ‘keep’ *rinaj* ‘wooden platform over hearth for keeping things’
jujɔ? ‘sweep’ *jɔnɔ* ‘broomstick’

The suffix *-nu* can form nouns from verbs as well in Gorum.

- (45) *aŋdi* ‘to play’ *aŋdinu* ‘game’

Some constructions have been borrowed from Indo-Aryan.

- (46) *amaru* ‘to shave’ *amarulɔk* ‘barber’
amaŋ ‘front’ *amaŋnulɔk* ‘ancestor’ (literally front-GEN/ATTR-person)
panu ‘sew’ *panukar* ‘tailor’
ɔa? ‘dance’ *ɔa?kar* ‘dancer, actor’

Note the following derivationally related set of forms:

- (47) *kun* ‘sing’ *kunkun* ‘song’ *kunkunkar* ‘singer’

Noun+Noun compounding with both elements appearing in a free form is also attested in Gorum.

- (48) *kumda* *don-ay* *ɖu* *kumda* *kol* *bag* *don-ay* *ɖu*
 pumpkin take-CLOC/IMP and pumpkin leaf two take-CLOC/IMP and
ɖa-ɖa²d *ɖu* *tut* *tariŋ* *ɖu* *don-ŋa*
 REDPL-chop and there cook and take-IMP.AFF
 ‘Bring pumpkin and then bring some pumpkin leaves and then chop them
 up and cook them here and take them.’ (Aze and Aze 1973:318)

3.1.11 Adjectives

The lexical category adjective, if it exists at all, consists of a few lexemes like *asa²j* ‘black’, *asel* ‘white’, *lup* ‘big’, and *asu?* ‘small’. There is a limited number of

further candidates but the category adjective is definitely a small closed set. Adjectives precede the noun they modify if they are used attributively and are combined with the verb *ɖuku* ‘to be, to stay’ when they are used predicatively. In contrast to other attributes they are not specially marked in this function.

- (49) *lup saka-ɖɔy yuʔ-ru gaʔ-ru sun-ru*
 big brother-3POSS remove-PST eat-PST say-PST
 ‘The big brother took it and ate.’ (Aze and Aze 1973:322)

More generally modifiers precede nouns and nearly always bear the genitive/attributive modifier suffix *-nu*. This may attach to plural or case-marked nouns as well as to whole phrases or sentences.

- (50) *banʒiʔ-nu lɔk en ɔr-ɖa-ey=gi*
 lazy-MODFR folk this NEG-do-3PL=3PL
 ‘lazy folk will not do this.’ (Aze and Aze 1973:314)

3.1.12 Adverb(*ial*)s

Adverbials in Gorum are uninflected elements that precede the verbs they modify. Adverbials are not frequent and many words used as adverbials are loanwords like in the example below, *begi* ‘quickly’ a word used in Telugu and Oriya, which entered the Gorum lexicon probably via Desia where it is also frequently used. A variety of expressive formations serve as adverbial elements in Gorum like in the second example *tsɔptsɔpa* (see also section 3.3) as well.

- (51) (a) *begi zel-i-leŋ-ay.*
 quickly tell-EPEN-1PL.OBJ-CLOC
 ‘Tell us quickly!’
 (b) *tsɔptsɔpa gaʔ-r-ey laʔ-r-ey*
 quickly eat-PST-3PL.SUBJ hit-PST-3PL
 ‘They gobbled it up.’

3.2 Verbal morphology

3.2.1 Subject

The person and number categories for ‘subject’ (or S/A, i.e. Gorum shows nominative/accusative alignment) are marked by affixes on the verb. These affixes are given below.

- (52) *sg pl*
 1 *ne- le-*
 2 *mɔ- bɔ-*
 3 *-ey/=gi*

The first and second person are marked by prefixes, which are always in the first slot.

- (53) *ne-ab-sɔʔj-t-ɔm*
 1SG-CAUS-learn-NPST-2OBJ
 ‘I will teach you.’

The third person subject marking differs from the first and second person marking. Third person singular is zero marked and overt third person plural marking is

exceptional. There are two plural subject markers in Gorum, *-ey* and *=gi*. The suffix *-ey* is the default marker. It follows the tense marking on the verb. This is the place taken by the object markers, as can also be seen from the fact that *-ey* never occurs together with object markers. With respect to its morphological characteristics, the subject marker *-ey* behaves more like an object marker, even in some syntactic respects, for example, in the progressive construction (discussed below under section 3.2.12), its behavior is parallel to the object markers. But syntactically it agrees with subject pronouns. Additionally it differs from object markers in that it cannot occur together with the cislocative affix *-ay*, which directly follows the object marker. If either an object affix or the cislocative affix is present, or both, the usage of *-ey* is blocked and the marker *=gi* is used. This marker always occupies the last position on the verb (except when followed by the progressive clitic *=ni*). In the case of some complex predicates consisting of more than one finite verb in contrast with the other person markers which occur on both verbs, it normally occurs only once (on the last verb). The following sentences are examples for the behavior of *=gi*. The morphosyntactic properties of *=gi* suggest that it is more likely a clitic.⁷

- (54) (a) *eniŋ laʔ-r-iŋ=gi*
 OBJ:I hit-PST-1OBJ=3PL
 ‘They hit me.’
- (b) *enuŋ-nu lək uqubun dɔuŋ-ŋ-ay*
 Enung-ATTR people yesterday depart-PST-CLOC
bay=gi
 come/PST=3PL.SUBJ
 ‘The people from Enung came here yesterday.’

The two plural suffixes in Gorum verbs very seldom occur at the same time on the same predicate as in the example below. Whether this double subject marking on the verb is meaningful is unknown.

- (55) *baŋgiʔ-nu lək en ɔr-də-ey=gi*
 lazy-MODFR folk this NEG-do-3PL=3PL
 ‘lazy folk will not do this.’ (Aze and Aze 1973:314)

3.2.2 Object types

The ‘object’ (or OBJ) is realized through a series of suffixes. Only first and second person object are marked. In addition to indexing arguments encoding patient, recipient, beneficiary, and even possessor, the ‘object’ series of inflections in Gorum also refer to experiencer arguments as well.

- (56) *aqaʔ-r-iŋ-ay*
 thirsty-PST-1OBJ-CLOC
 ‘I was thirsty.’
- (57) Gorum
- | | | |
|---|------------|--------------|
| | <i>sg</i> | <i>pl</i> |
| 1 | <i>-iŋ</i> | <i>-ileŋ</i> |
| 2 | <i>-ɔm</i> | <i>-ibeŋ</i> |
| 3 | <i>-Ø</i> | <i>-Ø</i> |

These markers coincide in the main in both form and function with both Sora and Juang, suggesting that this likely is a retention of an archaic feature from Proto-Sora-Gorum and Proto-South-Munda (Anderson 2007). Gorum encodes primary objects in the verb but may encode features of the subject as well; that is, the Gorum verb is maximally bi-personal.⁸

- (58) *bɔŋtel mɔ-taʹj-j-iŋ*
 buffalo 2-give-PST-1OBJ
 ‘you gave the buffalo to me.’

In addition to indexing the person and number of subcategorized referents, in Gorum it is also possible to index the animate possessor of the referent that logically fills the verb’s subcategorization frame.

- (59) (a) *miŋ pensil(-nɔm) ne-giʔ-t-ɔm*
 I pencil(-2POSS) 1-see-NPST-2OBJ
 ‘I will see your pencil.’ (Aze 1973:284)⁹
 (b) *puʔipuʔi-nɔm ir-ɔm luʔr-ɔm*
 heart-2POSS beat-2OBJ PROG-2OBJ¹⁰
 ‘your heart is beating.’ (Aze 1973:284)

As these examples suggest, both object and subject ‘possessor raising’ is found in Gorum. Such constructions serve to maximize the overt morphological encoding of highly salient animate referents. Note that as in Sora, many of the possessor-raising constructions are found with verbs exhibiting incorporation of a nominal, frequently a body part. These kinds of incorporative complexes are motivated by the desire to maximize the highly salient, animate possessors in the formal referent-indexing machinery of the verb, which in head-marking languages tends to be the central (and often the only obligatory) part of the clause (see Anderson 1995, 1997).

3.2.3 Tense

Gorum distinguishes two tense categories: past and nonpast. In an affirmative sentence, tense is expressed by a suffix. The past tense suffix in Gorum is *-ru*, a descendent form of the Proto-Sora-Gorum PST **-le*.

- (60) (a) *kapi e-niŋ tagu-r-iŋ-ay*
 coffee OBJ-1 burn-PST-1OBJ-CLOC
 ‘coffee burnt me.’
 (b) *payʔi qiʔ-ru*
 work finish-PST
 ‘s/he finished (that) work.’

When following a liquid, a nasal or /ʔj/, the /r/ is phonologically completely assimilated to the preceding sound. If the word is pronounced in isolation the consonant in the syllable onset can be heard but in actual discourse it becomes acoustically nearly nonexistent, except for an occasional slight lengthening of the consonant.¹¹

- (61) *gulɔŋ-ŋ-iŋ*
 call-PST-1OBJ
 ‘he called me.’

When overtly indexed, the category NPST is marked with the suffix *-tu* in Gorum:

- (62) (a) *miŋ ne-i-tu*
 I 1-go-NPST:AFF
 'I'll go.'
 (b) *mɔ-ta²j-t-iŋ*
 2-give-NPST-1OBJ
 'you will give (it) to me.'
 (c) *ne-la²-tu*
 1-hit-NPST:AFF
 'I'll hit (myself).'

There are however some otherwise finite forms with no overt tense marking like the following one. How this fact fits into the whole tense system is unknown. Maybe it constitutes a further tense category 'neutral'.

- (63) *miŋ asuŋ ne-ku*
 I house 1-be/AFF
 'I am in the house.' (Aze 1973:266)

3.2.4 Aspect

Aspect is expressed in Gorum by auxiliary verb constructions or other complex predicate types, which will be discussed under section 3.2.12. There is however one morphological device for encoding aspect. The progressive clitic *=ni*, attached to past forms of verbs, is an areal phenomenon occurring not only in Gorum but at least also in Gutob, Remo, and some varieties of Desia Oriya. Its origin is unknown, since the construction seems peripheral in the grammars of all the three languages and all three languages possess alternative constructions for expressing a progressive (see Anderson and Harrison this volume-a, Griffiths this volume, Gustafsson 1989).

- (64) *mita²d zi²d nisɔt dɔ²-r-iŋ=ni*
 today EMPH laziness feel-PST-1OBJ=PROG
 'I am feeling tired today.'

3.2.5 Mood

Mood is not an obligatory morphological category in Gorum. The system for expressing modality is rather formally varied, including mood constructions involving affixation, particles, and suppletion. The conditional mood is a reading of the topicalizing construction formed by the suffix, enclitic, or particle *nen/qen*.

- (65) *iŋkɔ, ir nen zi²g si² liga²d-tu ui-tu*
 no jump if leg arm break-F/AFF go-F/AFF
 'No. If I jump my arms and legs will get broken.' (Aze and Aze 1973:226)

It should be noted, however, that as in many Eurasian languages these sentences normally also have a temporal reading. A formally and functionally cognate element is found in Sora (Anderson and Harrison this volume-b), but the morphosyntax of the construction differs in some respects.

As in various other Munda languages, there are suppletive singular imperative forms for a small number of common high-frequency verbs, for example, ‘go’ which is usually *ui* appears as *yɔ*.

- (66) *yɔ kumda kɔl-dʒi ba? yɔ*
 go.IMP pumpkin leaf-FOC place go.IMP
 ‘go, go to the pumpkin leaf place.’ (Aze 1973:318)

Non-singular imperatives are formed by suffixing a marker to the verb, as in Sora (and other Munda languages), and in the first person additionally by prefixing the person marker, and in the third person plural by adding the suffix *-ey*. In the following example, *uay* ‘come’ receives the suffix *-bu* to form the second person plural imperative, which as in many languages can function as a polite imperative form. The verb *ui* ‘go’ is also irregular in the plural, so in the first person plural imperative the root *ui* disappears completely between the prefix *le-* and the suffix *-b[u]*, which thus appears as *leb* ‘let us go’.

- (67) (a) *ana ana begi uay-bu ana e-niŋ*
 elder.brother COPY quickly come-PL.IMP elder.brother OBJ-I
tagu-r-iŋ-ay
 burn-PST-1OBJ-CLOC
 ‘Elder brother, elder brother, come quickly, it is burning me!’ (Aze and Aze 1973:320)
- (b) *aŋi mali kunda-n leb*
 that Mali hill-LOC go/1PL.IMP
 ‘Let us go to the Mali hill over there.’

There are slight differences in inflection depending on whether the verb displays the *n*-element, marking low transitivity, the cislocative marker *-ay*, both, or neither of them. The different imperative paradigms for such verbs are as follows:

- (68) **-u**
- | | SG | DL | PL |
|---|---------------|---------------|-----------------|
| 1 | | <i>le-X-u</i> | <i>le-X-bu</i> |
| 2 | <i>X-∅</i> | | <i>X-ab(u)</i> |
| 3 | <i>X-e-tu</i> | | <i>X-e-t-ey</i> |
- nu?**
- | | SG | DL | PL |
|---|-----------------|-----------------|--------------------|
| 1 | | <i>le-X-nu?</i> | <i>le-X-nu?-bu</i> |
| 2 | <i>X-na?</i> | | <i>X-n-ab(u)</i> |
| 3 | <i>X-ne?-tu</i> | | <i>X-ne?-t-ey</i> |
- ay**
- | | SG | DL | PL |
|---|----------------|----------------|-------------------|
| 1 | | <i>le-X-ay</i> | <i>le-X-ay-bu</i> |
| 2 | <i>X-ay</i> | | <i>X-(ay)-bu</i> |
| 3 | <i>X-ay-tu</i> | | <i>X-ay-t-ey</i> |
- nu?-ay**
- | | SG | DL | PL |
|---|------------------|-------------------|---------------------|
| 1 | | <i>le-X-na?j-</i> | <i>le-X-na?j-bu</i> |
| 2 | <i>X-na?j-</i> | | <i>X-na?j-bu</i> |
| 3 | <i>X-na?j-tu</i> | | <i>X-na?j-t-ey</i> |

Both *-n-* and *-ay-* inflectional patterns are found in Sora; this suggests that an archaic Proto-Sora-Gorum system has been preserved in Gorum. Please note that these paradigms with first dual imperatives are the only place where Gorum distinguishes between dual and plural.¹²

3.2.6 Orientation/directionality

The verb system of Gorum formally encodes categories of verbal orientation or directionality. As is typically the case cross-linguistically, (former) serial verb constructions perform these functions in Gorum (for more examples of these see section 3.2.12 below). The translocative/itive formation is relatively straightforwardly a serial formation, where V_2 is a ‘go’ verb. It marks orientation or motion away from the subject or deictic center.

- (69) (a) *uriʔ-n ui-tu sun-ru bɔbɔ uriʔ-n ʉi*
 other-LOC go-NPST/AFF say-PST dear other-LOC go/AFF
sun sun-ru
 when say-PST
 ‘it’s going off mark, dear, going off mark.’ (Aze and Aze 1973:328)
- (b) *sat bay sun-ru ɔrel dar-r-ey ui-ey sun-ru*
 seven brother say-PST plough grasp-PST-3PL go-3PL.AFF say-PST
 ‘the seven brothers went ploughing.’ (Aze and Aze 1973:315)

The cislocative presents a more complex picture. Formally speaking it appears to be a suffix, appearing usually after tense and object markers. It might derive from a serialization of the verb *uay* meaning ‘come’ grammaticalized to the cislocative suffix *-ay*. Such a development is typical in languages with a grammaticalized cislocative or ventive category (Anderson 2004a, 2006). The situation is however far from clear, since the Gorum verb meaning ‘come’ is highly suppletive. The most frequent nonpast form is *baʔtay* and consists of the cislocative marker *-ay*, the nonpast suffix *-t*, and a root *baʔ*, which might be identical with the noun *baʔ* ‘place’. Whether this is accidental or not is an open question. Occasionally the root *uay* also appears in the nonpast. So *uay* might either be the source of the cislocative marker or on the contrary a frozen combination of a (reduced) root *u* and the cislocative *-ay*. Regardless of its origin, the cislocative suffix *-ay* in Gorum, depending on the verb stem involved and the particular discourse context, marks motion toward the speaker, orientation toward the speaker, and by extension, motion/orientation toward the ‘locus of discourse focus’. In some instances, it has also taken on a meaning of a first person singular referent, which the cognate element clearly has done in certain conjugations in Sora and Juray; a similar functional development has apparently occurred in Iwaidjan languages of Northern Australia as well (Nick Evans, p. c.). It additionally can express among other things spatial or temporal remoteness; so the full range of its semantics are still not sufficiently known.

The prototypical meaning of the cislocative can be seen in the following pair:

- (70) (a) *dɔn-t-ey*
 take-NPST-3PL
 ‘they will take it.’ (Field Notes)
- (b) *dɔn-t-ay=gi*
 take-NPST-CLOC=3PL
 ‘they will bring it.’ (Field Notes)

Another function of *-ay* seems to be connected with deontic modality:

- (71) (a) *miŋ ne-i-tu* (b) *miŋ ui-t-ay*
 I 1-go-NPST/AFF I go-NPST-CLOC/AFF
 ‘I will go.’ (Aze 1973:275) ‘I must go (now).’ (Aze 1973:275)
- (c) *le-i-t-ay*
 1PL-go-NPST-CLOC/AFF
 ‘we must go’ (Aze 1973:275, 277)
- (d) *nɔ̃ʔd gi-n-ay luʔ-r-ay*
 he see-ITR-CLOC/AFF PROG-PST-CLOC
 ‘he (remote) is seeing himself’ or ‘he can be seen (by someone).’ (Aze 1973:289)

In still other Gorum forms, it appears to be taking on the function of the version construction (Anderson and Gurevich 2005); see section 3.2.7 below:

- (72) (a) *gurɔ̃ʔj dɔʔ-r-iŋ* (b) *gurɔ̃ʔj dɔʔ-r-iŋ-ay*
 shy feel-PST-1OBJ shy feel-PST-1OBJ-CLOC
 ‘I felt shy.’ ‘I felt shy.’¹³ (Aze 1973:275)
- (c) *saybu-dʒi zel-iŋ* (d) *saybu-dʒi zel-iŋ-ay*
 master-FOC tell-1OBJ master-FOC tell-1OBJ-CLOC
 ‘the master told me’ ‘the master told me’
- (e) *udubun zalapuʔ ne-ɟ ne-k-ru ɟu nay mad adaʔ-r-iŋ-ay*
 yesterday Z 1-go.AFF 1-AUX-PST.AFF AUX what much thirst-PST-1OBJ-CLOC
 ‘yesterday when I went to Jalaput, I was so thirsty.’ (Aze 1973:262)

As mentioned above, the cislocative element can also be used in Gorum to mark remoteness in time or space. In the following pair of sentences the event in the first sentence is understood to take place in the near future or at a near place, whereas in the second the event will take place in a more distant future or at a remote place.

- (73) (a) *nɔ̃ʔd ɔaʔd laʔ-tu* (b) *nɔ̃ʔd ɔaʔd laʔ-t-ay*
 he dance hit-NPST he dance hit-NPST-CLOC
 ‘He will dance.’ ‘He will dance (later and/or over there).’
 (Aze 1973:274)

3.2.7 Voicelversion

Categories of argument manipulation, voice categories like causative, reflexive, etc. as well as version (Anderson 2001c, Anderson and Gurevich 2005) are an integral part of Gorum grammar. To the former belongs the causative prefix/infix and the detransitivizing *-n* suffix, to the latter, grammaticalized creaky/glottalized voice register.

Like most other South Munda languages, Gorum preserves both prefixal and infixal allomorphs of the Proto (South) Munda (and Proto-Austroasiatic) causative, with the prefix used with original monosyllabic/mono-moraic stems and an infix used with bi-moraic/syllabic ones (Anderson 2004b, Anderson and Zide 2001). Double causatives are always formed by the outer prefix as the second causative marker, regardless of the original means to mark the category and of course of the resulting stem shape.

- (74) *ab-giy-u* ‘show’ < *giʔ-u* ‘see’
bɔptɔŋ ‘frighten’ < *butɔŋ* ‘fear’
ligad-nu ‘to break’ (ITR)
li(b)gad-u ‘to break something’
ab-ligad-nu ‘cause to break, get broken’
ab-li(b)gad-u ‘cause s.o. to break something’ (A. Zide 1972)

The suffix *-n* may be added to stems in Gorum (or was added historically in most instances, as the process is only quasi-productive), to mark various kinds of detransitivized voice categories, for example, reflexive or passive, the latter, as is commonly the case in Munda, with modal capabilitive (or attemptive) semantics as well.

- (75) (a) *nɔʔd giʔ-n-ay luʔr-ay*
 he see-ITR-CLOC:AFF AUX-CLOC
 ‘he is seeing himself, can be seen’ (Aze 1973)
- (b) *bubɔŋ-di ɔriŋ-u luʔru* (c) *bubɔŋ-di ɔriŋ-nu luʔru*
 baby-FOC walk-u AUX baby-FOC walk-ITR.AFF AUX
 ‘the baby is walking.’ ‘the baby is trying to walk (but falls).’
 (Aze 1973)
- (d) *ɔr-ɔriŋ-nu* (e) *ar-samaʔd-n-iŋ*
 NEG-walk-ITR:AFF NEG-recognize-ITR-1OBJ:AFF
 ‘can’t walk’ ‘it was not recognized by me.’ (Aze 1973)

This multiplicity of functions associated with Gorum *-n* (and with parallels in Sora) is paralleled elsewhere in Munda, for example, in various Kherwarian languages (and probably Proto-North Munda too).

- (76) Bhumij
jom-ke-n-a-iŋ
 eat-ASP-ITR-FIN-1
 ‘I was eating.’ ‘I ate.’ (Ramaswami 1992:99)
- (77) (a) Mundari (b) Mundari
suku-le-n-a-ko *diku-n itu-a-n-a*
 happy-ASP-ITR-FIN-PL Hindi-1 teach-APPL/BEN-ITR-FIN
 ‘they had been happy.’ ‘I have been taught Hindi.’
 (Osada 1992:106) (Osada 1992:98)
- (c) Mundari (d) Mundari
tusiŋ-en-tan-a-eʔ *hayagu-n-ja-n-a-ko*
 put.ON-RFLXV-PROG-FIN-3 get.DOWN.-RFLXV-ASP-ITR-FIN-PL
 ‘he is putting it on himself.’ ‘they have just got down by
 (Osada 1992:92) themselves.’ (Osada 1992:92)

In addition to these seemingly related functions, the inflectional affix *-n-* has several other functions in Sora and Gorum, and presumably in Proto-SJG as well. It seems that a particular class of verbs (albeit ones whose semantics are in line with the general ‘middle’ or ‘de-/intransitive’ function of **-n-*) require this *-n-* in all forms throughout the paradigm, including imperative formations.

- (78) sg dl pl sg dl pl
 1 ? ə-X-(n)-ay ə-X-(n)-ay-ba ? le-X-ay le-X-ay-bu
 2 X-(n)-ay X-(n)-ay-ba X-ay X-(ay)-bu
 3 X-(n)-ay-te X-(n)-ay-te-yi X-ay-tu X-ay-t-ey

It has been argued (Anderson 2007, Anderson and Gurevich 2005) that the complex of notions that may be marked by the creaky voice vowel feature in Gorum can be meaningfully subsumed under the category of *version*, in particular subject (or actant) version, which contrasts with neutral (or action) version (ii). Version represents a grammaticalized discourse notion of ‘primary affectedness’ or ‘discourse salience’. Version relates to arguments and their relation to the event encoded by the predicate but it differs from voice manipulation as it neither adds/deletes, nor equates or switches grammatical relations or status of arguments, as traditionally understood voice categories do, which in this way belong to a grammatical layer of lexical event structure; version, much like passive argument defocusing (or in fact functionally just the opposite), highlights arguments in their relation to the particular discourse structure of the narration of the event. In other words voice is a grammaticalized system belonging to the *argument structure* of predicates and is obligatory, while version rather is a grammaticalized system of participants in an *event structure* and is generally optional (it may become obligatory/codified with individual predicates in various different languages (e.g. Turkic, Burushaski)). See Anderson (2001c), Anderson and Gurevich (2005) for more.

Version in Gorum covers a range of formal and functional contrasts. As noted above, it encodes the notions of primary affectedness, discourse salience, and discourse deictic orientation. The *-u* creaky voice segment performs a wide range of functions in Gorum, including antipassive, auto-poesis, passive potential, subject version, and object version.

(79) ◇ Subject affecting

- (a) *B. gosay gaʔa-ru-ni* vs. (b) *gosay gaʔa-ry-ni*
 B rice eat-PST-PROG rice eat-PST:AFF-PROG
 ‘B is eating the rice.’ ‘B is eating the rice.’ (Aze 1973:256)
- (c) *miy ne-aʔa-ry ne-k-ry*
 I 1-thirst-PST.AFF 1-AUX-PST.AFF
 ‘I am thirsty.’ (Aze 1973:255)
- (d) *miy tayku ne-məmɔʔ-ry*
 I cooked.rice 1-REDPL:smell-PST.AFF
 ‘I smelled the cooked rice.’ (Aze 1973:255)
- (e) *aʔkaŋ-nəm-gi ze bəbɔ sunen sun-ru inkuʔ*
 elder.brother’s.wife-2-PL EMPH dear when say-PST NEG
puʔa-di bey-beʔj-di ʔa-r-ey ʔuk-ey
 stomach-FOC REDPL-vomit-FOC do-PST-3PL.AFF AUX-3PL.AFF
 ‘Your wives, my dears, have sick stomachs.’ (Aze 1973:320)

(80) ◇ Object affecting

- (a) *bɔʔj amɔn e-niŋ bam-(m)-iŋ ʔuk-iŋ*
 one arrow OBJ-1 hit-T-1OBJ:AFF AUX-1OBJ:AFF
 ‘(an arrow) has hit me.’ (Aze 1973:298)
- (b) *tileʔj-di neʔ zel-ɔm tay-ɔm-t-ɔm ʔɔm*
 old.man-FOC it tell-2OBJ AUX-2OBJ-NPST-2OBJ:AFF MOD
 ‘it’s the old man, let him tell you.’ (Aze 1973:278)

(81) ◇ Passive Potential

- (a) *aʔi taʔleʔ luʔb-tu* (b) *aʔi taʔleʔ ɔr-luʔb-nu*
 that tablet swallow-NPST.AFF that tablet NEG-swallow-ITR.AFF
 ‘that tablet will/can be swallowed.’ ‘..will/can not.’ (Aze 1973:286)

- (82) \diamond Antipassive
nɔʔd turyaʔ-n-u luʔru
 he kick-ITR-AFF AUX
 ‘he is kicking (in general).’ (Aze 1973:284)
- (83) \diamond ‘semi-reflexive’ or auto-poesis
- | | |
|---|--|
| (a) <i>miŋ siʔ-ɖoy ne-poʔ-tu</i>
I arm-3 1-stab-NPST
‘I will stab his arm.’
(Aze 1973:281) | (b) <i>miŋ siʔ-niŋ ne-poʔ-tu</i>
I arm-1 1-stab-NPST.AFF
‘I will stab my arm.’
(Aze 1973:281) |
|---|--|

This Gorum development appears to be a functional shift from a serialized verb formation encoding a deictic orientation/directionality category (‘come’ > venitive/cislocative) to a category of version or discourse affectedness. A similar grammaticalization path seems to underlie the development of both the frozen, now opaque and lexicalized system of version in Burushaski (Anderson in press, Anderson and Gurevich 2005, cf. Bashir 1985) which like Munda appears to reflect a historical serialized formation with ‘come’, and Siberian Turkic languages (Anderson 2001c, 2004a), which rather show version structures deriving from serialized formations involving the deictic verb ‘take’.

Note that version and indirect ‘experiencer’ subject constructions appear to be in complementary distribution in Gorum.

- | | |
|---|---|
| (84) (a) <i>aɖaʔ-r-iŋ</i>
thirst-PST-1
‘I was thirsty.’ (Aze 1973:307)
oblique subject | (b) <i>ne-aɖaʔ-ryu</i>
1-thirst-PST.AFF
‘I was thirsty.’ (Aze 1973:307)
affected direct subject (+AFF) |
|---|---|

One and the same root which takes its subject marking from the object series of markers (i.e. with ‘oblique’ subject marking) are like the experiencer–subject doublets, that is, they may have optional affective marking as well. They are thus unlike the overt agent passives which similarly use ‘object’ suffixes as subject markers, but obligatorily have version marking.

- | | |
|--|--|
| (85) (a) <i>biɔʔgi payʔi ɖa-u baɖu ɖɔʔ-t-iŋ</i>
tomorrow work do-INF reluctant feel-NPST-1OBJ
‘I will fell reluctant to work tomorrow.’ (Aze 1973:263) | (b) <i>kilnu payʔi ɖa-u baɖu-n-iŋ luʔ-r-iŋ</i>
now work do-INF reluctant-ITR-1OBJ.AFF AUX-PST-1
‘I am reluctant to work now.’ (Aze 1973:263) |
|--|--|

Note that apart from obligatory, grammaticalized uses of version marking as on the overt agent passive constructions mentioned above, version marking on a verb is not an indication of the grammatical relations of the verbal actants involved in Gorum, like a voice category, but rather encodes their status in the discourse space.

- | | |
|--|---|
| (86) (a) <i>nɔʔd Bolram-ɖi etur ɖaʔb-u gotuŋ baʔtur</i>
he B-FOC OBJ COVER-TR cloth with
‘he covered B with a cloth.’ (Aze 1973:256) | (b) <i>B-ɖi ɖaʔb-u ɖuku</i>
B-FOC COVER-TR:AFF AUX.AFF
‘B is covered.’ (Aze 1973:256) |
|--|---|

- (c) *gotuŋ B-d̩i etur d̩aʔb-u*
 cloth B-FOC OBJ COVER-TR:AFF
 ‘the cloth covered B.’ (Aze 1973:256)

In each of the preceding sentences, there are three ‘arguments’ implicit in each form, though only the first expresses all overtly. The grammatical relations of the nominals concerned change from active to passive forms, but not the transitivity marking or the focus status. What changes is the affectedness of the patient Bolram from the perspective of discourse salience. In the last two examples, Bolram is marked as affected, though in the second example he is the subject and in the third the patient. Thus, affectedness has nothing to do with grammatical relations or semantic roles.

3.2.8 Finiteness

Gorum has no special means for marking finiteness. Although it is not always easy to determine what a finite verb in Gorum is, good indications for finiteness are tense marking and the presence of subject markers. On the other hand, object marking does not seem to be restricted to what we would count as finite verbs. From a South Asian perspective Gorum uses finite verbs in positions where non-finite verb forms would be expected. Much remains to be investigated in this domain of Gorum grammar.

3.2.9 Negation

Negative formations in Gorum are relatively straightforward. Negative forms are marked by the prefixes *ar-*, *ɔr-*, and if preceded by a subject prefix, the negative prefix is *r-*. The negation prefix is placed after the subject prefix slot but precedes the causative prefix *ab-*. Negative verb forms take subject and object affixes but do not display the normal tense marking. In the third person, nonpast tense is indicated by the negative prefix *ɔr-*. Past tense is marked by the prefix *ar-*. This distinction is neutralized, following the subject prefix: in the first and second person both forms become *-r-*.

The prefix *ar-* also appears with aspect forms and some modal forms like the *nen/den* forms below.

- | | |
|--|--|
| <p>(87) (a) <i>ar-ui-d̩en</i>
 NEG-go-COND
 ‘if you don’t go’</p> <p>(c) <i>ar-samaʔd-n-iŋ</i>
 NEG-recognize-ITR-LOBJ
 ‘it was not recognized by me.’</p> | <p>(b) <i>mɔ-r-zel-d̩en</i>
 2-NEG-say-COND
 ‘if you do not say’</p> <p>(d) <i>gɔʔzɔŋgi-nen ar-man-ey-iŋ-ay</i>
 TROUSER-THIS NEG-SUIT-LV-LOBJ-CLOC
 ‘these trousers do not fit me.’
 (Aze and Aze 1973:228)</p> |
|--|--|

The other prefix, *ɔr-*, is found in various modal formations, for example, prohibitives and capabilitives in addition to its future/nonpast meaning.

- | | |
|--|---|
| <p>(88) (a) <i>iŋkuʔ ɔr-ɔriŋ-nu</i>
 NEG.COP NEG.NPST-walk-ITR/AFF
 ‘No, (s)he can’t walk.’</p> <p>(c) <i>baŋgiʔ-nu lɔk en ɔr-d̩a-ey=gi</i>
 lazy-ATTR folk this NEG.NPST-do-3PL=3PL
 ‘lazy folk will not do this.’ (Aze 1973:314)</p> | <p>(b) <i>ɔr-alam</i>
 NEG.NPST-touch
 ‘don’t touch’ (Aze 1973:283)</p> |
|--|---|

An alternative prohibitive form, *ambu* is possibly old and/or a loan, cf. Kharia *abu Remo*, etc.

- (89) *ambu alam*
 PROHIB touch
 'don't touch'

The negative copula form is *iyku[ʔ]* or sometimes *iykɔ[ʔ]*. It cannot take tense or person affixes. However, it appears in clause-final position like normal verbs and is used for the negation of existence and the negation of properties.

- (90) (a) *e-nɔʔd nisɔt iykɔ ɔlɔs iykɔ*
 OBJ-s/he him lazy NEG.COP slow NEG.COP
 'He is neither lazy nor slow.'
- (b) *asuy-di baʔ sat asuy ne-daʔj ne-giʔ*
 house-FOC place seven house 1SG.SUBJ-climb 1SG.SUBJ-see
kɔndɛk kumaʔb iykuʔ
 little powder NEG.COP
 'I climbed into the lofts of seven houses and saw that there was not even a little powder.' NB: *asuy* (Aze and Aze 1973:320)

3.2.10 Derivation

Verb derivation as such has not yet been studied extensively in Gorum. There are a number of nominal and verbal stems that are related semantically to one another and for which a denominal or deverbal process of verb derivation can be posited. There are however not many derivational affixes in Gorum, which is not very surprising given the not very pronounced noun/verb distinction in the whole Munda family (for a discussion of the lexical categories of Munda see: Bhat 1997, Evans and Osada 2005, Peterson 2005). The most important type of derivation is probably a zero derivation by which a very large group of nominal and adjectival roots can also be used as verbs, like the pair below.

- (91) *kuilaʔ* 'old woman'
aya-nom kuilaʔ-ru ui
 mother-2SGPOSS old.woman-PST go/PST
 'Your mother has become old (an old woman).'
- (92) *lup* 'big'
aqʉʔd dɔgin lupu-r-ay baʔj=gi=ni
 boy all big-PST-CLOC come=3PL.SUBJ=PROG
 'The boys are growing up.'

As can be seen from the examples above, as a general rule the concept expressed in the nominal usages is associated with an inchoative in the verbal usage. It is unknown which nominal/adjectival roots can be used as verbs.

There are also various processes of compounding or lexically restricted affixation types that may be found in the make-up of seemingly derived verb stems in Gorum.

- (93) *ani* 'louse' *aniṭaṅ* 'press nails to kill lice'

These words may be best analyzed as (frozen) noun incorporations, and are discussed in the next section.

Loan verbs in Gorum tend to take the suffix *-ey* which directly follows the root. Similar phenomena are found in other Munda languages as well.

- (94) *bubɔŋ-dʒi zɛl-iŋ dʒuk-iŋ gɔʔzɔŋgi-niŋ lɔbɔʔ-n*
 baby-FOC tell-1SG.OBJ AUX-1SG.OBJ.AFF trouser-1SG.OBJ ground-LOC
gɔsr-ey-nʉ ui-nʉ luʔ-ru
 land-LV-INF.AFF AUX-INF.AFF AUX-PST
 ‘The youngster was telling me “my trousers reach the ground.”’ (Aze and Aze 1973: 230)

3.2.11 Noun incorporation and combining forms

One of the salient features of Gorum is the presence of a developed system of nominal incorporation (Mithun 1984). As has been described previously (A. Zide 1976, 1997) that speaks to a (in Gorum formerly) elaborate system of nominal combining forms, which in some instances (a) may reflect an older phonological or morphologically undervived form than the corresponding free-standing stem, (b) may be a later phonological clipping, or (c) may be in a suppletive relationship with the corresponding free form.

- (95) *-suŋ* ‘firewood’ *aŋal* ‘firewood’, cf. Kharia *sɔŋɔl, -sɔl*

These short or combining forms appear both in verbal constructions of various types, and in various derived nominalizations and non-finite forms as well.

While incorporation is found in Sora, Juray and Gorum, and is, thus, clearly reconstructable for Proto-Sora-Gorum, it seems to be much less developed in Gorum than in Juray or Sora, or at least more lexically restricted. At least today, in the present stage of language death, all verbs involving noun incorporation are frozen lexemes. There seems to be no mechanism to derive new verb/noun combinations or to derive combining forms from new nouns. The list of verbs involving noun incorporation is, however, rather long and there are candidates of which not every part is semantically transparent, like in the following example:

- (96) *gal-baʔa*
 tie-head
 ‘tie a turban’

The word *daʔ* ‘water’ has come to be used as an almost classifier-type element in a number of Gorum verbs. The following list is far from complete but may provide an impression of how productive the process once must have been.

- (97) *aɖaʔ* ‘be thirsty’ *ɖiɖaʔ* ‘to soak’
guɖaʔ ‘to open up a gully’
zɔɖaʔ ‘to (white) wash the walls’ cf. *zɔɖ* ‘to wipe off’
riɖaʔ ‘to wash (clothes)’
giɖaʔ ‘to drink (water)’
gɔɖaʔ ‘to rinse’
tiɖaʔ ‘to bathe (buffalos) in a tank’

Similar classificatory incorporation is found in Nicobarese as well and may be an archaic feature of Austroasiatic morphosyntax (Anderson 2004b).

3.2.12 Auxiliary verb constructions and complex predicate types

Like most other Munda languages, Gorum makes extensive use of verb+verb structures. Some of the combinations show a functional restriction that suggests these combinations fall under the heading of auxiliary, light, compound, or explicator formations common throughout South Asian and Eurasian languages broadly speaking. Others retain a type of combinatorial semantics that is rather suggestive of serialized formations. The range of formal and functional subtypes of complex or multi-verb predicates in Gorum is briefly presented in the following paragraphs.

Usually such verb+verb formations consist of two verbs, although complex formations with three or more verbs may be found.

Although, generally speaking, taken to mean functional complex predicate subtypes where there is a degree of functional semantic bleaching associated with the second or head element, the auxiliary verb constructions (AVCs) in Gorum, discussed below, themselves fall into several definable formal subgroups. In terms of their grammatical semantics, the functions associated with AVCs in Gorum include a range of *Aktionsart*, aspectual and modal categories, in particular, and also orientation/affectedness/version or voice notions as well. Examples demonstrating functional categories such as deliberate action (98), non-volitional action (99), vigorous action (100), and temporary action (101) include the following:

- (98) ... *ne-gur-r-ay* *ne-ta²j-ay*
 1-enter-PST-CLOC 1-give-CLOC
 'I deliberately entered (your house).' (Aze 1973:278)
- (99) *kula ne-gi²-sun miŋ ne-buton²-tu ne-i-tu*
 tiger 1-see-when I 1-fear-NPST:AFF 1-go-NPST:AFF
 'when I see the tiger, I'll be afraid.' (Aze 1973:279)
- (100) *miŋ ne-ga²-ru ne-la²-ru*
 I 1-eat-PST 1-hit-PST
 'I ate vigorously.' (Aze 1973:279)
- (101) *in²i basa-n le-reŋ-u le-ku²d-u*
 this base-LOC 1PL-leave-TR 1PL-AUX-TR
 'we temporarily leave (our stuff) at this base.' (Aze 1973:279)

From the perspective of the inflectional typology of AVCs (Anderson 2006), Gorum shows a wide range of patterns. Unlike most Munda languages which prefer an AUX-headed inflectional pattern, with obligatory inflectional categories realized on the auxiliary and the lexical verb appearing in a constructionally predetermined form (including bare stem forms), Gorum prefers the doubled pattern of inflection (probably reflecting their origin in Gorum not in a complement structure, but rather a serialized formation, see below). Various split and split/doubled constructions and even the rare LEX-headed pattern are attested in various Gorum AVCs.

The different auxiliaries, which occur in AVCs in Gorum, are given in the list below together with their lexical meaning and a rough label of the function of the AVCs as well as the inflectional type. The auxiliaries *pu²d* 'flower' and *ku²d* 'to sit a child on

one's hip' listed by Aze (1973:279) and Aze and Aze (1973:215) are suspicious. These verbs do not usually occur as auxiliaries cross-linguistically, and are used only rarely as such in Gorum. Combinations involving these two verbs may better be dealt with as lexical phenomena rather than as AVCs.

(102)	<i>taʹj</i>	'give'	benefactive	doubled inflection
	<i>taʹj</i>	'give'	deliberate action	doubled inflection
	<i>ui</i>	'go'	unintentional action	doubled inflection
	<i>uay</i>	'come'	unintentional action	doubled inflection
	<i>laʹ</i>	'hit'	vigorous action	doubled inflection
	<i>puʹd</i>	'flower'	'action moving outwards from a center'	doubled inflection
	<i>kuʹd</i>	'to sit a child on one's hip'	'action of short duration'	doubled inflection
	<i>quku</i>	'be'	'stative'	AUX-headed/doubled
	<i>luʹd</i>	'lift'	'progressive'	AUX-headed
	<i>ɔlku</i>	(no related verb)	'progressive'	AUX-headed

The AUX-headed patterns in Gorum require either an infinitival form of the lexical verb in the case of *luʹd* 'lift' or a past tense (or 'participial') form as in some cases of *quku* 'be'.

- (103) (a) *bileŋ qa-u leʹ-ru*
 we DO-INF 1PL:AUX-PST
 'we are doing it.' (Aze and Aze 1973:314)
- (b) *nɔŋia zɔʹ-ru quku*
 coconut fruit-PST AUX:AFF
 'There are coconuts growing.' (Aze and Aze 1973:222)

The progressive in the first of the two preceding examples is a typical complement construction, while in the AVC involving *quku* 'be' the first verb (V_1) preserves most characteristics of a finite verb as can be seen from the following example, although with regard to tense it is subordinated to the AUX, as we have seen before.

- (104) *ne-aqaʹ-ru ne-k-ru*
 1-thirsty-PST:AFF 1-thirsty-PST:AFF
 'I was thirsty.' (Aze 1973:296)

Doubled person formations, subject, object, or both like in the benefactive AVC below, are especially frequent in Gorum complex predicate formations. They occur in all AVCs except the AUX-headed and the LEX-headed progressive.

- (105) ... *miŋ kiaŋ ne-silay-t-ɔm ne-taʹj-t-ɔm-ay*
 I when 1-sew-NPST-2OBJ 1-give-NPST-2OBJ-CLOC
 'When can I sew for you?' (Aze and Aze 1973:229)

Fully doubled tense/subject forms are also attested in various Gorum auxiliary verb constructions. This is generally the case with the double inflected constructions and with the nonpast construction of the auxiliary *quku* resulting in a habitual construction. Contrary to this, the present perfect as seen above as well as the past perfect forms are formed by an AUX-headed formation and thus the *quku* construction is heterogeneous inflectionally. In other words, this auxiliary verb has been grammaticalized into two separate constructions, an AUX-headed perfect formation and a doubly inflected habitual

formation (cf. the two different AUX-headed formations found in English with the auxiliary *be*, the *be -ing* progressive and the *be -ed/en* passive).

The cislocative marker *-ay* and the affectedness or version marker are encoded either in a split or a doubled manner in different Gorum formations. Each one can appear on either the lexical verb or the auxiliary (or, in various constructions, both).

Other complex split/doubled patterns may be found as well in individual constructions of Gorum. For example, the third plural subject enclitic =*gi* appears preferentially on the auxiliary (or leftmost or head position). The semi-bound conditional/subordinate *nen/qen* marker may also appear in a doubled formation.

- (106) *biti qen ui qen sun-ru bɔbɔ sun-r-ay sun-ru*
 tired if/AFF go if/AFF say-PST dear say-PST-CLOC say-PST
 ‘When they became tired, she said “Dear”.’ (Aze and Aze 1973:328)

The LEX-headed inflectional pattern, where the lexical verb is fully inflected but accompanied by a functional auxiliary head, is found in a small number of Gorum paradigms, including *ɔlku* and arguably the semi-auxiliary modal form *dɔm*.

- (107) *qa-ru ɔlku* ‘s/he was doing.’
qa-tu ɔlku ‘s/he will be doing.’

Thus Gorum possesses three different kinds of progressive, the (possibly borrowed or fused LEX-headed) clitic =*ni* (see section 3.2.4), the AUX-headed *lu²d* ‘lift’ construction and the LEX-headed *ɔlku* progressive.

The origins of inflectional patterns are to be found in the type of structure that gave rise to the grammaticalization of a particular AVC. Thus, structurally similar formations may be found even if the construction itself is an auxiliary or serial formation.

- (108) (a) *ijkɔ, ir nen zi²g si² liga²d-tu ui-tu mita²d*
 no jump if leg arm break-NPST/AFF go-NPST/AFF today
ta ne-ki²d-tu
 EMPH 1SG.SUBJ-die-NPST/AFF
 ‘No. If I jump my arms and legs will get broken. Today I will surely die.’ (Aze and Aze 1973:226)
- (b) *enuŋ-nu lɔk uqubun quŋ-ŋ-ey ui-j-ey*
 Enung-ATTR people yesterday depart-PST-3PL.SUBJ go-PST-3PL.SUBJ
 ‘The people from Enung went away yesterday.’

To be sure, the origin of the doubled inflectional pattern so characteristic of Gorum multi-predicate formations is to be found in two possible source constructions. In the first source, it is frequently the case that certain concepts appear as lexically doubled predicates in Gorum, with two semantically similar inflected verbs, here ‘pound’ and ‘grind’ simply juxtaposed with no intervening conjunction like *qu*.

- (109) *miŋ ui-t-a²j qɔm qu tay-t-ay ɔ²d-t-ay*
 I go-NPST-CLOC/AFF must and pound-NPST-CLOC grind-NPST-CLOC
qɔm qu miŋ qɔniŋ dɔn-t-ay sun-ru sun-ru
 must and I meal take-NPST-CLOC say-PST say-PST
 ‘I must go, I must pound and grind and take the picnic lunches, she said.’
 (Aze and Aze 1973:317)

The other common source construction for doubly inflected AVCs in Gorum is a core serialized formation (Bril 2004, Crowley 2002). Here verbs are concatenated in a combinatorial semantic sequence, each bearing their own inflection. It is relatively straightforward to see how the second of these forms, when it functionally specializes and grammaticalizes into an auxiliary, would simply carry over its original morphosyntax into the target formation.

- (110) *mɔ-tariŋ-tu mɔ-dɔn-tu sun-r-ay sun-ru*
 2-cook-NPST 2-take-NPST say-PST-CLOC say-PST
 ‘you will cook the stuff and take it.’ (Aze and Aze 1973:318)

This is only a rough picture of the AVCs and other complex predicates of Gorum, which are, as we have tried to show, formally very varied, ranging from auxiliary verb constructions to serial verbs constructions and LEX-headed more particle-like constructions, either semi-univerbated as with =*ni* or free-standing as with *ɔlku*.

3.3 Expressives (echo and tag forms)

Although expressive words and related phenomena are a characteristic feature of Gorum and Munda languages in general, they have not been studied in Gorum. Only their existence has been acknowledged and they have been called echo forms, tag formations, or tag word among others. In general, they consist of the augmenting of an element in certain expressive contexts in a lexically determined manner. This augment may be a partially altered reduplication or ‘echo’ form, or a suppletive, opaque element always used immediately following the element in question. Both strategies also exist in Desia Oriya and in other languages of the region.

There are at least two different kinds of usage for echo words: the first is reduplicated words used as a kind of manner adverbial. In these cases, the reduplicated word itself is often semantically and etymologically opaque. The other usage is reduplicated nouns, in this case the reduplication indicates an unspecific quantity of what the respective noun denotes. Expressive forms of both kinds are presented in the following sentences. The last sentence is an example for the suppletive type of echo word.

- (111) (a) *tsɔp-tsɔpa gaʔ-r-ey laʔ-r-ey*
 ‘gulping’-ECHO eat-PST-3PL.SUBJ hit-PST-3PL.SUBJ
 ‘They gobbled it up.’
 (b) *kusa-kasa zel-ey sun-ru uaʔj sun-ru*
 ‘quietly’-ECHO tell-3PL.SUBJ/AFF say-PST come.IMP/AFF say-PST
 ‘they spoke in whispers, come, he said.’ (Aze and Aze 1973:324)
 (c) *kɔŋki bɔʔj dar-ab, gudi bɔʔj dar-ab,*
 mattock one carry-HORT iron.rod one carry-HORT
tsatni-tsutna ɖigin dar-ab
 basket-ECHO COLL carry-HORT
 ‘Let us carry a mattock, an iron rod and some baskets.’
 (d) *ɔ bubɔŋ-gutɔr ɖa-r-ey ɖu sun-ru*
 that baby=ECHO do-PST-3PL.SUBJ/AFF and say-PST
 ‘he had many children.’ (Aze and Aze 1973:333)

4 SYNTAX

4.1 Syntax of the simple sentence

Gorum syntax is fairly typical for South Munda languages. It has clause-final verbs preceded by mainly left-branching modifiers and arguments. That is, Gorum is a predominantly head-final language (with some notable deviations mentioned below). So the main clause pattern is SOV. Of course, as in most natural languages, transitive clauses with two realized noun phrases are not so frequent in actual Gorum discourse.

- (112) (a) [...] *kaɾa-d̪i durdant siɔ̃ rɔza-d̪i etur zel-u*
 rabbit-FOC with.terrible.teeth lion king-FOC OBJ tell-PST
 '[...] the rabbit told the awesome lion lord.' (Aze and Aze 1973:238)
- (b) *bɔlram kristel etur lɔruŋ li²d-ru*
 Bolram Christel OBJ oil rub-PST
 'Bolram rubbed oil on Christel.' (Aze 1973:249)

From the preceding examples we can also see that additional arguments or adjuncts, like instruments and other more verb dominated functions tend to take the interjacent place between object and verb and are not marked morphologically. Adverbials with a scope over the whole clause, like temporals or locatives, also frequently occur in clause initial position. These functions also sometimes occur in the post-verbal position, which appears to be discourse active, and are discussed in section 5 below. In the following example the temporal *mita²d* 'today' is in clause-initial position.

- (113) *mita²d, siɔ̃ rɔza-d̪i etur ne-d̪i²-r-ay ne-ta²j-ay.*
 today lion lord-FOC OBJ 1-finish-PST-CLOC 1-AUX-CLOC
d̪u n-a²j-ay, miŋ
 and 1-come-CLOC I
 'Today I finished off the lion lord and I have come. Yes, me!' (Aze and Aze 1973:241).

Yes/no questions may be marked by intonation or by combining the assertive sentence with the question tag *ki iŋkɔ*. This tag consists of the Gorum negative copula *iŋkɔ* and the question marker particle *ki*. This particle is a loan from Desia Oriya. There may be a pragmatic difference, perhaps in respect to the presence of a bias, between the two questions or they may be total equivalents. As in many languages, such questions are typically answered by repeating the verb in question with a change of subject.

- (114) (a) *lai mɔ-ga²-ru ki iŋkɔ*
 millet.gruel 2-eat-PST Q NEG.COP
 'Have you eaten the millet gruel (, or not)?'
- (b) *za²g mɔ-katr-ey-u ne-katr-ey-u*
 bone 2-chew-LV-PST 1-chew-LV-PST
 'have you chewed the bones?' 'I have.' (Aze and Aze 1973:331–332)

Existential and copular-type sentences, including negative copular forms, appear with the copular element if present in final position as well.

- (115) *aytarɔm ɔtur nai paiŋi iŋkɔ*
 sunday from what work NEG.COP
 'From sunday on there will not be any work (to do).'

4.1.1 Typological features

As expected in a Eurasian SOV language, the internal structure of Gorum predominantly shows an Operator–Operand or Dependent–Head pattern.

- (116) *may e-nij mɔ-zel-l-ij*
 you OBJ-I 2-tell-PST-1OBJ
 ‘you told (to) me.’ (Bhattacharya 1975:161)

Typologically, Gorum is, in contrast with most South Asian (or even Eurasian) languages, not a language which predominantly employs converbs as a clause linkage strategy, but prefers the combination of finite verbs sometimes showing more resemblance to serializing languages. This can among others be seen from the linkage types involved in the formation of the auxiliary verb constructions discussed in section 3.2.12.

Historically, the syntax may not have been always so predominantly of the SOV type. As discussed under section 3.1 the verb has an object suffix but, with the exception of the third person, a subject prefix. The fused demonstrative *nen* which has been grammaticalized as a topicalizing and subordinating device follows the noun, which, if it has not evolved from a combination of a word or phrase marked with the attributive *-nu* and the demonstrative *en*, could be a hint that historically the order was noun demonstrative and not demonstrative noun as it is in present-day Gorum. Noun-numeral order is found in some Munda languages, however (e.g. Remo), at least as an alternate order.

- (117) *gɔʔzɔŋgi-nen ar-man-ey-ij-ay*
 trouser-this NEG-suit-LV-1OBJ-CLOC
 ‘these trousers do not fit me.’ (Aze and Aze 1973:228)

The same variation in the order of dependent and head is seen with possessive constructions. The series of possessive markers discussed in section 3.1.3 may reflect a possible earlier possessum–possessor configuration, and the possession formation with the general modifier or attributive element *-nu* shows the synchronically predominant dependent-head order.

- (118) *aya-dɔy zel-u dʌ²d mɔdu-n gɔʔzɔŋgi-dɔy*
 mother-3.POSS tell-PST for Modu-ATTR trousers-3.POSS
taʔ-ru dʌ phedʒi baʔ tɔ²b-u ta²j
 come.out-PST and trunk place put-PST AUX
 ‘because his mother told him this Modu took off his trousers and put them in the trunk.’ (Aze and Aze 1973:229)

4.2 Complex sentence structure

A variety of strategies are attested that serve to encode a range of functional subtypes of complex sentences in Gorum. While coordination mostly involves either the marker *dʌ* or mere juxtaposition, subordination structurally speaking usually entails a preposed quasi- or overtly nominalized verb form or a verb in the infinitive form followed by a postpositional element. Quotatives usually mark speech but are also used for a variety of other functions. The particle or clitic *dʌ* coordinates clauses mainly with subject coreference. The different clause combining strategies are discussed in the following sections.

4.2.1 *Relative-type clauses*

Clauses in the delimiting relative-type function in Gorum consist of preposing a clause with all its attendant internal morphosyntax and adding the suffix/clitic *-nu*. This phrase now modifies the noun it precedes.

- (119) *e-nɔʔd* *tiŋ-ey* *laʔ-r-ey-nu* *lɔk*
 OBJ-S/he shoot-3PL AUX-PST-3PL-ATTR folk
 ‘the folks who shot her.’ (Aze and Aze 1973:333)

The attributive function performed by this kind of clauses is also reflected in the fact that words, phrases, and clauses are all marked by the same formative *-nu* when used attributively (see also above section 3.1.11). Verbs in the infinitive form occasionally occur in this function as well, as in the following sentences where the infinitive form of the verbs is formed by *-u*. Intransitive verbs, however, form their infinitive with the suffix *-nu*, but this may derive from **-n-u*.

- (120) (a) *bileŋ* *baŋgiʔ=buzɔl* *nɔyte* *ɖa-u* *lɔk* *ɖu* *ɖa-u*
 we lazy=ECHO not do-INF folk and do-INF
le-ʔ-ru
 1PL.SUBJ-AUX-PST
 ‘we are not lazy folk who do nothing.’ (Aze and Aze 1973:314)
- (b) *tɔnan-ɖɔy-nu* *sisiʔd* *zum-u* *sama*
 sister-3.POSS-GEN meat eat-INF story
 ‘the story of eating their sister’s flesh.’ (Aze and Aze 1973:315)

The last example shows that – the genitive marker is also *-nu* as already mentioned under section 3.1.2. So the relative clause-like function and attributively functioning nouns or noun phrases are marked by *-nu* as are possessors and infinitives of intransitive verbs. Infinitives, as seen above, occur sometimes in environments where the relative marking *-nu* would be expected. More investigation is needed to show the various uses of the marker *-nu* and how they are related.

4.2.2 *Other subordinate clauses (time, manner, cause, purpose)*

Subordinate clauses in Gorum may either consist of a clause with an infinitive verb directly combined with the finite verb or the clause may be marked by postpositional elements like *ɖaʔd* ‘because, for’ which appears with a verb in the infinitive form, or with clitics like *nen/ɖen* ‘if/when’, which generally follow an unmarked form of the verb; these may be combined to form the complex subordinator *ɖaʔdnen* ‘because’.

- (121) (a) *bileŋ* *uʔa* *yuʔ-u* *le-j*
 we mango remove-INF 1PL-GO:PST:AFF
 ‘We went to collect mangoes.’ (Aze 1973:304)
- (b) *dinek* *kuntur-ɖi* *kinte* *zum-u* *ɖaʔd* *taʔ-r-ay*
 one.day rat-FOC grass eat-INF for come.out-PST-CLOC
 ‘One day the shrew came out to eat some grass.’ (Aze and Aze 1973:270)
- (c) *biba* *asuj* *ui* *nen* *e-niŋ* *liɖa-t-iŋ*
 wedding house go if OBJ-I ridicule-NPST-IOBJ

- laʔ-t-iŋ=gi*
 AUX-NPST-1OBJ=3PL
 ‘If I go to the wedding, they will laugh at me.’ (Aze and Aze 1973:228)
- (d) *silay-iŋ tay-iŋ sun-ru ɖaʔdnen kuilaʔ-dʒi zel-u*
 sew-1OBJ AUX-1OBJ say-PST because old.woman-FOC tell-PST
- miŋ tɔ din rati zakɔ e-niŋ ɖimaʔd ɖuk-iŋ-ay*
 I EMPH day night item OBJ-I sleep AUX-1OBJ/AFF-CLOC
- miŋ kiaʔ ne-silay-t-ɔm ne-taʔj-t-ɔm-ay*
 I when 1-sew-NPST-2OBJ 1-AUX-NPST-2OBJ-CLOC
 ‘because he said “sew it for me” the old woman said “Me?! Day and night I have to sleep. When can I sew for you?!”’ (Aze and Aze 1973:229)

There is a further subordinator *ziʔ*, sometimes realized as *zi*, meaning roughly ‘if’ or ‘although’ or ‘after’.

- (122) *bar-bɔʔj ɔl-ay ziʔ ne-r-ui.*
 again-one write-CLOC after 1-NEG-go
 ‘After/although he wrote again, I still did not go.’ (Aze and Aze 1973:249)

The verb *sun* ‘say’ marks complement and quotative structures. It appears in certain narrative genres in practically every line. Combined with the subordinator *nen* the resulting *sunen* seems to function almost identically like the simple *nen*.

- (123) *inʒɔʔ sunen, gagaʔ-gigiʔ gaʔ-tu laʔ-tu.*
 not when.if cooked.rice-ECHO eat-NPST AUX-NPST
 ‘Should this not be done, then they [the dogs] will eat all the rice meals.’ (Aze and Aze 1973:286)

Its bare form *sun* may be used in a subordinating function as well. Its other functions are discussed under section 4.2.3.

- (124) *le-luʔd-ru sun ayuʔ baʔj=gi sun,*
 1PL-lift-PST when fish come=3PL when
le-leʔd-tu, garmaʔd baʔ bɔʔj le-ɔrɔŋ-tu
 1PL-catch-NPST bamboo.bag place one 1PL-pack-NPST
le-taʔj-tu
 1PL-give-NPST
 ‘When we lift it up, when the fish come, we will catch them and pack them into the bamboo bag.’ (Aze and Aze 1973:279)

Other subordinator combinations of, yet, unknown functional types are also found in Gorum, for example, the complex *ɖen sun-ru* in the following example.

- (125) *dar ɖen sun-ru nɔʔd-gi sun-ru aʔi bal-u*
 grasp COND say-PST s/he-PL say-PST that cook-INF
bal-u-nu-ʔa sun-ru sɔbu zum-u laʔ-u luʔr=ey sun-ru
 COOK-INF-MODFR-EMPH say-PST all eat-INF AUX-INF AUX=3PL say-PST
 ‘when he held it, they ate heartily, all that had been warmed up.’ (Aze and Aze 1973:331)

One peculiar feature of Gorum complex sentence structure is that the conditional/subordinate clitic or particle *nen/den* may appear in a doubled inflectional pattern in auxiliary structures. The consequences for the analysis of the syntax of *nen* and the auxiliary verb constructions have not been fully explored.

- (126) *silay nen ta²j nen, bar tsɔt qa-ru uj*
 sew if AUX if again short do-PST/AFF AUX/AFF
 ‘If it was sewn, it would become even shorter.’ (Aze and Aze 1973:230)

4.2.3 Coordination and switch reference

Like most South Munda languages, Gorum has certain functional elements that, in addition to or instead of their typical conversational genre, are found in particular narrative discourse genres and seem to be embedded within a switch reference system. The elements are the same subject conjunctive *qu* and the different subject conditional/subordinator *nen/den*.

- (127) *ta² nen bɔ²j kua kuntur-qi etur qu²b-u dɔn-ru*
 come.out DS/COND one crow rat-FOC DO peck-PST take-PST
lɔm-u dɔn-ru
 bite-PST take-PST
 ‘it came out, and a certain crow pecked the shrew, bit the shrew, and took it away.’ (Aze and Aze 1973:270)
- (128) (a) *qu²-ay ba²j qu tarbulqa²d ɔtur zalapu² qu²-u uj*
 depart-CLOC come and Talabireda from Jalaput depart-PST go/AFF
 ‘Having departed and come, he departed from Talabireda and went to Jalaput.’ (Aze and Aze 1973:249)
- (b) *zoipur ne-ma²ta-ru² qu marba²ri dugan*
 Jeypur 1-lower-PST:AFF and Marwari shop
ne-i² qu bɔ²j tsɔka ne-apa²d-ru
 1-go/AFF and one shirt 1-sew-PST
 ‘Having got off (the bus) at Jeypur I went to Marwari’s shop and had a shirt made’ (Aze and Aze 1973: 251)

Sometimes same subject clauses can be simply juxtaposed without a conjunction or subject coreference marker like the take-wash-take sequence in the following example.

- (129) *le-ga²d-ru le-ta²j qu bar galmɔ²j le-qa-ru*
 1PL-cut-PST 1PL-AUX and again yank 1PL-do-PST
qu kin²qa²-n le-dɔn-r-ay le-ti²qa²-r-ay
 and river-LOC 1PL-take-PST-CLOC 1PL-wash-PST-CLOC
le-dɔn-r-ay qu le-gandi²-ru pɔrbu u²an
 1PL-take-PST-CLOC and 1PL-fasten-PST:AFF festival time
 ‘We deliberately cut it and then we did it up into yanks, we took it to the river and washed it, and then we put it on at a festival time.’ (Aze and Aze 1973:275)

In many languages, among them Eurasian languages of the Indo-Aryan and Dravidian families, various forms of the verb ‘say’ have been grammaticalized into a variety of subordinate clause markers. Such functions typically include quotatives, complementizers, narrative devices, and evidentials. These perhaps represent

one original grammaticalization into the marker of a quotative, with three separate subsequent chains of semantic development into evidentials, complementizers, and other not yet fully explored discourse-pragmatic functions (e.g. quotatives > (a) evidential; >> (b) complementizers; >>> (c) narrative devices). The form in question in Gorum is *sun-ru* and it has in various examples the function of introducing actual speech, serves to introduce the content of any verb of mental action, serves to hedge on the information source of a narrative chunk, re-focus on particular referents in the discourse temporarily, serves as a kind of general topic marker, and serves to mark a range of complement or subordinate clauses. These descriptions of the functions of *sun-ru* are rather impressionistic. A full survey of *sun-ru*, however, has not been made, not to mention a thorough analysis. A few examples of the diverse functions of *sun-ru* are given below. Large numbers of clauses strung together with *sun-ru* are characteristic of the texts of Aze and Aze (1973) and Gorum discourse, in general, with the element performing different functions in the same sentence as well.

- (130) (a) *kati bam den ari miam sun-ru kumda kol-di*
 knife hit COND/AFF that blood say-PST pumpkin leaf-FOC
ba? sun-ru bam-u
 place say-PST hit-TR:AFF
 ‘When she was struck by the knife, her blood went on the pumpkin leaves.’ (Aze and Aze 1973:318)
- (b) *guni sun-ru karu-di sun-ru aliñ da²j duk-u*
 girl say-PST young.girl-FOC say-PST in[side] climb AUX-AFF
 ‘The girl, the young girl, was climbing up the tree.’ (Aze and Aze 1973:327)
- (c) *lag-ey-ey duk-ey one? otur*
 repeatedly.say-LV-3PL:AFF AUX-3PL:AFF over.there from
sun-ru ta asu? adu²d-di sun-ru ta qunu-di
 say-PST EMPH small boy-FOC say-PST EMPH bow-FOC
kand-di tsel-di kepa-di nali-di dasta-di
 arrow-FOC spear-FOC spear-FOC gun-FOC what.all-FOC
dar-r-ay sun-ru ua²j-ay
 grasp-PSTST-CLOC say-PST come(.AFF)-CLOC
 ‘They said that repeatedly; then from over there the small boy came holding a bow and arrow, spears and a gun and what-all.’ (Aze and Aze 1973:327)
- (d) *t²b den ta²j den sun-ru no²d-gi sun-ru denqin*
 put COND AUX COND say-PST s/he-PL say-PST cook-ECHO
da-r-ey den-ey sun-ru bal-ey sun-ru bobo
 do-PST-3PL:AFF COOK-3PL.SUBJ say-PST heat-LV say-PST dear
duk-ey sun-ru katar bafe gitor sonde sun-ru
 AUX-3PL:AFF say-PST story way song join say-PST
duk-ey sun-ru
 AUX-3PL:AFF say-PST
 ‘when I put it there, they did the cooking and heating. Dear, they stayed and made up songs.’ (Aze and Aze 1973:333)
- (e) *bileñ q² k²j le-i-tu sun-r-ay sun-ru*
 we there:AFF place 1PL-GO-NPST:AFF say-PST-CLOC say-PST
 ‘We will go there, they said.’ (Aze and Aze 1973:326)

- (f) *a²jnu* *ɖa* *ɖen* *sun-ru* *a²jnu* *ɖa-ru* *sun* *zi?*
 like.this do COND say-PST like.this do-PST when after
tiŋ-u *sun* *zi?* *sun-ru* *uri?-n* *ui-tu* *sun-ru*
 shoot-PST when after say-PST other-LOC go-NPST:AFF say-PST
amon-ɖi
 arrow-FOC
 ‘He did like this; when he did like this, it went somewhere else, the
 arrow.’ (Aze and Aze 1973:328)

5 SEMANTICS/DISCOURSE

5.1 Semantics

Lexical semantics have not been investigated in Gorum yet. In most cases only rough English equivalents are known. In some cases we also have the Desia Oriya equivalent, but lack precise lexical semantic knowledge of the element. The grammatical semantics are discussed throughout sections 3 and 4, but as mentioned in these passages their semantics are often also insufficiently understood. The major problem here, as with many other endangered languages, is that with the decreasing number of fully competent speakers the semantics of words and constructions are increasingly assimilated to the semantics of the presumptive equivalents in the superceding language, in this case Desia Oriya.

5.2 Discourse

The topic of Gorum discourse is far too great, for anything substantial here to be stated, and our knowledge of Gorum discourse is limited to a very small set of text genres comprising different types of narratives. However, a few brief, if somewhat, random comments are made below. In narratives, verbal repetition is not atypical to indicate extended action relevant to the discourse.

- (131) *ũ* *gi?-ey* *gi?-ey* *gi?-ey* *biti-r-ey* *ui-ey* *sun-ru*
 yes see-3PL.SUBJ COPY COPY tire-PST-3PL:AFF AUX-3PL:AFF say-PST
 ‘yes, they looked and looked until they became tired.’ (Aze and Aze 1973:316)

Little is really known yet about the information structure of Gorum sentences as these have not been systematically investigated to date. What Aze calls the *-ɖi* ‘focus’ element does appear to mark definiteness or specificity (see section 3.1.4), maybe with some sort of contrastiveness, which may be implied by the label ‘focus’. It may occur more than once in a sentence and mark different referents.

- (132) (a) *dinek* *kuntur-ɖi* *kinte* *zum-u* *ɖa²d* *ta?-r-ay*
 one.day rat-FOC grass eat-INF for come.out-PST-CLOC
 ‘one day the shrew came out to eat some grass.’
 (b) *ɔna?-ɖi* *amɔn-ɖi* *nɔli-ɖi* *dɔn-ay*
 bow-FOC arrow-FOC gun-FOC bring-CLOC/IMP
 ‘bring me a bow and arrow, and a gun.’ (Aze and Aze 1973:324)

A common device used in Gorum narrative discourse is a so-called head-to-tail linkage. This consists of rote, mechanical repetition of the finite verb of a preceding sentence (often at the right edge periphery of the clause) as a non-finite form

at the left edge periphery of the next sentence. As aforementioned, in Gorum, one typically finds *ɖu* if the subject of the sentence is the same as the form in the head-to-tail linkage construction and *nen/ɖen* if it is not. Compare the use of *nen* and *ɖu* in the following pair of successive sentences:

- (133) (a) *taʔ nen bɔʔj kua kuntur-ɖi etur ɖub-u*
 come.out DS/COND one crow rat-FOC DO peck-PST
dɔn-ru lɔm-u dɔn-ru
 AUX-PST bite-PST take-PST
 ‘it came out, and a certain crow pecked the shrew, bit the shrew and took it away.’
 (b) *dɔn-ru ɖu araʔ aliŋ ab-kɔkɔ-ru*
 take-PST and/SS tree on CAUS-sit-PST
 ‘having taken it, the crow put it in a tree.’ (Aze and Aze 1973:270)

Similar alternations are found in Gtaʔ texts as well (Anderson this volume). Other discourse-sensitive elements of a currently poorly understood nature in Gorum include *laʔki*, exclamations like *ale* and *si* and possibly more.

- (134) *ũ ale dantɔn-nɔ lɔki dantɔn-nɔ bileŋ*
 yes DISC clean.teeth-IMP:AFF DISC clean.teeth-IMP:AFF we
tay-i-leŋ-ay ɖu sun-r-ay sun-ru
 give-EPEN-1PL.OBJ-CLOC and say-CLOC-PST say-PST
 ‘“Yes, clean your teeth, then give the food to us,” he said.’ (Aze and Aze 1973:322)

Gorum is a nearly rigidly verb final language, there is however a post-verbal field in which elements can be placed. Mostly continuous topics are placed there; they also often have the character of what have been called after-thought topics.

- (135) (a) *Kɔria laʔ-nɔ ɖu ɖaʔj; ale taŋk gaʔ*
 waistcloth hit-IMP:AFF and climb/IMP DISC cooked.rice eat/IMP
miŋ n-ariʔ ana taŋk gaʔ-u
 I 1SG.SUBJ-refuse elder.brother cooked.rice eat-INF
 ‘“Tuck your waistcloth round you and then climb. Eat the rice first.”
 “I refuse to eat the rice, elder brother.”’ (Aze and Aze 1973:326)
 (b) *miŋ baʔ sɔbu bele lɔk baʔj=gi=ni saibu*
 I place all time people come=3PL.SUBJ=PROG westerners
ɖigin baʔj=gi=ni seday ɖigin baʔj=gi=ni
 COLL come=3PL.SUBJ=PROG constable COLL come=3PL.SUBJ=PROG
miŋ baʔ
 I PLACE
 ‘People always come to me, Westerners come to me, police men come to me.’

As all Gorum speakers are bilingual in Desia Oriya, code-mixed utterances are frequently attested in connected speech, as in the following Gorum–Desia sentence.

- (136) *tɔr maki ne-i mɔ-i-tu ki iŋkɔʔ*
 (D)you (D)your.mother (D)take-CV 2SG.SUBJ-GO-NPST:AFF Q NEG
 ‘taking your mother (=Desia), will you go or not?’ (Aze and Aze 1973:324)

6 LEXICON

6.1 Austroasiatic/Munda components

Numerous Gorum lexical elements can be traced back not only to Proto-South Munda or Proto-Munda origins, but also to old Austroasiatic forms. This is the case for many body part terms as well as for the (now mainly lost) lower numerals and other basic vocabulary, for example, *bɔʔj* ‘one’, *siʔ* ‘hand’, or *qaʔ* ‘water’.

6.2 Loan strata

In the current state of endangerment, Gorum discourse is full of Indo-Aryan words. Depending on the individual speaker, the topic of the discourse and other factors, the percentage of Indo-Aryan words can be extremely high so that the line between borrowing and code switching is sometimes hard to draw or may not even be relevant here. Nevertheless, different loan strata can be distinguished.

The vast majority of loanwords come from Desia Oriya, the lingua franca of the whole region and the first or second language of every Gorum speaker. Many of the loans from Desia seem to be rather recent, since Gorum possesses alternative words of Munda origin for them. Loans from Desia are not restricted to any lexical field or lexical category. With the exception of verbs, which get the suffix *-ey* attached to their stem, the loanwords remain morphologically and phonologically unchanged.¹⁴

There are also words from other languages in Gorum, mostly from Telugu, Oriya, and English. The loans from Telugu like *dabu* ‘money’ and from other languages entered Gorum via Desia Oriya. As far as we know there are no English or Telugu loans, which are not also present in Desia Oriya.

There is an older loan stratum with words presumable mostly of Indo-Aryan origin. Some of them have no direct correspondent in Desia, so they probably come from another source. One word of this group is *mersa* ‘chili’. For words of this group there exists no alternative Gorum word and speakers regard them as native Gorum. There are also some verbs, the most frequent being *dar* ‘to grasp, to hold’, which are loaned from Desia but do not receive the loan verb suffix *-ey*, maybe because they are also older loans.

There may be some older loanwords of Dravidian origin, but Dravidian–Munda contact has not been studied thoroughly. A preliminary attempt at certain possible avenues may be seen in Zide (1991) and Anderson (2003).

In discourse, Indo-Aryan loans sometimes constitute half or more of the lexemes of a sentence, like in the following example:

- (137) *emti ar-qa-ey-gi sun-ru kemti ki ate pani*
 like.that NEG-DO-3PL.SUBJ=3PL.SUBJ say-PST how Q hand water
gɔʔe pani neʔ sun-ru ate taŋk gɔʔe taŋk neʔ
 foot water it say-PST hand cooked.rice leg cooked.rice it
qa-t-ey duk-t-ey
 DO-PST-3PL.SUBJ AUX-PST-3PL.SUBJ/AFF
 ‘she waited on them hand and foot, serving water and rice.’ (Aze and Aze 1973:315)

7 BRIEF ANALYSED TEXTS

Text 1: The Shrew that became a Tiger from (Aze and Aze 1973: 270–274) (in normalized transcription)

- (i) *biel bɔʔj luʔg ɖuku-ry*
field one hole be-PST.AFF
‘there was a hole in a field.’
- (ii) *luʔg mɔ-gulɔm-u*
hole 2-know-PST
‘a hole, you know?’
- (iii) *luʔg aluŋ bɔʔj tsunʃia kuntur ɖuku-ry*
hole inside one shrew rat be-PST.AFF
‘inside the hole there was a shrew.’
- (iv) *dinek kuntur-ɖi kinte zum-u ɖaʔd taʔ-r-ay*
one.day rat-FOC grass eat-INF for come.out-PST-CLOC
‘one day the shrew came out to eat some grass.’
- (v) *taʔ nen bɔʔj kua kuntur-ɖi etur ɖuʔb-u ɖɔn-ru;*
come.out COND one crow rat-FOC DA peck-PST AUX-PST
lɔm-u ɖɔn-ru
bite-PST take-PST
‘when it came out, a certain crow pecked the shrew, bit the shrew and took it away.’
- (vi) *ɖɔn-ru ɖu araʔ aliŋ ab-kɔkɔ-ru*
take-PST and tree inside CAUS-sit-PST
‘having taken it, the crow sat in a tree.’
- (vii) *zum-t-ay sun-ru ɖu miŋeʔj kua-ɖi miŋeʔj ɖu*
eat-NPST-CLOC say-PST and happy.AFF crow-FOC happy.AFF and
kua-ɖi besi miŋeʔj ɖu kaakaa sunru
crow-FOC very happy.AFF and ‘Kaa-Kaa’ say-PST
amtɔm-y taʔj tsunʃia kuntur-ɖi tɔʔb-ɖɔy baʔ
open.mouth-(PST).AFF AUX shrew rat-FOC mouth-3.POSS place
ɔʔtur suŋ-y lɔbɔʔ-n
from fall-(PST).AFF ground-LOC
‘“I must eat you” he said and the crow was very happy; he was happy and said “Kaa-Kaa” and as he opened his mouth, the shrew fell from his mouth to the ground.’
- (viii) *suŋ-y ɖaʔd aʔi tsunʃia kuntur-ɖi bɔʔj rusi tɔpɔsia*
fall-(PST)-AFF because that shrew rat-FOC one priest worship
ɖa-ru ɔlku
do-PST AUX
‘that shrew fell while a certain priest was doing his worship.’
- (ix) *butɔŋ zibɔn ɖa-ry ɖu aʔi rusi-ɖi-nu ziʔg aki*
fear life do-PST.AFF and that priest-FOC-GEN leg below

- yi du sɔɔn qa-ru*
 go:AFF and worship do-PST
 ‘the shrew feared for his life and went to the priest’s feet and did worshipped.’
- (x) *rusi-qi gi? du tsunʃia kuntur-qi etur leʔd-ru qɔn-ru du*
 priest-FOC see and shrew rat-FOC OBJ catch-PST AUX-PST and
e-nɔʔd zɔɔn qa-ru
 OBJ-s/he guardian do-PST
 ‘the priest saw the shrew, caught and took him and looked after him.’
- (xi) *zɔɔn qa-ru du rusi-qi babɛy ɔ kuntur e-nɔʔd*
 guardian do-PST and priest-FOC think.AFF that rat OBJ-s/he
etur kua-nen zum-tu laʔ-tu
 OBJ crow-this eat-NPST AUX-NPST
 ‘having taken care of him, the priest thought “that shrew, this crow will eat him up, for sure.”’
- (xii) *e-nɔʔd bɔʔj rumaŋ qa-t-ay taʔj-t-ay sun-ru du*
 OBJ-s/he one cat do-NPST-CLOC AUX-NPST-CLOC say-PST and
kuntur-qi etur rumaŋ ɣyir taʔj
 rat-FOC OBJ cat change-PST AUX
 ‘he must become a cat; then he created a cat from the shrew.’
- (xiii) *rumaŋ ɣyir nen taʔj nen dinek aʃi rumaŋ-qi etur*
 cat change COND AUX COND one.day that cat-FOC OBJ
kusɔʔd san-ru qɔn-ru leʔd-t-ay zum-t-ay sun-ru du
 dog chase-PST AUX-PST catch-NPST-CLOC eat-NPST-TLOC say-PST and
 ‘when he created the cat, one day a dog chased and took the cat and said “I must catch and eat you.”’
- (xiv) *kusɔʔd san-u-ta gi? du bɔyragi-qi rumaŋ-qi etur kusɔʔd*
 dog catch-LOC-EMPH see and priest-FOC cat-FOC OBJ dog
aʔyir-u taʔj
 change-PST AUX
 ‘the priest saw the dog, the hunting one, and created a dog from the cat.’
- (xv) *kusɔʔd ɣyir-u taʔj du kusɔʔd-qi etur bar-bɔʔj qruka*
 dog change-PST AUX and dog-FOC OBJ again-one tiger
bɔʔj leʔn-u zum-u qaʔd aram qa-ru
 one catch-INF eat-INF for idea do-PST
 ‘he created a dog and a tiger had the idea to catch and eat the dog.’
- (xvi) *ent zi? bɔyragi-qi gi? du kusɔʔd-qi etur qruka*
 that also priest-FOC see and dog-FOC OBJ tiger
ɔb-il-u taʔj
 CAUS.AFF-create-PST AUX
 ‘the priest also saw that and caused the dog to become a tiger.’
- (xvii) *kusɔʔd-qi qruka il-u du bɔyragi-qi-nu asrɔm aʔsuŋ*
 dog-FOC tiger create-NF:AFF and priest-FOC-GEN ashram house

baʔ uʃi kurɔŋ-ke pɔi pɔi du ɡɔrɔn ɔa-ru ɔa-ru
 place four way-each turn.AFF REDFL and roar do-PST do-PST
ɔuku-ry ɔaʔn bɔyragi-ɔi-nu sisu rusi zel-nu
 AUX-PST.AFF because priest-FOC-GEN student priest tell-n:AFF
sun-nuʔ lɔk aʃi bɔyragi-ɔi-nu kuʃia mɔy ar-u-i-n-ey
 say-n:AFF folk that priest-FOC-GEN hut who NEG-go-n-PL.AFF
butɔŋ-ey
 fear-PL.AFF

‘the dog, having become a tiger, walked around the four sides of the ashram building and was roaring and roaring; the priest and his students talked to each other “who will go to the priest’s hut?” – they are afraid.’

(xviii) *bɔyragi-ɔi dinek zel-u*
 priest-FOC one.day tell-PST
 ‘one day the priest told them’

(xix) *ɔruka-nen bɔʔptɔŋ-t-ay du sɔbu lɔk miŋ lɔge mɔy*
 tiger-this see-CLOC and all folk I near who
ar-u-ay-gi ɡialpɔn sɔbu-ke bɔyragi-ɔi luʔ-ru
 NEG-come-PL blighter all-each fear.CAUS-PST AUX-PST
 ‘when everyone see this tiger, who will come near me? The blighter is frightening them all.’

(xx) *sun-ru du aʃi lɔk ɔigin etur zel-u ɔruka-nen sɔt*
 say-PST and that people COLL OBJ tell-PST tiger-this real
ɔruka ɔr-ɔa-nu bɔʔj tsunʃia kuntut enu
 tiger NEG-do-n:AFF one shrew rat that
 ‘so he told those folk “the tiger is not a real tiger, it is a shrew.”’

(xxi) *miŋ mɔntrɔ ne-ɔa-ru du ɔruka ɔb-il-u ne-ku*
 I spell 1-do-PST and tiger CAUS.AFF-create-PST 1-AUX:AFF
bɔ-r-butɔŋ sun-ru du bɔyragi-ɔi aʃi lɔk ɔigin etur zel-u
 2PL-NEG-fear say-PST and priest-FOC that folk CLFR OBJ tell-PST
 ‘I made a spell and I caused this tiger to appear; don’t be afraid.’

(xxii) *laʔki ɔruka-ɔi babey ɡialpɔn miŋ-tɔ bɔʔj bɔn zɔntu*
 DISC tiger-FOC think.AFF blighter I-EMPH one wild animal
ɔruka e-niŋ etur amnaʔj mɔy bɔ-r-butɔŋ-nuʔ sun-ru zel-u
 tiger OBJ-I OBJ how who 2PL-NEG-fear-n:AFF say-PST tell-PST
 ‘Then the tiger thought “Blighter! I myself am a wild animal!” “How is it that you should not be afraid of me?” he said.’

(xxiii) *miŋ bɔʔptɔŋ-t-ay bɔyragi-ɔi etur sun-ru ɔruka-ɔi dinek ɡɔrɔn*
 I frighten-NPST-CLOC priest-FOC OBJ say-PST tiger-FOC one.day roar
ɔa-ru ɔa-ru du bɔyragi-ɔi etur leʔn-u zum-u ɔaʔd u
 do-PST do-PST and priest-FOC OBJ catch-INF eat-INF for go:AFF
 ‘“I will frighten you” said the tiger to the priest; one day he roared and went to catch and eat the priest.’

(xxiv) *guṭguṭ-ey* *ḍu le²d-u* *zum-u ḍa²d yi* *ḍa²d arre*
 REDPL.FOAR-LV.AFF and catch-INF eat-INF for go:AFF because ‘wow’
tsunṭia gialpən etur miṅ məntrə bəle *kula ne-ḍa-ru ne-ku-ry*
 shrew blighter OBJ I spell due.to tiger 1-DO-PST 1-AUX-PST.AFF
mita²n-nen e-niṅ etur le²d-u *zum-u əbay luṛ-ay*
 today-this OBJ-I OBJ catch-INF eat-INF come.INF AUX-CLOC
 ‘he roared and went to catch and eat him; then the priest said “Wow! due to my spell I made that blighter of a shrew into a tiger; this day he is coming to catch and eat me.”’

(xxv) *amna²j aṛmaṅ tsunṭia mə-ku-ry* *ayte?* *tsunṭia ḍa-na*
 how first shrew 2-be-PST.AFF like.that shrew DO-IMP.AFF
yə sun-ru ḍu bar bər ta²j la²-ru
 go.IMP say-PST and again spell give AUX-PST
 ‘“become like the shrew you were previously” he said and again he cast a spell.’

(xxvi) *bər ta²j-u ḍa²n bar ḍruka-ḍi zemti aṛmaṅ-nu tsunṭia*
 spell give-PST because again tiger-FOC how first-MODFR shrew
rup-u zemti rup ḍa-ry yi
 change-INF how change DO-PST.AFF AUX:AFF
 ‘because he cast a spell again, the tiger was changed back into the shrew he was at first.’

(xxvii) *rup ḍa-ry yi ḍa²d ruməṅ yi ḍu aṛi*
 change DO-PST.AFF AUX:AFF because cat go:AFF that
tsunṭia kuntur-ḍi etur le²n-ru zum-u la²-ru
 shrew rat-FOC OBJ catch-PST eat-PST AUX-PST
 ‘because he was changed, a cat caught that shrew and ate him up.’

(xxviii) *sun-ry*
 say-PST.AFF
 ‘The End’

Text 2: Bedbugs (from Aze and Aze 1973:296–299)

(i) *bileṅ gərum ḍigin ba? kət inku?*
 we Gorum COLL place cot not
 ‘there are no beds in our Gorum houses.’

(ii) *ayne? sila²d ḍuk-tu*
 like.this mat be-NPST.AFF
 ‘there is a sleeping mat.’

(iii) *bileṅ gərum ḍigin ba? kət inkə?*
 we Gorum COLL place cot not
 ‘there are no beds in Gorum places.’

(iv) *baiṅ sərkar lək ba? kət ḍuku-tu*
 you.PL business folk place cot be-NPST.AFF
 ‘in your houses there are beds.’

- (v) *bileŋ gɔrum ba? sila²d*
 we Gorum place mat
 ‘in our Gorum places there are mats.’
- (vi) *ayne? sila²d ba? zi? sima²d dɔk-t-ey*
 like.this mat place also bedbug be-NPST-PL.AFF
 ‘Bedbugs stay in mats likes these.’
- (vii) *tin din-ke tsari din-ke gɔdɑ? nen dɔn nen ne?*
 three day-each four day-each wash COND take COND it
kɔndɛk una-t-ey
 little decrease-NPST-PL.AFF
 ‘if we take them and wash them every three or four days then there are fewer bedbugs.’
- (viii) *ink? sun enen-gi enen-gi ura dɔgin sen sirim*
 not when this-PL this-PL body COLL this.side pinch
sun-ta-y lɔm-t-ay-gi sen sirim ananananay
 say-NPST-CLOC bite-NPST-CLOC-PL this.side pinch elder.brother
sun-u ɔyɔyɔyɔyɔy
 say-INF Oy...
 ‘if not, then they scratch their sides as the bedbugs bite their bodies;
 “brother, brother” one says “Oyoyoyoy.”’
- (ix) *kine? laka bubɔŋ dɔgin sunen ayne? inku?*
 this.tall item baby COLL when/if like.this not
zum-t-ey la²-t-ey
 eat-NPST-PL AUX-NPST-PL
 ‘when the babies are this big, it is not like this, for the bedbugs will eat them.’
- (x) *rata rata dɑ-tu ui-tu*
 red.spot RDPL do-NPST-AFF AUX-NPST-AFF
 ‘red spots will appear.’
- (xi) *inɔi kulan-gi ba²-nu guni dɔgin kulan-gi ba²-nu guni*
 this kulan-PL place-GEN girl COLL kulan-PL place-GEN girl
dɔgin sunen enten-gi
 COLL when/if that-PL
 ‘this happens to the girls at this place, Kulans’ place, those girls.’
- (xii) *ura-iri rata-tu ui-tu*
 body-ECHO red.spot-NPST-AFF AUX-AFF
 ‘their bodies are covered in red spots.’
- (xiii) *a²jtɔ lɔm-t-ey sima²d*
 like.that bite-NPST-PL bedbugs
 ‘bedbugs will bite like that.’
- (xiv) *aa*
 yes
 ‘yes’

- (xv) *e-nɔʔd aka lɔm-u luʔ-r-ey ɖu nɔʔd luʔp lɔk ɖu*
 OBJ-S/he EMPH bite-INF AUX-PST-PL and s/he big folk and
man-ey ar-luʔd gɔʔtuŋ gusaʔj ɖuku aʔi
 sense-LV NEG-AUX cloth wrap.around AUX:AFF that
 ‘when they are biting him, he doesn’t feel it as he is an adult and is wrapped in his blanket.’
- (xvi) *idiʔa guni ɖigin ɖaʔd lɔm-ey sun aaaa sun-t-ey*
 small girl COLL because bite-PL when aaaa say-NPST-PL
mɔmɔɔɔ mɔmɔmɔmɔɔmɔ sun-t-ey
 suckling.noise say-NPST-PL
 ‘when they bite little girls, the say “aa give me a suck, give me a suck” (of mother’s breast).’
- (xvii) *aʔjtu-ʔa simaʔd*
 like.that-EMPH bedbugs
 ‘bedbugs are like that.’
- (xviii) *simaʔd-nen bɔʔj-bɔʔj aʔsuŋ pampu laʔ-ey taʔj-ey luʔr-ey pampu*
 bedbug-this one-one house pump hit-PL AUX-PL AUX-PL pump
ɖigin inɖi ayneʔ kabri kabri gi-n-ay luʔr-ay pampu
 COLL this like.this spotted COPY see-n-CLOC:AFF AUX-CLOC pump
ɖigin ɖɔn-r-ay(-gi) laʔ-r-ey taʔj-ey sun ta abuʔd
 COLL take-PST-CLOC-PL AUX-PST-PL AUX-PL when EMPH much
ɖa-t-ey simaʔd
 do-NPST-PL.AFF bedbug
 ‘in each house they are spraying these bedbugs and the wall looks like this, all spotted; when they bring pump to spray, there will be many bedbugs.’
- (xix) *lɔk din-ke aʔsuŋ zɔɖaʔ-u*
 folk day-each house paint.w.cowdung water-INF
 ‘every day people cowdung wash there houses.’
- (xx) *din-ke saʔa gɔbɔr saʔa ɖa-u*
 day-each cowdung.water cowdung cowdung.water do-INF
laʔ-u ɖu din-ke silaʔd goɖaʔ-ay tebe simaʔd inkɔʔ
 AUX-INF and day-each mat wash-CLOC then bedbug not
 ‘each day they do it with cowdung water (with wood chippings) and then they wash the sleeping mats, and then there are no more bedbugs.’
- (xxi) *aa*
 yes
 ‘yes’
- (xxii) *inkʔ sun munaʔd aka yeʔ-u laʔ-u ɖa-tu lɔm-ay-gi*
 not when myself EMPH cry-INF AUX-INF AUX-NPST.AFF bite-CLOC-PL
sun lɔbulɔb lɔbulɔb lɔbulɔb lɔm-ay-gi sun
 when nibble bite-CLOC-PL when
 ‘when it is not so, we ourselves cry out when they bite us, nibble-nibble.’

- (xxiii) *hayhayhayhayhayhay sun-u qa-tu?*
 ‘Ha... say-INF AUX-NPST.AFF
 ‘“Ha” we say.’
- (xxiv) *sɔka-siki kuy nen duk nen ne? gɔtuŋ=gɪtiŋ ayne? gusaʔj*
 shirt=ECHO wear COND.AFF AUX COND.AFF it cloth=ECHO like.this wrap
nen du gɔtuŋ-ɖi aʔynu bil nen du gɔn nen
 COND.AFF and cloth-FOC like.this lay.out COND and lie.down COND
taʔj nen ne? kɔndek ɔ-lɔm-ay-gi
 AUX COND it little NEG-bite-CLOC-PL
 ‘if we are wearing shirts, or if we are wrapped up in blankets like this, and then if the blanket is laid out and one lies down, the bedbugs will not bite so much.’
- (xxv) *bar ɖumɖa ura gɔn-ng-t-ay ɖɔm*
 again top.half.naked body lie.down-IMP.AFF-NPST-CLOC.AFF must
 ‘if you lie down with a bare back.’
- (xxvi) *lɔm-lɔm-t-ay laʔ-t-ay-gi*
 RDPL-bite-NPST-CLOC AUX-NPST-CLOC-PL
 ‘bedbugs will bite (you) repeatedly.’
- (xxvii) *saklya giʔ nen emti zuʔb inku?*
 morning see COND.AFF in.this.manner itch not
 ‘if you see yourself in the morning, don’t itch yourself so much.’
- (xxviii) *nay-ʔa lɔm-ij rey sun-u qa-tu*
 what-EMPH bite-1 EXCLAM say-INF AUX-NPST.AFF
 ‘“what bit me” one will say.’
- (xxix) *aa*
 yes
 ‘yes’
- (xxx) *inkuʔ; yeʔ-u laʔ-u inkuʔ sina; aʔjtu-ʔa*
 no, cry-INF AUX-INF NEG EMPH like.that-EMPH
 ‘No, one mustn’t cry. It is like that.’
- (xxxi) *aa*
 yes
 ‘yes’

NOTES

* Data with no source specified was collected by Felix Rau during his fieldwork. Most of the data was collected on a field trip financed by the Gesellschaft für Bedrohte Sprachen (German Society for Endangered Languages), which he gratefully acknowledges. Felix Rau would also like to thank the people of Bod Kicob, Orissa and especially Loikon Kirsani for patiently teaching him Gorum.

1 A. Zide (1982) lists also the allophone [j] for /z/, but it is in free variation with the others and neither Aze nor the authors have found it.

- 2 Generally Gorum distinguishes only between singular and plural, but in the imperative verb paradigm a separate first person dual form also occurs.
- 3 Note that this =*gi* is part of the verbal agreement paradigm and not number marking of the noun.
- 4 This marker is called ‘classifier’ by Aze and Aze (1973:214) because of its function ‘to denote a collective noun indicating the whole class of things’. This label is confusing and should be regarded as a misnomer.
- 5 Present informants only know of the object marking function of *etur* and use only *entur* in the ‘to, towards’ sense. Dative/Accusative conflation is common in the South Asian linguistic area however, and found in numerous other Munda languages as well, for example, Gta? (Anderson, this volume) or Korku (Zide, this volume).
- 6 Only one recent work (Mahapatra 1995) has *galgi* instead of *gulgi*.
- 7 As has been already mentioned under section 3.1.1 the relation between the nominal plural marker *-gi* and the third person plural subject marker =*gi* is an open question. Historically they seem to be related, and cognate with plural markers found in both and nominal and verbal forms across the Munda languages (Anderson 2007). Their phonology and morphosyntax differ however.
- 8 Three participant events with a first or second person as a theme-like participant and not as the recipient are avoided through various strategies, like grouping the theme-like participant together with the agent making both the subject.
- 9 This example is not acceptable to all speakers. The ambiguous status might result from the fact that normally only inalienable possession is marked by the possessor affixes and that most speakers refuse to construe the possession of a pencil as inalienable.
- 10 Note the double marking of object in this possessor raising construction in Gorum. Most auxiliary verbs in Gorum show such doubled inflection (of both ‘subjects’ and ‘objects’). Gorum is thus close to a canonical ‘doubled’ inflectional system in auxiliary verb constructions (see Anderson 2006).
- 11 These forms of the past tense suffix are treated as a zero past tense form by Aze (1973) and Aze and Aze (1973).
- 12 A similar archaism is attested in contemporary varieties of Xakas as well (Anderson 2005).
- 13 It is possible that this use of the *CLOC* reflects a Munda-type analog to the common Indian areal experiencer constructions with dative subjects.
- 14 It should be noted that the phoneme inventories of Gorum and Desia Oriya are very similar. For the phoneme inventory of Desia Oriya see Gustafsson (1989).

REFERENCES

- Anderson, G.D.S. (1995) ‘Ditransitives, possessor raising, copying-to-OBJ: Animacy in morphosyntax’, *CLS*, 31: 1–17.
- (1997) ‘On “animacy maximization” in Fox (Mesquakie)’, *International Journal of American Linguistics*, 63(2): 227–247.
- (2001a) ‘A new classification of Munda: Evidence from comparative verb morphology’, *Indian Linguistics*, 62: 27–42.
- (2001b) ‘Santali’. in J. Garry and C. Rubino (eds) *Encyclopedia of the World’s Major Languages*, New York/Dublin: H. W. Wilson Press, 623–628.
- (2001c) ‘Subject version and object version in Tofa auxiliary verb constructions’, *Turkic Languages*, 5(1): 240–269.
- (2003) ‘Dravidian influence on Munda’, *International Journal of Dravidian Linguistics*, 32(1): 27–48.

- (2004a) *Auxiliary Verb Constructions in Altai-Sayan Turkic*, Wiesbaden: Harrassowitz Verlag.
- (2004b) 'Advances in Proto-Munda reconstruction', *Mon-Khmer Studies* 34: 159–184.
- (2005) *Language Contact in South Central Siberia*. Wiesbaden: Harrassowitz Verlag.
- (2006) *Auxiliary Verb Constructions*, Oxford: Oxford University Press.
- (2007a) *The Munda Verb. Typological Perspectives*, Trends in Linguistics Studies and Monographs 174, Berlin: Mouton de Gruyter.
- (this volume, Chapter 13) 'Gta?'.
- (2007b) 'Burushaski morphology', in A. Kaye (ed.) *Morphologies of Asia and Africa*, Winona Lake: Eisenbrauns, 1233–1275.
- Anderson, G.D.S. and O.P. Gurevich (2005) 'Towards a typology of version: Formal and functional', to appear in *CLS 41*, Chicago: Chicago Linguistic Society.
- Anderson, G.D.S. and K.D. Harrison (this volume-a, Chapter 11) 'Remo (Bonda)'.
- (this volume-b, Chapter 7) 'Sora'.
- Anderson, G.D.S. and N.H. Zide (2001) 'Recent advances in the reconstruction of the Proto-Munda (Austroasiatic) verb', in L.J. Brinton (ed.) *Historical Linguistics 1999*, Amsterdam: John Benjamins, 13–30.
- (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The Bimoraic constraint', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Anderson, G.D.S., T. Osada and K.D. Harrison (this volume, Chapter 5) 'Ho and the other Kherwarian languages'.
- Aze, (F.) R. (1971) *Parengi (Gorum) Phonemic Summary*, Kathmandu: Summer Institute of Linguistics mimeograph.
- (1972) 'The status of glotalization in Parengi (Gorum)', *Proceedings of All India Conference of Linguists*, 1972, 84–95 Poona: Deccan college.
- (1973) 'Clause patterns in Parengi-Gorum', in R. Trail (ed.) *Patterns in Clause, Sentence and Discourse in Selected Languages of India and Nepal, Part 1*, Kathmandu: SIL and Tribhuvan University, 235–312.
- Aze, R. and Aze, T. (1973) 'Parengi texts', in R. Trail (ed.) *Patterns in Clause, Sentence and Discourse in Selected Languages of India and Nepal, Part 3*, Kathmandu: SIL and Tribhuvan University, 213–362.
- Bashir, E. (1985) 'Towards a semantics of the Burushaski verb', in A. Zide, D. Magier, and E. Schiller (eds) *Proceedings of the Conference on Participant Roles: South Asia and Adjacent Areas*, Bloomington: Indiana University Linguistics Club, 1–32.
- Bhat, D.N.S. (1997) 'Noun-verb distinction in Munda languages', in A. Abbi (ed.) *Languages of Tribal and Indigenous Peoples of India: The Ethnic Space*, Delhi: Motilal Banarsidass, 227–251.
- Bhattacharya, S. (1954) 'Studies in the Parengi language', *Indian Linguistics*, 14: 45–63.
- (1975) *Studies in Comparative Munda Linguistics*, Simla: Indian Institute of Advanced Study.
- Bril, I. (2004) 'Complex nuclei in Oceanic languages: Contribution to an areal typology', in I. Brill and F. Ozanne-Rivierre (eds) *Complex Predicates in Oceanic Languages*, Berlin: Mouton de Gruyter, 1–48.
- Crowley, T. (2002) *Serial Verbs in Oceanic*, Oxford: Oxford University Press.
- Diffloth, G. (1989) 'Proto-Austroasiatic creaky voice', *Mon-Khmer Studies*, XV: 139–154.
- Dryer, M. (1986) 'Primary objects, secondary objects, and antitativity', *Language*, 62: 808–845.

- Evans, N. and T. Osada (2005) 'Mundari: The myth of a language without word classes', *Linguistic Typology*, 9(3): 351–390.
- Gordon, R. G. [ed.] (2005) *Ethnologue*, 15th edition, Dallas: Summer Institute of Linguistics International.
- Grierson, G. A. [ed.] (1906) *Linguistic Survey of India. Vol IV: Munda and Dravidian*, Calcutta: Office of the Superintendent of Government Printing.
- Griffiths, A. (this volume, Chapter 12) 'Gutob'.
- Gustafsson, Uwe (1989) *An Adiwasi Oriya-Telugu-English Dictionary*, Mysore: CIIL.
- Mahapatra, K. (1995) *Parenga*, Bhubaneswar: Tribal Affairs Office.
- Mithun, M. (1984) 'The evolution of noun incorporation', *Language*, 60(4): 847–894.
- Osada, T. (1992) *A Reference Grammar of Mundari*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- Parkin, R. (1991) *A Guide to Austroasiatic Speakers and Their Languages*, (Oceanic Linguistics Special Publication 23), Honolulu: University of Hawaii Press.
- Peterson, John (2005) 'There's a grain of truth in every "myth", or, why the discussion of lexical classes in Mundari isn't quite over yet', *Linguistic Typology*, 9:391–405.
- Ramaswami, N. (1992) *Bhumij Grammar*, Mysore: Central Institute of Indian Languages.
- Rau F. (2004) 'Complex predicates in Gorum', Presented at the Munda Meeting, Leipzig.
- Senkuttuvan, R. (2002). 'Parengi', in B. Mahapatra (ed.) *Linguistic Survey of India, Special Studies*, Orissa, Calcutta: Office of the Registrar General India, 424–439.
- Sitapati, G.V. (1933) 'Parēṅ', in *A Miscellany of Papers Presented to Rao Sahib Mahopadhyaya Gidugu Venkata Ramamurti Guru*. Madras Government office Press, 145–165.
- Smith, K.D. (1975) 'The velar-animal prefix relic in Vietnam languages', *Linguistics in the Tibeto-Burman Area*, 2(1): 1–18.
- Solntseva, V. N. (1996) 'Case-marked pronouns in the Tao-ih language', *Mon-Khmer Studies*, XXVI: 33–36.
- Subba Rao, V. and D.R. Patnaik (1992) *Gadaba. The Language and the People*, Amaravathi: Papayaradhya Sahiti Kendram.
- Thurston Edgar (assisted by K. Rangachari) (1909), *Castes and Tribes of Southern India*, Madras Government Press office, 7 volumes.
- Zide, A.R.K. (1963) *Parengi phonology*, MA Thesis, University of Pennsylvania.
- (1966) 'Some problems in the generation of Gorum obstruents and related glides', Unpublished manuscript.
- (1972) 'Transitive and causative in Gorum', *Journal of Linguistics*, 8(2): 201–216.
- (1976) 'Nominal combining forms in Sora and Gorum', in P. Jenner, L. Thompson and S. Starosta (eds) *Austroasiatic Studies. Part II*, (Oceanic Linguistics Special Publication no. 13), Honolulu: University of Hawaii Press, 1259–1293.
- (1982) *A Reconstruction of Sora-Gorum Morphology*, PhD Dissertation, University of Chicago.
- (1997) 'Pronominal and nominal incorporation in Gorum', in A. Abbi (ed.) *Languages of Tribal and Indigenous Peoples of India*, Delhi: Motilal Banarsidass, 253–261.
- Zide, N.H. (1976) 'A note on Gta? echo forms', in P. Jenner, L. Thompson, and S. Starosta (eds) *Austroasiatic Studies. Part I & II*, (Oceanic Linguistics Special Publication no. 13), Honolulu: University of Hawaii Press, 1335–1343.

-
- (1978) *Studies in the Munda Numerals*, Mysore: Central Institute of Indian Languages.
- (1991) 'Possible Dravidian sources of some Munda demonstrative bases', *Studies in Dravidian and General Linguistics* (Krishnamurti Festschrift), Hyderabad, 349–364.
- (this volume, Chapter 6) 'Korku'.
- Zide N.H. and Stampe, D.L. (1968) 'The position of Kharia-Juang in the Munda family', in B. Krishnamurti (ed.) *Studies in Indian Linguistics* (Professor M.B. Emeneau's *śaṣṭipurti* volume), Poona: Centres of Advanced Study in Linguistics, Deccan College, 370–377.

KHARIA

John Peterson

1 INTRODUCTION

Kharia is a South Munda language spoken primarily in the southwestern areas of the state of Jharkhand (primarily in Districts Simdega and Gumla)¹ and adjoining areas of the states of Chhattisgarh (primarily in Districts Surguja and Raigarh) and Orissa (especially in Sundargarh District). According to the internet version of the *Ethnologue*,² there is also a considerable number of speakers in Assam and smaller communities elsewhere, including West Bengal and the Andaman and Nicobar Islands.

In the genetic scheme of Munda given in Zide (1969), *Kharia* and Juang make up the South Central branch of Munda. In Anderson's (2001) scheme, however, Proto-Juang and Proto-Kharia parted company at an early date, although Kharia and Juang are noted to share many common traits. As Anderson (2001:32) notes, the question of internal relationships within South Munda is still very much an unsettled issue.

Kharia is often spoken in multi-lingual communities, where its speakers are in daily contact with speakers of Sadani,³ the traditional *lingua franca* of the region, and Hindi (both Indo-Aryan), Mundari (Munda) and Kurukh⁴ (North Dravidian). In Orissa speakers of Kharia are also in close contact with speakers of Oriya (Indo-Aryan). Most speakers of Kharia are multi-lingual and speak Kharia, Sadani and Hindi more or less fluently, and in Orissa also Oriya.

The standard name of the language and people is *kharīya* [*k^hə.ɽijá*], but many speakers feel that the form *kheṛiya* [*k^hɛ.ɽijá*] is more correct. The etymology of *kharīya* is unclear. Konow (Grierson 1906 [1994]:190) notes a suggestion which connects the term to the Proto-Munda word for 'person'. Although this is uncontroversial in the case of the North Munda languages and peoples, who refer to themselves and their languages as either *hɔɽ* or *ho* (both of which mean 'man, person' in the respective languages), the cognate word in Kharia is *kaɽ*, and there remain a number of problems with such a derivation, especially the aspiration in the name *kharīyal kheṛiya*. A number of other popular or 'folk' etymologies have also been proposed (see Kullū 2000:3–7 for an overview); perhaps the most interesting is the one in which the name is claimed to derive from the Kharia word *khori* 'village section', as the Kharia often live together in individual sections of multi-lingual villages.

There are three groups which are generally classified together as Kharia – the *dudh* Kharia (cf. Hindi *dūdh*, Sadani *dudh* 'milk'), the *qelki* or *qhelki* Kharia, and the Hill Kharia. Of these three groups, apparently only the Dudh and D(h)elki Kharia speak Kharia, whereas the Hill Kharia now speak Indo-Aryan languages closely related to Bengali and Oriya.

It is generally assumed that there are two main dialects of Kharia corresponding to the division between the Dudh and D(h)elki Kharia, but this has yet to be confirmed.

The present study deals only with Dudh Kharia as it is spoken in southwest Jharkhand and northwest Orissa.

According to the 1971 census, 191,421 people spoke Kharia as their native language at that time, out of a total Kharia population of 321,190 (data from Abbi 1993:543), whereas Grimes (1988:471), quoting another work, gives 111,000–160,000 speakers for roughly the same period. The present Kharia-speaking population has undoubtedly increased since then, although it is not clear how many speakers there are. The internet versions of the *Ethnologue*⁵ give a total of 278,500 speakers for 1994 (14th edition 2000) and 292,000 in India for 1997 (15th edition 2005).⁶

The question of whether Kharia is endangered is difficult to answer. In the short term, it cannot be considered endangered, although the long-term chances of survival are more difficult to assess. As early as Grierson (1906 [1994]:190) Kharia was declared to be a dying language, a view which also found support in Pinnow (1965a:ix, 4ff.). Nevertheless, the language probably has more speakers now than at any other point in its history due to the size of the ethnic Kharia population, although it is my impression that an ever-increasing percentage of this population is choosing to rear their children in either Hindi or Sadani.

Conversely, there is a growing movement to promote Kharia, both through various organizations as well as in the schools. On the other hand, the increasingly high level of education is bringing the Kharia in their daily lives more and more into contact with Hindi and other languages, while opportunities for speaking Kharia are becoming fewer. Only time will tell to what extent present attempts at preserving the language are successful. Although the language has certainly taken on a large number of traits from Indo-Aryan (cf. Abbi 1993, 1997 and Malhotra 1982), it has survived up to the present and the number of people speaking it has even increased substantially in the past 30 years, despite all earlier forecasts.

2 PHONOLOGY

Ironically, although modern linguistic work on Kharia began with Pinnow's (1959) monograph on Kharia phonology, predominantly from a historical viewpoint, work on Kharia phonology has progressed little since then. As work on this area is currently in progress, the information provided in the following pages is still somewhat tentative. The analysis here is based largely on Rehberg (2003).

2.1 Vowel inventory

Monophthongs. Kharia has the following five distinctive (monophthongal) vowel phonemes, given here along with their most common allophones:

- | | | |
|-----|-------------|----------------|
| (1) | i [i, i(:)] | u [u, u(:)] |
| | e [ɛ, e(:)] | o [ɔ, o(:)] |
| | | a [ɑ, a, ə, ʌ] |

In (2), I give a few minimal pairs to demonstrate the status of these vowels as phonemes.

The status of [ə] is unclear: For Biligiri (1965:16f.) it can be phonemic but is generally an allophone of /a/, while Rehberg (2003:7) considers it an allophone of /e/, /a/, /o/ or /u/. I consider it an allophone of /a/. [ʌ] occurs only in loanwords.

Vowel length is not phonemic in Kharia. There is a strong tendency for word-final vowels to be lengthened. Otherwise, it is not yet possible to give precise rules for the distribution of these allophones, although the first allophone given in (1) for each phoneme may be considered the unmarked form.

Biligiri (1965:18) and Pinnow (1959:29) also consider nasalization (marginally) phonemic. I consider it non-phonemic, as I know of no genuine minimal pairs which are distinguished by nasalization alone.

- (2) /a/ vs. /i/: *anaj* ‘we (DL INCL)’ / *aniŋ* ‘we (PL INCL)’
 /a/ vs. /o/: *jaʔ* ‘pull out of the ground’ / *joʔ* ‘sweep’
 /i/ vs. /e/: *ɲim* ‘thatch’ / *ɲem* ‘warm’, *dɲɪn* ‘draw, pull’ / *dɲɛn* ‘cook (pulse, vegetables)’
 /u/ vs. /a/: *uluʔ* ‘boil’ / *ulaʔ* ‘leaf’
 /o/ vs. /e/: *koloŋ* ‘bread’ / *koleŋ* ‘king’, *goʔj* ‘die’ / *geʔj* ‘pluck (fruit)’
 /e/ vs. /a/: *=te* ‘present active’ / *=ta* ‘present, middle’, *ter* ‘give’ / *tar* ‘kill’
 /o/ vs. /u/: *koyo* ‘wind’ / *kuyu* ‘pot’, *moʔ* ‘smoke’ / *muʔ* ‘emerge’
 /i/ vs. /u/: *gil* ‘beat’ / *gul* ‘cry out’

Diphthongs. Rehberg (2003:6) lists the following five diphthongs. The phonetic values given here are from my own ongoing work and appear to be their most common realizations:

- (3) /ae/ [aɛ, əɛ eɪ], /ao/ [aɔ], /ou/ [ɔʊ], /oi/ [ɔɛ] and /ui/ [tʃɪ]

As I follow here a transliteration of the written language, I will write these as /ay/, /aw/, /ou/, /oy/ and /ui/, respectively, unless they occur in closed syllables, where the first two are transliterated as /ai/ and /au/.

The following presents a few minimal pairs between ‘diphthongs’ and ‘monophthongs’:

- (4) /o/ vs. /ou/: *=jo* ‘additive focal particle’ / *=jou* ‘as long as’, *ho* ‘that’ / *hou* ‘yes’
 /o/ vs. /oy/: *ho* ‘that’ / *hoy* ‘become’; *koy* ‘shave’ / *=ko* ‘contrastive particle’
 /u/ vs. /ui/: *lebu* ‘man; person’ / *lebui* ‘love’
 /e/ vs. /ei/: *besɖaʔ* ‘well (adv.)’ / *beisɖaʔ* ‘the water which is poured into rice for cooking’
 /a/ vs. /ay/: *=ta* ‘present, middle marker’ / *tay* ‘ablative’; *ɖaʔ* ‘water’ / *ɖay* ‘woman’ (near-minimal)

The status of these as diphthongs is doubtful: It seems more appropriate to me to treat them simply as a vowel in the nucleus and a glide in the coda rather than as diphthongs, at least in the native vocabulary. This is primarily due to the incompatibility of these combinations with a consonant in the coda (cf. Rehberg 2003:17f.). Although such examples are found, all of these examples are either loanwords from Indo-Aryan (generally Sadani) or place names of uncertain origin: *cair* ‘four’, *j(h)ai̯t* ‘animal’, *sou²b* ‘all’, *bhoir* ‘entire’, *buidh* ‘intelligence’ and *digduin* ‘Digduin (name of a town)’.

2.2 Suprasegmental inventory

Kharia has no phonemic tones or registers. In Rehberg’s (2003:23ff.) analysis, Kharia has a non-distinctive pitch accent, marked by a low-tone on the first syllable

of the word. After this syllable, the pitch then gradually rises on the following syllable(s). This fits in well with the fact that Kharia generally does not allow monosyllabic words (see section 3.1.10).

In the present study, I do not assume a low-tone accent in Kharia. Rather, I propose that it is the ‘low → high’ prosodic pattern itself which defines a phonological word in Kharia rather than a low-tone accent. Thus, each phonological word in Kharia begins with a low-tone pitch which then gradually rises throughout the remainder of the phonological word. If the word is monosyllabic, then we find a rising contour:

- (5) *rocho²b* ‘side’ [rò.ch^hɔ²b^{ˈm}], *lay* ‘tongue’ [lǎŋ]

There is no expiratory accent in Kharia and any syllable of the phonological word may be more ‘prominent’ with respect to volume (Rehberg 2003:24ff. and 53ff. (tables)).

2.3 Consonants

Table 9.1 presents the consonant phonemes of Kharia. The status of forms in ‘()’ as phonemes is uncertain. Where transliteration and IPA values differ, the IPA value is given in ‘[]’. Aspiration is indicated in the transliteration by an ‘h’, for example, /th/ is pronounced [t^h], /bh/ as [b^h], etc.

In the following I present minimal pairs for some of these phonemes.

- (6) /g/ vs. /k/: *goŋ* ‘cook’ / *koŋ* ‘know’, *gone* ‘tooth’ / *kone* ‘mouse’
 /d/ vs. /dʰ/: *bida* ‘farewell’ / *biḍa* ‘rice seed’, *kuda* ‘millet’ / *kuḍa* ‘a kind of black berry’
 /b/ vs. /bh/: *bu?* ‘beat a drum’ / *bhu?* ‘bark’(v.)
 /t/ vs. /th/: *tomsij* ‘wasp’ / *thomsij* ‘9’
 /t/ vs. /dʰ/: *tuta* ‘bottom’ / *tuḍa* ‘tomorrow’
 /m/ vs. /ŋ/: *oŋem* ‘warm oneself’ / *oŋeŋ* ‘give back’, *ḍam* ‘arrive’ / *ḍaŋ* ‘send’
 /ɲ/ vs. /ŋ/: *eŋ* ‘open (of clouds), clear’ / *eŋ* ‘return’ (ITR)
 /m/ vs. /n/: *nom* ‘bask in the sun’ / *non* ‘drive (an ox-cart)’
 /l/ vs. /r/: *gul* ‘cry out’ / *gur* ‘fall’
 /r/ vs. /s/: *irin* ‘plum’ / *isin* ‘cook (rice)’
 /r/ vs. /dʰ/: *iri²b* ‘forget’ / *idi²b* ‘night’
 Near minimal pair: /dʰ/ vs. /d/: *ḍaŋu* ‘fool’ *daru* ‘tree’

- What are generally referred to as retroflex plosives in studies on Kharia are actually often – perhaps generally – realized as postalveolars. The exact status

TABLE 9.1: THE CONSONANTS

	Labial	Dental	Retroflex/ Postalveolar	Palatal	Velar	Glottal
Plosives	p b bh	t th d dh	ʈ ʈh ɖ ɖh	c ch j [j] jh [j ^h]	k kh g gh	(ʔ)
Nasals	m	n	(ɳ)	ɲ	ŋ	
Flaps		r [r]	(ɽ) (ɽh)			
Approximants	w	l		y [j]		
Fricatives	ph [f]	s				h [ɦ]

and air flows through the nasal cavity, producing a homorganic nasal; that is, what is written as ²*b* is pronounced [ʔb^m].

In order to differentiate between the (marginally phonemic) plosive /ʔ/ and the (non-phonemic) pre-glottalized plosives, the former will be consistently written as ʔ while the latter will be represented as ²*d*, etc. For example, what is written *poʔda* ‘village’ has the phonemic structure /poʔda/ or, perhaps, /pogda/. On the other hand, what is written *qoʔqna* ‘to carry’ has the phonemic structure /qoqna/.⁹

2.4 Syllable structure

Kharia words of native origin all have the syllable structure (C)V(C), where ‘C’ is a consonant and ‘V’ a vowel or nasal.¹⁰ In addition, there are the following general constraints:

- Words may begin with any consonant other than /t/, /tʰ/ or /ŋ/ (Rehberg 2003:15). /w/ is only found word-initially in Indo-Aryan loans. [ʔ] occasionally occurs in the onset in the spontaneous speech of some speakers (Rehberg 2003:23) but is never phonemic in this position and is in free alternation with an empty onset.
- The nucleus may consist of a vowel or nasal. /ŋ/ is attested in this function in my data (/laphŋa/ ‘cave’ [lə.pŋ.ɡa]) and Rehberg (2003:14) cites the form /phophnda/ ‘fungus’ from Biligiri (1965:11), presumably with [ŋ] in the nucleus.
- The oppositions [±voiced], [±aspirated] and dental/retroflex are neutralized in the coda (see under paragraph head ‘Pre-glottalized stops’).
- There are no native words with an /s/ or /h/ in the coda. Where these are found, the lexeme is invariably of Indo-Aryan origin, such as *bes* ‘good’, *jimis* ‘animal, thing’, *sandeh* ‘doubt’, etc.

2.5 Intonation

Little work has yet been done on Kharia intonation. Rehberg (2003:37ff.) has a few preliminary remarks on the topic, although she cautiously refers to these as speculative.

With respect to declaratives, Rehberg notes that the utterance is relatively monotonous and decreases gradually both in intensity and frequency towards the end of the sentence.

In interrogatives, on the other hand, the pitch rises and falls considerably in the utterance, and each low-tone pitch (marking the beginning of a phonological word) is clearly articulated. Sentence-pitch only falls towards the end of the utterance.

2.6 Morphophonology

Morphophonological phenomena can conveniently be divided into two groups.

- (i) In the first group, these are restricted to a particular morpheme or morphemes.

The past, active marker =oʔ. Before this marker a stem-final plosive is devoiced and aspirated, *oʔj* ‘drive’, *och=oʔ* ‘she/he drove’. Stem-final [ʔ] (< /g/) is realized as [k^h]: *noʔ* ‘eat’, *nokh=oʔ* ‘she/he ate’ (sections 3.2.3–3.2.5).

After vowels, the glide -y- is inserted before the past active marker to avoid a hiatus. For the sake of glossing I will consider the past, active morpheme in these environments to have the form =yoʔ: *thisa=yoʔ* ‘she/he called out’, *yo=yoʔ* ‘she/he saw’.¹¹

The first and second persons in the past active have slightly irregular forms: $=o^?j$ ($[\text{ɔ}^?j^{\text{m}}] < *o^?j$) and $=o^?b$ ($[\text{ɔ}^?b^{\text{m}}] < *o^?b$), respectively.

qel ‘come’ and *col* ‘go’. The final *l* in *col* ‘go’ and *qel* ‘come’ is dropped before all enclitics beginning with *n*/, for example, *co=na* ‘to go’ (INFINITIVE), *qe=na=ki* (come=M.IRR=PL) ‘they will come’ vs. *col=ki* (go=M.PST) ‘she/he went’ and *qel=ki* (come=M.PST) ‘she/he came’. This does not hold for other lexemes ending in *l*/: *ol* ‘bring’, *ol=na* ‘to bring’ (INFINITIVE).

- (ii) The second group is more phonologically conditioned, that is, the changes here are due only to the phonotactic rules of the language.

/gl/ > [g] / [ʔ]. When a stem ending in $[ʔ]$ ($< /gl/$) is followed by an enclitic beginning with a vowel (except for the past active marker $=o^?$), $[ʔ]$ may be realized as $[g]$, for example, *o?* ‘house’, *og=a?* ‘house=GEN’, *no?*=*te* (eat=A.PRS) ‘she/he eats’ vs. *nog=e* (eat=A.IRR) ‘she/he will eat’. However, this rule is not obligatory for the genitive, and speakers tend to insert the glide */y/ [j]* after the glottal stop (and other plosives) and before the genitive marker $=a?$, for example, *qa?-y-a?* ‘of the water’. This process has now spread to most environments, so that it is now probably more appropriate to consider the underlying form of the genitive marker to be $=ya?$ instead of the etymologically expected form $=a?$.

The perfect. As noted already in Biligiri (1965:59), the perfect marker $=si^?(d)=$ can be realized as $[χi.ʔi]$ after a stem-final */j/*, *lgoj=si?* \Rightarrow $[\text{gɔ}^?j.χi.ʔi]$ ‘she/he has died.’ However, this is not obligatory, and the pronunciation $[\text{gɔ}^?j.si.ʔi]$ is also acceptable.

No complex onsets or codas. The progressive markers $=te^?jq$ (active voice) and $=ta^?jq$ (middle voice) are realized as such only before the markers of the first and second persons, singular, the only two personal markers which begin with a vowel. Otherwise the final *q* is dropped: *likha=te^?jq=ijn* ‘I am writing’ vs. *likha=te^?j=ki* ‘they are writing’ ($< *likha=te^?jq=ki$). This is due to the strict (C)V(C) structure of the native elements of the language, that is, complex codas (and onsets) are not allowed. The same distribution holds for the present copulae *ayi^?j(d)* (non-negative) and *umboɽi^?j(d)* (negative).

Similarly, the perfect marker $=si^?d$ is realized as $=si^?d$ in the first and second persons, singular, elsewhere as $=si^?$: *col=si^?d=ijn* ‘I have gone’, *col=si^?* ‘she/he has gone’.

No hiatus within words. Kharia does not allow a word-internal hiatus. When an enclitic beginning with a vowel attaches to a word ending in a vowel, either *-w-* (between two vowels belonging to the group */o/, /a/ and /u/*) or *-y-* (otherwise) is inserted.

- (9) *u* ‘this’ *u-w-a?* ‘of this’ *yo* ‘see’ *yo-y-e* ‘she/he will see’
 this-w-GEN see-y-A.IRR

Rarely, *-n-* is encountered with demonstratives instead of *-w-*: *u-n-a?* ‘this-n-GEN’ ‘of this’.

Merger of /e/ with similar vowels and palatal glides. The marker of the irrealis active, $=e$, is dropped after stems ending in $=e$, after the ‘diphthongs’ */ay/, /ui/, /oy/* and before person markers beginning with a vowel (i.e. $/=ijn/$ ‘1SG’ and $/=em/$ ‘2SG’), for example, */karay=e=em/* \Rightarrow *karayem* ‘you will do’, */ter=e=ijn/* \Rightarrow *terijn* ‘I will give’.

3 MORPHOLOGY

3.1 Nominal morphology

Box 9.1 shows the maximal structure of the NP. The order of the unbound morphemes is free to some extent and that given in Box 9.1 may be considered the unmarked order. ‘Lexeme(s)’ covers both the (possibly complex) lexical head as well as all attributes, which directly precede the head.

BOX 9.1: A SCHEMATIC OVERVIEW OF THE NP

Genitive Determiner DEM QUANT Genitive Determiner Lexeme(s)=POSS=NUM= CASE

- (10) *aḍi=ya?* *ghol* *beṭa=ḍom=ki=te*¹²
 ANAPH=GEN 10 SON=3POSS=PL=OBLQ
 ‘his ten sons’ (object)
- (11) *ho* *rusuṅ* *o?*
 that red house
 ‘that red house’

Although not a hard-and-fast rule, there is a clear preference for possessive genitive determiners to appear at the beginning of the NP with the second genitive determiner position reserved for denoting the material something is made of or where something is from:

- (12) *moṅ* *kinir=a?* *jantu*
 1 jungle=GEN animal
 ‘an animal of the jungle (i.e. a wild animal)’
- (13) *aḍi=ya?* *moṅ* *dhaṅgar=ḍom=te* *gam=o?* ...
 ANAPH=GEN 1 servant=3POSS=OBLQ say=A.PST
 ‘He said to one of his servants ...’ (adapted from Malhotra 1982:127)

Note that the markers to the far right in Box 9.1 are enclitic. For example, if the lexical base is not overtly present (e.g. if it is known), these markers simply attach to the right-most element of the phrase, regardless of its status, whether lexical or genitive attribute. That is, unlike affixes, these markers do not attach to stems. Rather, they attach to the entire phrase.¹³ Thus, in the following example, one may assume that the lexical base *lebu* ‘person’ is not mentioned as it is clear from context.

- (14) *ṃunu?siṅ* *rochob=a?*=*ki=ko* (cf. *ṃunu?siṅ* *rochob=a?* *lebu=ki=ko*)
 east side=GEN=PL=CNTR east side=GEN person=PL=CNTR
 ‘the easterners (contrasted)’

In fact, if there is no further information within the NP, as all this is known from context, then these markers attach directly to the demonstratives, which then serve as pronominals of the third person, for example, *ho=ki=te* (that=PL=OBLQ) ‘them’ from *ho* ‘that’.

With respect to the stem itself, it should be noted that the issue of parts-of-speech in Kharia, especially if these are taken to be lexical classes, is quite complex and cannot be dealt with here in detail. For the present I will only note that ANY lexeme,

but also any ‘NP’, may function in Kharia both as a predicate (without any light or ‘dummy’ verb), as the argument of a predicate, or as an attribute of either an argument or a predicate. However, if an element is overtly marked for the oblique case (or an adposition, see sections 3.1.2, 3.1.9), it is an ‘NP’.¹⁴

The only ‘derivational’ processes which could be taken as evidence for the presence of lexical classes in Kharia are the ‘nominalizing’ infix *-NV-* and reduplication in the free-standing form, both of which are almost entirely restricted to monosyllabic stems and usually function as attributes or complements of the predicate.¹⁵

As I argue in section 3.1.10, Kharia strongly disfavours monosyllabic words, hence there are a number of derivational strategies to avoid this, including infixation and reduplication (cf. also Anderson and Zide 2002). As the markers to the right in Box 9.1 may all be lacking (or ‘zero morphemes’), a monosyllabic word could result in the case of a nominal, which should be avoided. In fact, this is true of almost all words in Kharia, and only a few, very common words are monosyllabic, such as *mo²q* ‘eye’, *ti²* ‘hand’ or *o²* ‘house’, the vast majority being bisyllabic.

This is never possible with a predicate, which is always overtly marked for at least one grammatical category and, with that, is always at least bisyllabic. As such, there is no need to ‘strengthen’ predicates by adding a second syllable.

As space is limited here, the interested reader is referred to the discussion in section 3.2, where examples of ‘unusual’ predicates are given, as well as the sections dealing with the ‘nominalizer’ *-NV-* and the free-standing form (sections 3.1.10 and 3.2.8, respectively). A much more detailed discussion of this topic is given in Peterson (2007) and Peterson (2006).

3.1.1 Number

Kharia has three grammatical numbers, both on NPs and predicates: Singular (unmarked), dual (marked by =*kiyar*) and plural (marked by =*ki*).¹⁶ Number marking for non-singular entities is optional in the spoken language, especially if these are non-human. In written texts, however, number marking is generally considered obligatory, even if duality or plurality is indicated elsewhere in the clause.

- (15) *lebu* (*lebu*= \emptyset) ‘man, person’ *lebu*=*kiyar* ‘two men, two people’, *lebu*=*ki* ‘men, people’

The dual also functions as a means of expressing honorific status and is marked on both the NP and the predicate.

- (16) *ho-ka τ =a[?]* *tay konon bahin*=*kiyar* *tama in τ ermi \dot{q} iyat*=*a[?]*
 that-SG.HUM=GEN ABL small sister=HON now intermediate=GEN
paricha likha=*te*=*kiyar*.
 exam write=A.PRS=3.HON
 ‘Her younger sister (HON) is now writing the intermediate exams.’ (AK, 4:12)

3.1.2 Case

Case markers attach directly to the right-most element of the NP. There are three morphological cases:

- Unmarked form* or ‘direct case’ (\emptyset -marking) – found with subjects and non-definite objects;

- Oblique* (marked by =*te*) – marks definite objects, ‘indirect objects’ and adverbials;
Genitive (marked by =(y)aʔ)¹⁷ – the adnominal case, through which one NP becomes an attribute in a larger NP.

Case markers attach directly to the bare nominal (or rather, the last element of the lexical base of the NP). This differs from the adpositions (section 3.1.9), which either always take the genitive or which take the genitive with personal pronominals and proper names. An example:

- (17) *oʔ=te* ‘the house (object); in the house’ CASE MARKER
house=OBLQ
- (18) *og=aʔ bahare* ‘outside the house’ POSTPOSITION
house=GEN outside

There is, however, one construction in which the oblique marker =*te* can co-occur with the genitive: If the semantic head of the NP is not overtly expressed, the (enclitic) oblique marker =*te* attaches to the right-most element of the remaining lexical base of the NP, regardless of its status. If this element is a genitive determiner, this results in apparent ‘double case marking’. It is especially common with predicative alienable possession and with the meaning ‘at the home/place of’:

- (19) *ijn=aʔ=te saykal ayiʔj.*
1SG=GEN=OBLQ bicycle PRS.COP
‘I have a bicycle.’

This may be considered a ‘reduced’ form of the following construction, in which the semantic head is overtly mentioned:

- (20) *ijn=aʔ boʔ=te saykal ayiʔj.*
1SG=GEN place=OBLQ bicycle PRS.COP
‘I have a bicycle.’ (literally: ‘There is a bicycle at my place.’)

The same is also true of the meaning ‘at the home/place of’:

- (21) *am=aʔ=te or am=aʔ boʔ=te*
2SG=GEN=OBLQ 2SG=GEN place=OBLQ
‘at your place’

(19) and (21) are thus similar to (14): As the lexical head, *boʔ* ‘place’, is omitted in these two examples, the enclitic case marker simply attaches to the right-most element of the remaining lexical base, as was the case in (14) with respect to number marking.¹⁸

3.1.3 Person

Kharia distinguishes morphosyntactically between alienable and inalienable attributive possession. With alienable possession, the possessor appears in the genitive before the semantic head. With inalienable possession, instead of a genitive attribute the semantic head itself is marked for the possessive relationship by one of the enclitics in Table 9.2, which cross-reference the possessor (adapted from Malhotra 1982:96).¹⁹

TABLE 9.2: INALIENABLE POSSESSIVE MARKERS

	Singular	Dual		Plural	
		INCL.	EXCL.	INCL.	EXCL.
First person	(=na)= <i>n</i> , (=na)= <i>ɲ</i> (seldom: (=na)=(i)ɲ).	(=na)= <i>naɲ</i>	(=na)= <i>jar</i>	(=na)= <i>niɲ</i>	(=na)= <i>le</i>
Second person	= <i>nom</i>	(=no)= <i>bar</i>		(=no)= <i>pe</i>	
Third person	= <i>ɔom</i>	= <i>ɔom</i> (= <i>kiyar</i>)		= <i>ɔom</i> (= <i>ki</i>)	

As Malhotra notes, these forms are most commonly found with kinship terms, body parts and domestic animals.

- (22) *aba=n* *haɾ=nom* *ma=ɔom*
 father=1SG.POSS bone=2SG.POSS mother=3SG.POSS
 ‘my father’ ‘your bone(s)’ ‘his/her mother’

However, even with these nominals, inalienable possessive markers are not obligatory:

- (23) *am=a?* *ti?*
 2SG=GEN hand
 ‘your hand’

Malhotra (1982:97) notes that the forms for the third persons dual and plural are ambiguous – the number marking may refer either to the possessor or to the possessed entity. Context differentiates between them: *kulam=ɔom=ki* ‘their (PL) brother(s) / his/her brothers’.

There are considerable differences of opinion with respect to these markers: Pinnow (1966:159) lists merely one form each for the first, second and third persons, (=na)=*n*, =*nom* and =*ɔom*, respectively, regardless of number. Similar data are found in Mahapatra (1976:809). This was also confirmed by one speaker I questioned. Nevertheless, counterexamples are easy to find:

- (24) *gandur=na=niɲ* *laɾe=ki*, *laɾe dakha=ki hã?niɲ*.
 Gandur=1POSS=1PL.INCL fight=M.PST fight try=M.PST indeed
 ‘Our Gandur fought, he tried to fight, indeed.’ (Kerkeṭṭā 1990:8)

Data from other speakers I consulted were also in accordance with the data in Table 9.2. Obviously, both systems are in use.

The inalienable construction may also be ‘reinforced’ by the alienable construction:

- (25) *hote=ga manu purkha buɾha=? saw-ɾay=ɔom*
 there=FOC Manu ancestor old.man=GEN spouse-woman=3POSS
jiyom ter=o?
 life give=A.PST
 ‘Right there, the elderly ancestor Manu’s wife died (= ‘gave [her] life’).’
 (MT, 1:110)
- (26) *am=a?* *dhaɲgar=nom gil ɔom=sikh=o?*
 2S=GEN servant=2POSS beat PASS=PRF=A.PST
 ‘Your servant had been beaten.’ (Abbi 1993:551)

3.1.4 Definiteness

Kharia has no definite markers other than the oblique marker with definite direct objects, which is restricted to nominals with definite reference. With ‘indirect’ objects and adverbials, however, its use is more or less mandatory (section 3.1.2).

To refer to an unspecified, non-definite referential entity, *moŋ* or *ek* ‘one’ is used, as in the following example, where the cave is being referred to for the first time.

- (27) *sou²b se maha be²t=dom²⁰ simɽa, moŋ maɽa=te ɖa?*
 all ABL big son=3POSS Simra 1 cave=OBLQ water
kuy=o?
 find=A PST
 ‘His eldest son, Simra, found water in a cave.’ (AK, 1:22)

Word order may also be used to denote the topical status of a sentence constituent, with highly topical information appearing earlier in the sentence and new information to the right. For examples, see section 4.1.

3.1.5 Gender/nominal class

Kharia has neither grammatical gender nor any other type of morphosyntactic nominal classification.²¹ There are several ways of expressing whether a male or female is intended, for example, lexically *kokro siŋkoy* ‘rooster’, *kitur siŋkoy* ‘hen’, through the ‘suffix’ *-ɖay* ‘woman’, for example, *kulam* ‘brother’, *kulam-ɖay* ‘sister’, etc. (adapted from Malhotra 1982:66f.).²²

A number of nominals and attributes which are specified for gender have been borrowed from Indo-Aryan, such as *koɽhi* (m.) / *koɽhni* ‘lazy’ (f.). In the large majority of these forms, however, the masculine form ends in /a/ and the feminine in /i/, for example, *beɽa* ‘son’, *beɽi* ‘daughter’, *budɽha* ‘old man’, *budɽhi* ‘old woman’. These same two endings can also denote large/neutral vs. small entities. Again, this use has been borrowed from Indo-Aryan, such as *ɖoŋga* ‘boat’ vs. *ɖoŋgi* ‘small boat’ (see Biligiri 1965:144f.).

3.1.6 Pronominals

The free-standing pronominals are presented in Table 9.3. The use of these pronominals in Kharia is optional when the identity of the referent is not focused, that is, Kharia is a ‘pro-drop’ language.

As noted in section 3.1.1 above, the dual is also used to denote honorific status. When referring to oneself to someone to whom respect is due, the exclusive form of the first person, *iŋjar*, is used. This use, however, is restricted to singular entities in the first and second persons: There are no special honorific forms for first and second persons, dual and plural, although the ‘dual’ may be used to denote politeness with a third person, singular or plural entity.

In the non-singular first persons, there is an inclusive/exclusive distinction. The inclusive denotes that the speaker includes the addressee in the reference (‘we including you’), whereas with the exclusive, the addressee is excluded (‘we without you’). There are two further distinctions in the third persons:

Anaphoric vs. unmarked. There are two sets of forms for the third persons, one based on *aɖi*, the other on the demonstratives (see section 3.1.7). The difference between

TABLE 9.3: THE FREE-STANDING PRONOMINALS

	Singular	Dual / Honorific		Plural	
		INCL	EXCL	INCL	EXCL
1	<i>ɨn</i> (less commonly: <i>ɨj</i>)	<i>anaɨ</i>	<i>ɨɲjar</i>	<i>aniɨ</i>	<i>ele</i>
2	<i>am</i>	<i>am=bar</i>	<i>am=pe</i>		
3	<i>aɟi</i> , <i>ho-kaɟ</i> , <i>ho-je?</i> , <i>u-kaɟ</i> , <i>u-je?</i> , <i>hin-kaɟ</i> , <i>hin-je?</i> , <i>han-kaɟ</i> , <i>han-je?</i>	<i>aɟ(i)=kiyar</i> , <i>ho=kiyar</i> , <i>u=kiyar</i> , <i>hin=kiyar</i> , <i>han=kiyar</i>		<i>aɟ(i)=ki</i> , <i>ho=ki</i> , <i>u=ki</i> , <i>hin=ki</i> , <i>han=ki</i>	

these two groups is as follows: The forms based on the demonstratives, such as *ho* ‘that’, can be used in cataphoric, anaphoric and deictic function whereas those based on *aɟi* can only be used anaphorically and refer back to the topic.²³ Also, forms based on *aɟi* can only be used with reference to humans.

- (28) *sou²b se maha be²t=ɟom, simɽa, moɲ maɽa=te ɟa?*
 all ABL big son=3POSS Simra 1 cave=OBLQ water
kuy=o? **aɟi** *uɽh=o?*
 find=A PST ANAPH drink=A.PST
 ‘His eldest son, Simra, found water in a cave. **He** drank [it].’ (AK, I:22)

The unmarked pronominals (*hokaɟ*, *hoje?*) are literally ‘unmarked’ in this respect: In addition to their deictic and cataphoric functions, they can also be used anaphorically. In fact, they often refer to the same entity as *aɟi* within the space of a single sentence, as in the following example.

- (29) *ghaɟ aɟi=te baɽa kharab la?²=ki. tay moɲ moɲ*
 therefore ANAPH=OBLQ big bad EMOT=M.PST then 1 REP
bhai=ki=ya? nimi ɟho²ɟ=ta ɟho²ɟta ho-kaɟ
 brother=PL=GEN name take=CV REP that-SG.HUM
ɽhisa=na maɽe=yo?
 call.out=INF start=A.PST
 ‘Therefore, **he** became very sad. Then, taking the names of the brothers one by one, **he** began to call out.’ (AK, 1:27f.)

Human/Non-human. Third-person, unmarked pronominals distinguish between human and non-human reference in the singular. These pronominals derive from a demonstrative such as *u* ‘this’, *ho* ‘that’ (MEDIAL) or *han / hin* ‘that’ (REMOTE) plus a second, derivational element: The pronominals for human reference make use of *-kaɟ* ‘person’, which is quite common in compounds but whose use as a free morpheme is now largely obsolete. Pronominals for non-human reference make use of *-je?*. *-je?* is also often found alone in this function, but does not appear to have any lexical meaning.

- (30) *u-je?²=ga heke manus jati=ya? kahani.*
 this-SG.NHUM=FOC PRS.COP man ethnic.group=GEN story
 ‘This is the story of humanity.’ (AK, 3:17)

TABLE 9.4: INTERROGATIVES

Free morphemes		Bound morpheme
<i>ata</i> 'what?, which?'	<i>ber, behar</i> 'who?'	<i>a-</i> 'Q'
<i>atu</i> 'where?'	<i>i</i> 'what?'	
<i>ina</i> 'why?'		

For an example of *ho-kaṛ* with human reference, see example (29) above.

Interrogatives. There are six basic interrogative forms, given in Table 9.4.

a- and *i* combine with free morphemes and with case markers to derive other interrogative compounds:

- (31) *a-bo?* 'where?' (*bo?* 'place')
a=te 'where?' (=te 'OBLIQUE')
i-bhere 'when?' (*bhere* 'time')
i=te 'where?' (=te 'OBLIQUE')
i-ghay 'how?' (*ghay* 'way')

Interrogatives may also be used predicatively:

- (32) *am i=yo'b?*
 2SG what=A.PST.2SG
 'What did you do?'
- (33) *am=te i=ki?*
 2SG=OBL what=M.PST
 'What happened to you?'

Indefinites. Some interrogatives, such as *atu* 'where', combine with the additive focus marker =*jo* 'also' to serve as indefinites, for example, *atu=jo* 'wherever'. For some indefinites, however, there is no corresponding interrogative, such as *jahā* 'something' or *jahāy(=ga)* 'whoever'. *jahā* may also be used attributively. The most common indefinites are:

- (34) *jahāy, jahāy=ga, ber=jo* 'whoever' *atu=jo* 'wherever'
jahā, i=jo 'some(thing); whatever' *jahā bhere* 'whenever'

3.1.7 Demonstratives

Kharia has a three-way demonstrative system with evidence for an earlier four-way system:

- (35) *u* 'this', near to speaker
ho 'that', medial distance from speaker
hin, han 'that', farther away from speaker

hin and *han* differ little in meaning in actual usage, if at all. Along with Malhotra (1982:53), I can find no evidence for Biligiri's (1965:65) analysis of *han* as restricted to visible and *hin* to non-visible entities. However, unlike Malhotra, I find no evidence for *hin* and *han* being in complementary distribution. Instead, my data indicate that they are in free distribution, with *hin* being the more common of the two.

There is yet another demonstrative with the same meaning as *han* and *hin -se*, a borrowing from Sadani. It is only seldomly encountered and is generally only found with temporal adverbials, such as *se bhere* ‘that time’, *se dinu* ‘that day’, but cf. also *se lekhe* ‘like that’.

These demonstratives may appear at the beginning of an NP as a determiner, preceding all other NP constituents, or, with the exception of *se*, they may combine with *-kaɽ* ‘person’ and *-jeʔ* ‘3SG.NHUM’ to form pronominals (section 3.1.6). Again, with the exception of *se*, they may also combine with the oblique marker =*te* to form locative adverbials, with *ghay* ‘way’ to form manner adverbials, with *tiʔj* ‘side, direction’ to form goal-directed locative adverbials, or with *bhere* to form temporal adverbials:

- (36) *u goɽa duniya=te*
 this entire world=OBLQ
 ‘in this entire world’
- (37) *hin bhere* ‘(at) that time’
u-kaɽ ‘she/he (here)’, *ho-kaɽ* ‘she/he (there)’
hin-kaɽ, *han-kaɽ* ‘s/he (remote)’
u=te ‘here’, *han=te* / *hin=te* ‘there’
ho-ghay ‘that way, like that’
han-tiʔj u-tiʔj ‘this way and that’

3.1.8 Numerals and classifiers

There are two sets of numerals in Kharia. One, of Kharia origin, is no longer in common use and shows a great deal of variation. The other is borrowed directly from Sadani. Only the Sadani numerals are at all common, and even the lower numerals in actual speech are all borrowed from Sadani. Numerals of Sadani origin generally occur with numeral classifiers, although this is not obligatory.

*Kharia Numerals.*²⁴ There is great deal of uncertainty concerning these numerals, and only the forms for ‘1–3’ and ‘10’ are uniform for all speakers. Other numerals can vary greatly from speaker to speaker. In Box 9.2, beginning with ‘4’, the first numeral given is the most commonly cited form, followed by less common variants.

There are two systems for the numerals from ‘11’ to ‘19’. The first simply combines the word for ‘10’ with the respective single digit. There is also another system in which all numerals from ‘11’ to ‘19’ have their own designation. This system is given to the right in Table 9.5.²⁵

The designation for ‘20’ is *ekɽi* (< Indo-Aryan, cf. Bengali *kuɽi* ‘a score’ Sadani *kori* ‘a score’). Pinnow (1965a:6) also cites *ukɽi*. For ‘21–29’, the single digits are added to *ekɽi*, for example, *ekɽi moɽ* ‘21’, etc.

BOX 9.2: KHARIA NUMERALS 1–10

1	<i>moɽ</i> (NHUM), <i>muɽu</i> (HUM)	6	<i>tibru</i> , <i>tibhru</i> , <i>ɽibru</i>
2	<i>ubar</i>	7	<i>tham</i> , <i>thom</i> , <i>thoy</i> , <i>ghul</i>
3	<i>uʔphe</i>	8	<i>thom</i> , <i>ghal</i> , <i>tham</i> , <i>thomsij</i>
4	<i>iʔphon</i> , <i>tham</i>	9	<i>thomsij</i> , <i>ghal</i> , <i>tomsij</i> , <i>ghul</i>
5	<i>moloy</i> , <i>thum</i>	10	<i>ghol</i>

BOX 9.3: KHARIA NUMERALS 11–19

11	<i>ghol moŋ</i>	<i>ghul</i>	16	<i>ghol tibru</i>	<i>rabe?</i>
12	<i>ghol ubar</i>	<i>gholsiŋ</i>	17	<i>ghol tham</i>	<i>tarsi?</i>
13	<i>ghol u?phe</i>	<i>ʔak</i>	18	<i>ghol thom</i>	<i>qabayeʔj</i>
14	<i>ghol i?phon</i>	<i>ʔoŋa</i>	19	<i>ghol thomsisŋ</i>	<i>qubki</i>
15	<i>ghol moloy</i>	<i>raba?</i>			

TABLE 9.5: KHARIA NUMERALS 30–100

Vigesimal system	System of decades
30 (<i>moŋ</i>) <i>ekŋi ghol</i> , literally ‘(1) 20 10’	<i>u?phe ghol</i> ‘3 10’
31 (<i>moŋ</i>) <i>ekŋi ghol moŋ</i> , etc.	<i>u?phe ghol moŋ</i> , etc.
40 <i>ubar ekŋi</i> ‘2 20’	<i>i?phon ghol</i> ‘4 10’
50 <i>ubar ekŋi ghol</i> ‘2 20 10’	<i>moloy ghol</i> ‘5 10’
60 <i>u?phe ekŋi</i> ‘3 20’	<i>tibru ghol</i> ‘6 10’
70 <i>u?phe ekŋi ghol</i> ‘3 20 10’	<i>tham ghol</i> ‘7 10’
80 <i>i?phon ekŋi</i> ‘4 20’	<i>thom ghol</i> ‘8 10’
90 <i>i?phon ekŋi ghol</i> ‘4 20 10’	<i>thomsisŋ ghol</i> ‘9 10’
100 <i>moloy ekŋi</i> ‘5 20’	<i>moloy ekŋi</i> ‘5 20’

There are no simple words in the Kharia numeral system for ‘30’, ‘40’, etc. There are two different strategies for expressing numbers above ‘29’.

- The first, vigesimal system counts in scores and adds the appropriate numeral from ‘1’ to ‘19’ to the correct multiplication of ‘20’.
- The second system counts in decades, with multiples of ‘10’ plus the respective single digit. The fact that this system has no decadal term for ‘100’ suggests that it is the younger of the two.

These two systems are shown in Table 9.5.

Native ordinals. The native term for ‘first’ is generally *seŋ*, for example, *seŋ qel-qel=a? lebu* (first come-REDPL=FOC person) ‘the people who came first’. Alternatively, the genitive of *absiʔb* ‘beginning’ can also be used, although this is rare.

- (38) *absib=a? eto?*
 begining=GEN order
 ‘(the) first commandment’ (lit.: ‘the commandment of the beginning’)

Other ordinals are derived as follows: In reference to humans, *-kaŋ* is attached to the corresponding cardinal number:

- (39) *u?phe-kaŋ*
 3-SG.HUM
 ‘(the) third person’

For non-humans/inanimates, there are three strategies: First, the respective cardinal number is placed in the genitive:

- (40) *ubar=a? etoŋ*
 2=GEN order
 ‘(the) second commandment’

Alternatively, beginning with *ubar* ‘2’ the compound form *-si²b* (< *absi²b* ‘beginning’) is used to form ordinals, for example, *bar-si²b āk* ‘second act (of a play)’, etc. Note that the /u/ in *ubar* is dropped here.

In a third strategy, cardinal numbers are used, with no special marking: *ubar ebo²* ‘second act/scene’.

Finally, Malhotra (1982:126) writes of another general strategy for ordinals beginning with ‘second’, which is not found in my corpus. She notes that ordinals above ‘first’ are formed by using the postpositions ‘after’ or ‘behind’ following the numeral. Her two examples (adapted):²⁶

- (41) *boriya=?* *lo?dho* *moloy=a?* *kundā²b*
 both=GEN after five=GEN behind
 ‘third’ ‘sixth’

Indo-Aryan Numerals (from Sadani). As mentioned above, only the numerals which have been borrowed from Sadani are in common use. Box 9.4 presents the lower numerals and a few of the higher numerals.

Ordinals. Ordinals in common use have also all been borrowed from Sadani: first *pohila*, second *dusra*, third *tisra*, etc.

Collectives. With the exception of *bariya*, which is of Kharia origin, collectives are formed by adding =*o* directly to the Sadani numeral. This =*o* is probably identical with the homophonous classifier =*o* (see below), although it may be related to the Hindi oblique plural suffix *-ō*. Box 9.5 presents examples for the lower collectives.

Classifiers. There is a small number of optional classifiers in Kharia. These are found only with cardinal numerals and, rarely, with quantifiers such as *kai* ‘a few’. Here are the most common ones.

=*o*

=*o* is the general classifier and is used with all types of nominals. It is by far the most common classifier.

BOX 9.4: NUMERALS BORROWED FROM SADANI

0 <i>sun</i>	6 <i>chaw</i>
1 <i>ek</i>	7 <i>sat</i>
2 <i>dui</i>	8 <i>aṭh</i>
3 <i>tin</i>	9 <i>naw, nāw</i>
4 <i>cair, ceir</i>	10 <i>das</i>
5 <i>pāc</i>	Others: 100 <i>say, saw, sos</i>
	1,000 <i>hajar</i>
	100,000 <i>lakh</i>

BOX 9.5: COLLECTIVES

<i>bariya, beriya, boriya</i> ‘both’	<i>chaw=o</i> ‘all six’
(cf. <i>ubar</i> ‘2’)	<i>sat=o</i> ‘all seven’
<i>tin=o</i> ‘all three’	<i>aṭh=o</i> ‘all eight’
<i>cair=o</i> ‘all four’	<i>naw=o</i> ‘all nine’
<i>pāc=o</i> ‘all five’	<i>das=o</i> ‘all ten’

- (42) *duy=o daru* ‘two trees’ *chaw=o sorej* ‘six stones’ *tin=o bhai* ‘three brothers’
tho

tho is another all-purpose classifier, less common than =*o*. Before *tho*, ‘6’ is realized as *che*.

- (43) *pāc tho ore²j* ‘five oxen’ *che tho sorej* ‘six stones’ *kai tho ompay* ‘a few rivers’
jan l jhan

j(h)an is used exclusively for humans. It is less common than =*o* in this function and has been borrowed from Sadani *-jan l -jhan*.

- (44) *tin jhan lebu=ki* ‘three people’ *tin jhan bhai=ki* ‘three brothers’

j(h)an also attaches to names to denote the person who is explicitly mentioned as well as either those accompanying him or her at that moment or his or her family:

- (45) *birsi jhan* ‘Birsi and her friends (who are with her now).’

3.1.9 Adpositions

Kharia has a large number of postpositions and one preposition. Adpositions differ from case markers (section 3.1.2) in that adpositions generally require the genitive case when the semantic head of the NP is a pronominal or a proper nominal, otherwise the bare nominal. Case markers, on the other hand, attach directly to the last element of the bare NP. Here I present some of the more common adpositions.

The preposition enem ‘without’ This is the only preposition in the language. Normally, the NP appears in the direct case: for example, *enem go²jhuj* ‘without a path’, but it can also appear in the genitive:

- (46) *ani jahā-ti²j col=ta=niṅ. ute=ko enem*
come.on INDEF-side go=M.PRS=1PL.INCL here=CNTR without
ḍa²=ya² go²j go²ḍ=na=niṅ.
water=GEN die C:TEL=M.IRR=1PL.INCL
‘Come on! Let’s go somewhere. Here we will die without water.’ (TK, 1:7)

In subordination, an NP whose lexical head is a free-standing form (here: *aw-aw*)²⁷ appears in the oblique case to denote the adverbial status of the subordinate clause. The entire construction is then negated by *enem*.

- (47) *enem raja ro rani=kiyar=a² aw-aw=te khaṛiya=ki*
without king and queen=DL=GEN COP-REDPL=OBLQ Kharia=PL
kaṭi²j=ko ḍher=ga akbakay=ki=may.
somewhat=CNTR very=FOC be.in.a.flurry=M.PST=3PL
‘Without there being a king and queen, the Kharia panicked.’ [MT, 1:104]

Postpositions. There are simple and complex postpositions in Kharia. Most simple postpositions are not further analysable from a synchronic perspective. Here are some of the most common:

- (48) *buṅ* ‘with; INS; COM’ *tay, se* ‘ABL’
gha²ḍ ‘for; PURP’ (cf. *gha²ḍ* ‘opportunity’ (?)) *thom, thoṅ* ‘for; PURP’

As already noted, these postpositions generally take the genitive with pronominals and proper names and the unmarked form with lexical nominals:

- (49) *ijn=a?* *buŋ* *ho-kaɽ=a?* *tay* *kinir tay*
 1SG=GEN COM that-SG.HUM=GEN ABL forest ABL
 ‘with me’ ‘from him/her’ ‘from (the) forest’

Complex postpositions derive from oblique-cased relational nominals, whereby the oblique marker is often omitted.

- (50) *bahar(=te)* ‘outside of’ (*bahar* ‘outside’)
mugam(=te) ‘in front of’ (*mugam* ‘front’)
kunɽa²b(=te) ‘behind’ (*kunɽa²b* ‘back’)
tobluy(=te) ‘on top of, above’ (*tobluy* ‘top’)
mo?jhi(=te) ‘amidst, among’ (*mo?jhi* ‘middle’)
tuta(=te) ‘under, below’ (*tuta* ‘bottom’)

As with the simple postpositions, complex postpositions take the genitive with pronominals and proper names, but show a higher tendency to take the genitive with lexical nominals as well.

- (51) *sadhu=ki=ya?* *mo?jhite*
 holy.man=PL=GEN amongst
 ‘amongst the holy men’

However, they do not require the genitive: *tunbo? mo?jhite* ‘in the middle of the day’.²⁸

3.1.10 Derivation: ‘nominalizers’

There are two non-productive derivations which are generally considered nominalizers. The first, which is found in a number of common words, makes use of the infix *-NV-*, where *-V-* has the same quality as the vowel preceding *-N-* in all cases other than the slightly irregular *mo<ne>n* ‘one each’ from *moŋ* ‘one’. *-N-* is a nasal of indeterminate quality. It is usually realized as /n/, but /m/ is also occasionally found.²⁹ Box 9.6 presents a few examples.

In many cases the simple, non-derived form no longer exists:

- (52) **ga²d* ‘reap’ *ga<na>²d* ‘sickle’
 (**)sel* ‘pray’ *se<ne>l* ‘prayer’

BOX 9.6: THE ‘NOMINALIZER’ *-NV-*

<i>bel</i> ‘spread out (a mat or bedspread)’	<i>be<ne>l</i> ‘bedding’
<i>dɛ²j</i> ‘chop’	<i>dɛ<ne>²j</i> ‘large hatchet’
<i>ku²j</i> ‘dance’ (predicate)	<i>ku<nu>²j</i> ‘dance’ (nominal)
<i>moŋ</i> ‘one’	<i>mo<ne>n</i> ‘one each’
<i>ro²d</i> ‘be sad’	<i>ro<mo>²d-dɔ?</i> ‘tear’ (i.e. ‘sad-water’)
<i>si</i> ‘plow’ (predicate)	<i>si<ni></i> ‘plow’ (nominal)

There is no root **ga²d* in Kharia, but such a lexeme is found, for example, in the South-Munda language Sora with the meaning ‘cut’. Similarly, while *sel* with the meaning ‘pray’ is still found in the texts in Pinnow (1965a), speakers I questioned all rejected the form, and use *senel* both predicatively (‘pray’) and referentially (‘prayer’).

The designation of this infix as a nominalizer is not without problems: Modern Kharia has the general requirement that free-standing lexemes be at least bisyllabic, a requirement which derives from an earlier bimoraic constraint (see Anderson and Zide (2002), discussed in Peterson (2007)). For monosyllabic lexemes to be used as the head of an NP, this means that some strategy is needed to derive bisyllabic forms, since possessive, number and case markers may all be lacking (or ‘zero morphemes’). There are only very few exceptions to this rule, namely commonly occurring nominals, such as *o?* ‘house’, *ti?* ‘hand’ or *mo²d* ‘eye’, but the vast majority are bisyllabic. Nowadays, monosyllabic roots are reduplicated in these environments as free-standing forms (section 3.2.8) whereas the older language seems to have preferred the *-NV-* infix. At any rate, forms such as *mo<ne>n* ‘one each’ from *mon* ‘one’ call any analysis of this marker as a nominalizer into question (for a more detailed discussion of parts-of-speech in Kharia, see Peterson, 2006, 2007).

The second construction, discussed in Pinnow (1965a:161, note), combines the infix *-NV-*, just discussed, with the reduplication of the root, typical of free-standing forms. Instead of *-NV-*, *-NVm-* is also found, and the consonant in the coda of the lexeme is deleted. Table 9.6 presents the four examples for this construction which I have come across. The first three are from Pinnow, the last from my own fieldwork:

TABLE 9.6: INFIX *-NV-* WITH REDUPLICATION

Root	Derived form
<i>d²el</i> ‘come’	<i>d²e<nem>-d²el</i> ‘having arrived’
<i>go²j</i> ‘die’	<i>go<nom>-go²j</i> ‘deceased (person)’
<i>ra²b</i> ‘bury’	<i>ra<na>²b-ra²b</i> ‘burial ground; grave’
<i>kol</i> ‘count’	<i>ko<no>l-kol</i> ‘population’

It is extremely difficult to determine the function of this form. It does not appear to have ever been very productive and the four forms given here are the only ones I have been able to locate, compared with at least 45 of the type *-NV-*.

3.1.11 Nominal attributes

If one uses only morphosyntactic criteria as a guide to determining parts-of-speech, then there is little evidence for assuming a separate class of adjectives in Kharia. Instead, from a purely morphosyntactic point of view it is more convenient to consider simple nominal attribution to be expressed by a free-standing form, described in section 3.2.8, that is, stems appear in their basic form if they are polysyllabic, or are reduplicated if monosyllabic. Some examples:

- (53) *konon bhai* ‘little brother’ cf. *konon* ‘become small’ *konon* Free-standing form:
doko thāro ‘settled place’ *doko* ‘sit down; settle’ *doko*

<i>maha daru</i> ‘the big tree’	<i>maha</i> ‘become big’	<i>maha</i>
<i>ruʔ-ruʔ kaʔbʃo</i> ‘open door’	<i>ruʔ</i> ‘become open’	<i>ruʔ-ruʔ</i>
<i>rusuŋ oʔ</i> ‘red house’	<i>rusuŋ</i> ‘become red’	<i>rusuŋ</i>

Note that reduplication here is not in any way restricted to ‘verbs’ or lexemes denoting an event vs. a state. This is especially clear in a case like *qoko ro dhoʔ-dhoʔ thāʔo* (settle.down and take-REDPL place) ‘the place where [they] settled down and which [they] took’. As *qoko* is bisyllabic, it is not reduplicated, whereas the monosyllabic *dhoʔ* is.

Although no longer in current use, there also appears to have been a root *say* ‘become yellow’, which had the form *say-say* when used attributively, that is, ‘yellow’, and which is given with this meaning in Floor *et al.* (1934:112). This form has now been replaced by the Sadani term *piyar* ‘yellow’, as all colour terms except *osel* ‘white’, *mogher* ‘black’ and *rusuŋ* ‘red’ have been replaced by the corresponding Sadani equivalents,³⁰ and in current usage *say-say* means only ‘turmeric’, a meaning which it has retained from very early times (see Anderson and Zide 2002).

On the other hand, ANY lexeme that can be used attributively can also be used to denote a change-of-state (middle voice) or a transitive telic action (active voice) and vice versa, hence there seems to be little justification for assuming an adjectival class in Kharia.

There are only a few monosyllabic lexemes which do not fit this pattern and which do not reduplicate (except to denote plurality, etc. see below), all loans from Indo-Aryan: *khus* ‘(become) happy’, *lil* ‘(become) blue’ and *bes* ‘(become) good’, for example, *bes dinu* ‘good day’.

Finally, the repetition of attributes can denote both intensity (‘very’) and plurality. When used to denote plurality, the NP is not generally marked for number, as with *konon konon murti* in the following example:

- (54) *ho-kaʔ loʔkha qoʃh=oʔ ro loʔkha qoʔq=na loʔqho*
 that-SG.HUM soil take=A.PST and soil take=INF after
i-ghay aʔi=yaʔ rup buŋ=ga konon konon murti
 what-way ANAPH=GEN form INS=FOC small REP statue
bay=oʔ.
 make=A.PST
 ‘He took soil and, after taking soil he made small statues in (= with) his [own] form.’ (AK, 3:10)

Comparatives and superlatives. Comparatives and superlatives are formed as follows: The standard is followed by an ablative postposition and the lexeme denoting a scalar property, appearing in its base form. The two constructions are identical except for the fact that the standard in the superlative construction contains a quantifier denoting entirety, such as *jhaʔi* or *souʔb* ‘all’. Also, there is a distinct preference for the ablative postposition *tay* in the comparative construction and *se* in the superlative construction, although both may be used in both constructions.

Comparative:

- (55) *laʔ aʔi=yaʔ tay konon bhai, beghma, ho-kaʔ*
 then ANAPH=GEN ABL small brother Beghma that-SG.HUM
onqor=oʔ.
 hear=A.PST
 ‘Then his_i younger brother, Beghma_j, he_j heard [him_i].’ (AK, 1:30)

Superlative:

- (56) *sou²b se maha=ya? biha hoy=si?*
 all ABL big=GEN wedding become=PRF
 ‘The oldest has married (= The marriage of the oldest [daughter] has become).’ (AK, 4:9)

Note that this construction is used with all scalar lexemes and is not limited to ‘attributive’ lexemes, thus it cannot be considered a typical property of ‘adjectives’.

- (57) *sou²b se suru=te ponmesor ap aw=ki.*
 all ABL beginning=OBLQ God Father COP=M.PST
 ‘In the very beginning there was God the Father.’ (AK, 3:4)

3.1.12 *Adverb(ial)s*

The most common means of modifying a predicate or an entire clause is by suffixing *-bo?*, *-da?* or *-son* to a lexeme. The choice of one over the other is lexically determined. Some lexemes can also take more than one of these markers, with no difference in meaning.

- (58) *anjor* ‘front’ *anjor-da?* ‘forwards’
bes ‘good’ *bes-bo?*, *bes-da?* ‘well’
dular ‘love’ *dular-bo?*, *dular-son* ‘lovingly’
lere? ‘rejoice; joy’ *lere?-son* ‘joyously’
- (59) *ro tama am=pe u naw kuṭum=te=ga sadi biha*
 and now 2=2PL this 9 family=OBLQ=FOC marry marry
kerson=na=pe ro ... baru-bo? aw=na=pe.
 marry=M.IRR=2PL and good-INTENS live=M.IRR=2PL
 ‘And you will marry in only these nine families and you will live very well.’
 (AK, 1:69)

However, this construction is not exclusively used to mark adverbials, and forms marked by one of these markers can also appear in predicative function:

- (60) *ho-kaṛ iku²d jughay lere?-son=ki.*
 that-SG.HUM very much joy-INTENS=M.PST
 ‘She/he became very happy.’

What these apparent ‘adverbials’ in fact express is a slight degree of ‘intensity’, that is, they denote that an event or state is ‘more intense’ than the unmarked forms, similar to the function of *very* in English. Hence they are not restricted in any way to an adverbial function, although they are certainly compatible with this usage.

For other types of ‘adverbial’ modification, see the sections on the sequential and imperfective converbs (section 3.2.8) and ‘adverbial’ clauses (section 4.5).

3.2 Predicative morphology

Box 9.7 gives a simplified overview of the maximal possible structure of a non-negated, non-periphrastic predicate with a single stem. There is some freedom of order with respect to the markers between the stem and the perfect marker, all of which have the status of grammatical words. They correspond to what are generally referred to in South Asian linguistics as ‘V2s’ or ‘explicator verbs’ (section 3.2.10).

The reciprocal marker may also be considered a word as it shows some degree of freedom of movement (section 3.2.11 under ‘Semi-productive incorporation’) and has a rising contour, typical of monosyllabic phonological words (section 2.2).

Box 9.7 gives the general, unmarked order of some of these constituents, which are too numerous to all be portrayed here (‘< >’ denotes an infix).

BOX 9.7: A SCHEMATIC OVERVIEW OF THE NON-NEGATED MORPHOLOGICALLY FULLY FINITE PREDICATE IN KHARIA

RECIP CAUS-**Lexeme**-<CAUS> AKT BEN PASS/RFLX AUTOPOES TEL=PRF=TAM/VOICE=PERS/NUM/HON

The causative marker is an affix which may only appear either directly before and/or within the lexical head. The markers to the right, beginning with the perfect, are all enclitic markers and attach directly to the right-most element of the predicate phrase, regardless of its status (see examples in the following section beginning with [66]).

Types of predicating bases. Typical of Kharia is the fact that the predicate base may be any type of lexeme, a nominal in the genitive, an entire ‘NP’, or it may consist of several lexical morphemes. Let us now look at these in somewhat more detail.

In theory at least, any lexical morpheme may function as the head of a predicate without the need for a light verb. For example, free-standing interrogatives can function as the predicate of a clause in Kharia with no derivational marking. In these cases, the middle voice denotes a change-of-state, *i=na?* (what=M.IRR) ‘what will happen?’, whereas the active denotes an activity, *i=yem?* (what=A.IRR=2SG) ‘What will you do?’. Another attested example:

- (61) *bəʔo=te=ga tiri²b ɖal=e la? i-ghay=na?*
 sun=OBLQ=FOC cloud cover=A.IRR then what-way=M.IRR
 ‘If a cloud covers the sun, then how will it be?’ (Kerkeṭṭā 1990:5)

This is only true with dynamic predicates. If a state is described, then the copula is used:

- (62) *i heke?*
 what PRS.COP
 ‘What is [that]?’

Similarly, indefinites may serve as predicates:

- (63) *jahā guru?=may, anij=te sa?dhe=si?=may, ho-je?*
 INDEF OPT=3PL 1PL.INCL=OBLQ torment=PRF=3PL that-SG.NHUM
bhagwan=jo=ko yo=te²j.
 god=ADD=CNTR see=A.PROG
 ‘Let them do as they like (= let them whatever), they have tormented us, and God is watching that as well.’ (Kerkeṭṭā 1990:7)

Despite this theoretical possibility, speakers’ intuitions differ greatly on the extent to which typically referential entities may be used predicatively, at least in interviews. Consider the following example:

- (64) *bhagwan lebu=ki ro ɖel=ki.*
 God man=M.PST and come=M.PST
 ‘God became man and came [to earth] (i.e. Jesus).’ (Malhotra 1982:136)

While no speakers objected to the second predicate in example (64), that is, *qelki*, opinions varied somewhat with regard to the first predicate, *lebuki*: For some speakers I consulted, *lebuki* is a perfectly grammatical form, although not the most natural means of expressing this thought, as the use of the transitional copula *hoy* ‘become’ is preferred. However, some speakers rejected it outright.

Similar data were obtained for typically ‘pronominal’ deictics. Although all insisted that the use of the copula would be more appropriate, some speakers were reluctant to accept the following example, while others had no difficulty in doing so.

(in a play about me and you, in which both of us will be taking part):

- (65) *‘naʔak=te in=ga ho-kaʔ=na=in ro am=ga*
 play=OBLQ 1SG=FOC that-SG.HUM=M.IRR=1SG and 2SG=FOC
in=na=m.’ ‘umboʔ. am=na um=in pal=e. direktar
 1SG=M.IRR=2SG no 2SG=INF NEG=1SG be.able=A.IRR director
seŋ=gaʔ in=te ho-kaʔ=oʔ am=ga am=na=m’
 early=FOC 1SG=OBLQ that-SG.HUM=A.PST 2SG=FOC 2SG=M.IRR=2SG
 ‘“In the play I will be him and you will be me.” “No. I can’t be you. The director already made me him. You will be you.”’

(Pro)nominals marked for the genitive can also function as predicates. This is a fully productive process but only holds for dynamic predicates:

- (66) *in ho-kaʔ=te in=aʔ=yoʔj.*
 1SG that-SG.HUM=OBLQ 1SG=GEN=A.PST.1SG
 ‘I adopted him/her (i.e. I made him/her mine).’

However, this is not true of oblique-marked NPs with any meaning: **sahar=te=ki* ‘city=OBLQ=M.PST’, **sahar=te=yoʔ* ‘city=OBLQ=A.PST’. The reason for the different treatment of the two morphologically marked cases appears to be straight-forward: The oblique marker *=te* marks an entire NP, which plays a role – either as object or adjunct – at the clause level. The genitive on the other hand occurs within the NP – it is not relevant at the clause level but is comparable to any other nominal attribute in this respect, such as the so-called adjectives discussed in section 3.1.11, all of which can also function as predicates. This is demonstrated most clearly in examples such as the following: If a possessive relationship exists between two or more entities, then the entire ‘NP’, including the genitive attributes, can function as a predicate:

- (67) *ho-kaʔ in=aʔ natgot=ki.*
 that-SG.HUM 1SG=GEN family.relationship=M.PST
 ‘She/he became a member of my family (e.g. through marriage).’

Finally, a head ‘nominal’ with a quantifier (68) or with a demonstrative (69) may function as a predicate:

- (68) *ubar rochoʔb=ki=n.*
 2 side=M.PST=1SG
 ‘I moved to both sides (i.e. this way and then that).’
- (69) *ho rochoʔb=ki=n.*
 that side=M.PST=1SG
 ‘I moved to that side’

In summary, as long as it expresses a change-of-state or action, any lexeme or ‘NP’ can serve as the stem of a predicate without the need for a light or ‘dummy’ verb, provided that a proper context can be found.³¹

The semantics of these lexical units in different functions are completely predictable. Although we cannot go into detail here, any lexical morpheme which denotes a material entity *X* has the meaning ‘become *X*’ in the middle voice and ‘turn something into *X*’ in the active. Similarly, lexemes denoting a property *Y* have the meaning ‘become *Y*’ (middle) or ‘make something *Y*’ (active). As we shall see in section 3.2.8, lexical morphemes denoting an event *Z* also have a predictable meaning when used as complements – ‘the act of *Z*-ing’. For a more detailed discussion, see Peterson (2007).

Of course, convention and world knowledge will serve to restrict these theoretical possibilities in everyday speech. For example, *tebul* ‘table’ usually refers to an entity on which things may be placed, is made of wood, etc. Things seldom ‘become a table’, which is what the predicative use of this lexeme in the middle voice would mean, and one seldom ‘turns something into a table’, which is what it would mean in the active. It thus comes as no surprise that speakers only reluctantly accepted this lexeme as a predicate although, once a proper context was found (e.g. fairy tales), most did accept this usage.

I therefore conclude that the only reason speakers were either reluctant or unwilling to accept predicative uses of lexemes such as these was my inability to find a natural-sounding context in which an action or event of this type can be conceived. There are no lexical or morphosyntactic rules which disallow such predicates. There are, however, clear personal preferences.

Simple and complex stems. The predicate may contain one or more stems, each of which may be marked for derivational categories such as the reciprocal or causative. Some examples (stems are underlined):

Single stems:

- (70) *mu*?=ki=*may* *bho*<²*b*>*re* *go*²*d*=*sikh*=*o*?=*ki*
 go.out=M.PST=3PL fill.up-<CAUS> C:TEL=PRF=A.PST=PL
 ‘they went out’ ‘they had filled [it] up.’

Complex stems:

- (71) *musniŋ jhaŋi kopuŋu*?=*ki jume kati*²*b kon socay*=*o*?=*ki no*
 one.day all man=PL assemble gather SEQ think=A.PST=PL COMP
 ‘*ho dano*=*te i-gud*?=*ga tar o*?=*gur*=*e*=*niŋ*?’
 that demon=OBLQ what-like=FOC kill CAUS-fall=A.IRR=1PL.EXCL
 ‘One day, all the men came together (= “*having assembled* [and] *gathered*”) and thought “How will we kill that demon?”’ (= ‘*kill* [and] *cause to fall*’) (MT, 1:66)

Alternating stems. This group consists of a large number of lexical morphemes which have been borrowed from Sadani and which have two different forms, one in *-e* for the middle voice and one in *-ay* (= [aɛ] i.e. /a/ + /e/) for the active, with the exception of *badle* ‘change’ (ITR) (middle) / *badli* ‘change’ (TR) (active) and *bu?jhi* ‘understand’ (middle) / *bujhay* ‘explain’ (active). Table 9.7 presents a few examples.

The *-e* in Kharia in both stems perhaps derives from the infinitive marker *-ek* in Sadani, although this derivation is not without problems. At any rate, *-e/-y* here serves to mark the lexeme as a borrowed stem.³²

TABLE 9.7: THE *-EI-AY* ALTERNATION

Middle stem	Active stem
<i>bagre</i> ‘be(come) bad’	<i>bagray</i> ‘spoil (TR), ruin’
<i>cale</i> ‘function; go’	<i>calay</i> ‘drive (a car, etc.)’
<i>chu?te</i> ‘leave’ (ITR)	<i>chu?tay</i> ‘leave’ (TR)
<i>khole</i> ‘open’ (ITR)	<i>kholay</i> ‘open’ (TR)
<i>salge</i> ‘burn, catch fire’	<i>salgay</i> ‘light a fire’

The status of *-a* in Kharia is more problematic. In Sadani, the suffix *-ā* is, among other things, a causative marker, although its status there is still somewhat unclear.³³ In Kharia, *-a* is not productive outside of these lexemes.

While all of these lexemes are direct borrowings from Sadani, many lexemes marked by *-ay* in Kharia in fact derive from verbs in Sadani which are not marked there by *-ā* with the same meaning. Cf. Kharia *karay* ‘do’ but Sadani *kar-* ‘do’, and Kharia *socay* ‘think’ but Sadani *soc-* ‘think’.³⁴

It would be incorrect to assume that *-a* is simply a causative marker or transitivizer in Kharia: *-a* is also found in a number of lexemes which can appear only in the middle or active, but not in both, such as *khisay* ‘become angry’ (middle only) or *chikay* ‘sneeze’ (active only). These forms do not have a counterpart marked in *-e*, and neither fits the usual description of a ‘causative verb’.

The exact status of *-a* in Kharia is dealt with in more detail in Peterson (2006). For our purposes it will suffice to note that it appears to be related to animacy, whereas predicating lexemes without the *-a* (i.e. those in *-e*) generally have inanimate subjects.

3.2.1 Subject

Kharia is a nominative/accusative type language in terms of predicate agreement. The predicate generally (but not always, see below and section 5) agrees with the ‘subject NP’ (= S and A) in terms of person, number and honorific status. Table 9.8 gives an overview of subject-markers on the predicate.

As the following examples show, the predicate usually agrees with the ‘subject NP’, which is either S or A. That is, Kharia shows nominative/accusative alignment:

- (72) *la?* *sou?b=ga* *bhai=ki* *ho* *khajar* *tar=na* *ghaq* *juda*
 then all=FOC brother=PL that deer kill=INF PURP separate
 S
juda *mu?ki=may,* *kinir=te.*
 REP emerge=M.PST=3PL forest=OBL
PREDICATE

‘Then all the brothers set out separately to kill that deer, into the forest.’ (AK, 1:15)

- (73) *muda* *ha?qo=ki* *kha?iya* *so?j=te=may,* *adha=ki* *kha?iya*
 but half=PL Kharia understand=A.PRS=3PL half=PL Kharia
 A O **PREDICATE₁** A O
umay *so?j=te.*
 NEG.3PL understand=A.PRS
PREDICATE₂

‘But half understand Kharia, half don’t understand Kharia.’ (AK, 5:6)

TABLE 9.8: PERSON/NUMBER/HONORIFIC MARKING ON THE PREDICATE

	Singular	Dual (/HON)		Plural	
		INCL	EXCL	INCL	EXCL
1	=(<i>i</i>) <i>n</i> (less commonly: =(<i>i</i>) <i>ŋ</i>)	= <i>naŋ</i>	= <i>jar</i>	= <i>niŋ</i>	= <i>le</i>
2	=(<i>e</i>) <i>m</i>	= <i>bar</i>		= <i>pe</i>	
3	∅	= <i>kiyar</i>		= <i>ki</i> / = <i>may</i>	

The third person plural marker =*may* is generally found after the middle past marker =*ki*, thus avoiding the sequence =*ki*=*ki* ‘=M.PST=3PL’, although this sequence is occasionally found. =*ki* is generally used elsewhere although it is often replaced by =*may* in these environments as well, especially after the perfect marker =*si*?. Thus =*may* and =*ki* are identical in meaning when they mark a third person plural subject: The choice of one over the other in these environments is largely speaker-specific; compare, for example, the following example, where the two appear side-by-side:

- (74) *tay lam=na lamna lamna ho=ki ho jinis=te, khajar,*
 then search=INF REP REP that=PL that animal=OBLQ deer
kuy=oʔ=ki ro tar=oʔ=may
 find=A.PST=PL and kill=A.PST=3PL
 ‘Then searching, searching, searching they *found* that animal, a deer, and
killed [it]’ (AK, 1:53)

The only environment in which they are not in free distribution is with the negative morpheme *um*: When a clause with a third person plural subject is negated, only *umay* (< **um=may* ‘NEG=3PL’) may be used, never **um=ki* (section 3.2.9).

Note that, with the exception of =*may* ‘=3PL’, the agreement markers of the third persons are, properly speaking, number markers and are identical with number marking on the NP. For example, *lebu*(=∅) ‘man, person’, *lebu=kiyar* ‘men, people’ (DL) and *lebu=ki* ‘men, people’ (PL) (section 3.1.1). Hence I gloss these markers on the predicate as SG(Singular), D(ua)L and PL(ural), respectively, but do not gloss them for person.

‘Subject NP’ or incorporated pronominal subject? As we shall see below in sections 3.2.9 and 5, there is reason for considering the person marking on the predicate to be the subject and the overt ‘subject NP’ to be in apposition to this form. First, explicit mention of the ‘subject NP’ is generally only found when it is contrasted with another entity, when it is a new topic, or when its identity is not clear from the context, especially in the third persons. ‘Subject-less’ clauses abound in my texts: Here I will only refer the reader to example (180), in which the ‘object NP’ is also not explicitly mentioned. Thus, explicit mention of the ‘subject NP’ (or other arguments) is pragmatically motivated.

In addition, the person marker enjoys a certain amount of ‘freedom of movement’ in that it attaches in negation to the negative marker, leaving behind a semi-finite predicate form (section 3.2.9), for example, *koy=t[e]=jn* (KNOW=A.PRS=1SG) ‘I know’ but *um=jn koy=te* (NEG=1SG KNOW=A.PRS) ‘I don’t know.’

And finally, in section 5 we will see that Kharia has ‘semantic agreement’ in that the formal categories of person and number need not be the same for the ‘subject NP’ and the person marker on the predicate in cases such as ‘those of you’, ‘who (of you)?’, etc., where the predicate is marked as second person, plural, although the ‘subject NP’ is the third person (see e.g. examples (214)–(215)).

All this strongly suggests that person marking on the predicate is a kind of incorporated pronominal and that the overtly mentioned ‘subject NP’ is in apposition to this.

3.2.2 Object types

There are two types of subcategorized elements in Kharia which are not subjects: (direct) objects and oblique arguments, which I will refer to here simply as obliques.

The object is a clearly defined category in Kharia: When definite it appears in the oblique case, marked by =*te*. When it is indefinite, it appears in the unmarked form (= direct case). This category corresponds to O or P in typological studies. However, unlike North Munda languages or Juang, Sora or Gorum, Kharia does not encode features of the object in the predicate itself.

The second group, the obliques, is generally marked by =*te*, but not consistently, although the rules governing this marking are not altogether clear as yet. In a sense, these may be considered ‘obligatory adjuncts’, similar, for example, to the obligatory locative with predicates such as *reside* or *put* in English. Semantically, this group includes beneficiaries (e.g. with *ter* ‘give’) and locatives (e.g. with *maṛay* ‘place’), and perhaps others.

Although the oblique marker =*te* is used to mark both definite objects on the one hand and obliques (and also adjuncts) on the other, there is at least one difference between these two groups which justifies this distinction: Only the object can become the subject of the passive, obliques and adjuncts cannot. Thus the morphosyntactic behaviour of a recipient with *ter* ‘give’ is virtually identical to that of a locative adverbial (e.g. with *ḍoko* ‘sit down’), except that it is obligatory or rather, its presence is obligatorily assumed even if it is not overtly expressed: An oblique or adjunct may not be promoted to subject in the passive ((76) and (79)), whereas an object may (78).

(75) Active: *ho-kaṛ* *u* *ṭhāṛo=te* *ḍoko=ki*.
 that-SG.HUM this place=OBLQ sit.down=M.PST
 ‘She/he sat down in this place.’

(76) Passive: **u* *ṭhāṛo* *ḍoko* *ḍom=ki*.
 this place sit.down PASS=M.PST
 ‘This place was sat on.’

(77) Active: *ho-kaṛ* *am=te* *pothi=te* *ter=o?*.
 that-SG.HUM 2S=OBLQ book=OBLQ give=A.PST
 ‘She/he gave you the book.’

(78) Passive: *am=te* *pothi* *ter* *ḍom=ki*.
 2S=OBLQ book give PASS=M.PST
 ‘A book was given to you.’

(79) Passive: **am* *pothi=te* *ter* *ḍom=ki=m*.
 2SG book=OBLQ give PASS=M.PST=1SG
 ‘You were given the book.’

When the object of the predicate is known from context, it need not be explicitly mentioned. Similar to the ‘subject NP’ discussed in section 3.2.1, the object is only explicitly mentioned if it is focused, a new topic, or not recoverable from context. Some examples of this in the examples in this text are (85), (120) (in which the recipient or ‘indirect object’ is also not explicitly mentioned) and (180).

3.2.3–3.2.5 *Tense, aspect and mood*

Tense, aspect and mood in Kharia are expressed in a predicate-final, complex grammatical marker. This marker begins with the marking for the perfect, followed by portmanteau markers which express both the basic TAM categories and basic voice (active vs. middle) and finally, markers for person, number and honorific status. With the exception of the optative, all of these markers are enclitic, whereas the optative marker is a phonological word and can be separated from the unit before it, for example, by the enclitic focal particle =*ga*.

The basic TAM/voice categories. Almost all morphologically fully finite predicates in Kharia are marked as either active or middle, while non-finite forms are neutral with respect to these two categories. These two categories are marked by portmanteau morphemes which simultaneously express four of the seven basic TAM categories.

The remaining three categories, the ‘Past II’, perfect and optative, are neutral with respect to basic voice.³⁵ These forms are presented in Table 9.9.

Past and ‘Past II’. The past and ‘Past II’ are used for events which have already occurred at the time of the speech-act. They are most common in narratives in chains of events.

- (80) *rel chu?te=na thoŋ pōga baje=ki la? bhe?i*
 train leave=INF PURP horn be.sounded=M.PST then sheep
merom=ki ho-ti?j=ga col=ki=may.
 goat=PL that-side=FOC go=M.PST=3PL
 ‘The horn sounded for the train to leave, so the sheep and goats went in that direction.’ (RD, 1:11)

The ‘Past II’ derives from the ending of the past perfect, =*sikh=o?*, which always appears in the active.³⁶ It does not seem to differ at all semantically from the simple past, although the active/middle opposition is neutralized in this category, as the form is always =*kho?*.³⁷ Its use is considered incorrect, although it is very common

TABLE 9.9: THE BASIC TAM/VOICE CATEGORIES

	Middle	Active
Past (PST, aspectually neutral)	= <i>ki</i>	= <i>o?</i>
Present General Imperfective (PRS)	= <i>ta</i>	= <i>te</i>
Present Progressive (PROG)	= <i>ta?jd</i>	= <i>te?jd</i>
Irrealis (IRR)	= <i>na</i>	= <i>e</i>
‘Past II’	= <i>kho?</i>	
Perfect	= <i>si?(d)</i>	
Optative	= <i>gu?du? / gu?ru?</i>	

in the spoken language, especially among younger speakers. It appears to be a rather recent innovation.

- (81) *hote kheti uslo? bes bes aw=kho?*
 there field soil good REP COP=PST.II
 ‘The fields and the soil there were good.’ (MT, 1:76)

The ‘Past II’, like the simple ‘past’, is aspectually neutral, in opposition to the periphrastic past imperfective (section 3.2.12). The past imperfective, which is otherwise marked by the middle past marker =*ki*, can also be marked for the ‘Past II’.

- (82) *ho-kaṛ ḍisa? ḍisa? kho?tay=jo merom gupa=na la?=kho?*
 that-SG.HUM far REP up.to=ADD goat guard=INF IPFV=PST.II
 ‘He used to watch over the goats going (= up to) very far.’ (RD, 1:4)

The present. The present (general) imperfective is used for actions which occur on a regular basis, hold at the moment of utterance, as a ‘historical present’ or, very rarely, for imminent future actions. In the following, the speaker is telling us about the cooking habits of the Kharia:

- (83) *khoṛi poḍda bo?=ki=te tam jou khaṛiya=ki*
 village.section village place=PL=OBLQ now up.to Kharia=PL
kaḍom cakhna hinte saṅsaṅ umay may=te. ho=ki
 fish curry LOC turmeric NEG.3PL mix=A.PRS that=PL
tenton=ga may=te=ki.
 tamarind=FOC mix=A.PRS=PL
 ‘Up to the present day, in the villages and village sections, the Kharia do not mix turmeric into fish curries. They mix in tamarind.’ (MT, 1:115f.)

In its use as a ‘historical present’, the present often alternates freely with the past tense:

- (84) *la? porha=jo col=ki ro yo=te ho maṛa bo?te ḍam=ki*
 then Porha=ADD go=M.PST and see=A.PRS that cave LOC arrive=M.PST
ḍaḍa?te yo=yo?
 water=OBLQ see=A.PST
 ‘Then Porha also went and looks, [he] arrived at the cave [and] saw the water.’ [AK, 1:39]

As mentioned above, the use of the present tense for future actions is quite seldom:

- (85) *kongher rag badli kon gam=o? ‘ga?j=chi?q=iṅ, yo,*
 boy voice change SEQ say=A.PST cast=PRF=1SG mother
ol kay=t[e]=iṅ?’
 bring BEN=A.PRS=1SG
 ‘The boy, changing his voice, said “I have thrown [down a loaf of bread], mother, shall I bring [some more] for [you]?”’ (BB, 1:80)

It could thus be argued that what I term here the present (general imperfective) is more appropriately termed ‘nonpast’, as I have referred to this in previous studies (Peterson 2002, 2007, in press). However, as this function of the present is so seldom and the use of the irrealis so common for future time, I prefer to refer to it now as

the present tense and examples such as (85) must then be considered ‘exceptions’ which can be explained by the fact that they are considered imminent, as with the present progressive (see below).

The present progressive. The progressive is used for events which are in progress at the time of utterance (86) or are imminent (87). The progressive consists of the present marker =*te*/*ta* plus what appears to be a marker deriving from the present copula *ayi²jq*: *ayi²jq* is realized as *ayi²j* in all persons other than the first and second persons, singular. Similarly, the progressive markers =*te²jq* (active) and =*ta²jq* (middle) are realized as =*te²j* and =*ta²j* in these environments, respectively.

- (86) *ijn=te musa jeri kosu=te²j ro kulqa?=jo.*
 1SG=OBLQ today body get.sick-<CAUS>=A.PROG and fever=ADD
jab=si?
 catch=PRF
 ‘Today [my] body is hurting me and fever has also taken hold of (= caught) me.’ (BB, 2:23)

- (87) (The witch said:) *ijn gone cokhay ke qel=ta²jq=ijn=ga.*
 1SG tooth sharpen SEQ come=M.PROG=1SG=FOC
 ‘I’ll come right back after I sharpen my teeth.’ (BB, 1:65)

The progressive is restricted to ‘present’ contexts. For past, future and hypothetical progressive meanings, the periphrastic imperfective is used (section 3.2.12).

Irrealis. The irrealis is used for future actions and states (88), in the imperative (89), and in conditional and counterfactual clauses (section 4.5).

- (88) *tobqa? qa?=te i-ghay uq=e=niŋ?*
 mud water=OBLQ what-way drink=A.IRR=1PL.INCL
 ‘How will we drink muddy water?’ (MS, 1:126)

- (89) *kulu gam=o? ‘baru kayom! ijn=te=jo am=bar=a?*
 turtle say=A.PST good speech 1SG=OBLQ=ADD 2=2DL=GEN
sori qoŋ=e=bar!’
 with take=A.IRR=2DL
 ‘The turtle said “Great idea! Take me along with you as well!”’ (TK, 1:11)

Perfect. The present perfect denotes that an event has occurred which is of direct relevance to a later situation, usually – but certainly not always – the moment of the speech act. For example, in the following story, nine sons have been sent out by their father to hunt down a certain animal when suddenly one of them, Porha, realizes that they do not know which animal they are supposed to hunt – none of them can remember what their father has said. Thus, the fact that the father may have said that they should hunt down a deer is of direct relevance to the moment of speech (although English does not allow the perfect here). The present perfect is marked for neither tense nor active/middle.

- (90) *poŋha gam=o? no ‘sahayda, ikon, khajar, hiran=te,*
 Porha say=A.PST COMP maybe umh deer deer=OBL
tar=na gam=si? hoy’.
 kill=INF say=PRF INFER
 ‘Porha said “Maybe, umh, he said to kill a deer.”’ (AK, 1:13)

The past perfect is found only in the active. It usually designates a past event which preceded another event in the past.³⁸

- (91) *am=pe, iŋ gam=sikh=o²j ho-ghay=ga am=pe col=ki=pe.*
 2=2PL 1SG say=PRF=A.PST.1SG that-way=FOC 2=2PL go=M.PST=2PL
 ‘You, you went just as I had told you.’ (AK, 1:63)

The past perfect may also be used as a simple past tense. From this usage, the ‘Past II’ discussed above has developed.

- (92) *ho-kaŋ ho jara? daru=ya? khoŋri=te murti=te*
 that-SG.HUM that banyan tree=GEN hollow=OBLQ statue=OBLQ
ko<?>sor=na un=sikh=o².
 dry-<CAUS>=INF place=PRF=A.PST
 ‘He placed the statues in the hollow of that banyan tree to dry [them].’
 (AK,3:13)

For examples of the irrealis perfect, see (205) and (215).

Optative. The optative is used to denote a strong wish, obligation or a third person imperative. It may not be used in conjunction with first or second persons. Unlike all other basic TAM categories, *guŋu?* / *guḍu?* may be separated from the stem by the focal particle =*ga* and is hence a separate word. However, it has no independent lexical meaning.

- (93) *aniŋ=a² konsel kongher=ki khaŋiya=ki=ya? bair nāḍani=te*
 1PL.INC=GEN girl boy=PL kharia=PL=GEN old history=OBLQ
koŋ guŋu?²=may ro khaŋiya=ki=ya? paŋom ro main=te
 know OPT=3PL and Kharia=PL=GEN strength and honour=OBLQ
koŋ kon oḍo? meson ho-je? arjay guŋu?²=may.
 know SEQ once again that-SG.NHUM revive OPT=3PL
 ‘Our girls and boys should know the old history of the Kharia and, having learned the honour and strength of the Kharia, they should revive it once again.’ (Kerkeṭṭā 1990:ii)

3.2.6 Orientation/directionality

Kharia has no general marker of directionality on the predicate. However, motion towards the deictic centre of the narrative can be expressed by using a predicate whose second lexical stem is *ḍel* ‘come’ as a kind of ventive.

- (94) ... *ro botoŋ go²ḍ=ki ro botoŋ ke eŋ ḍel=ki.*
 and fear C:TEL=M.PST and fear SEQ return come=M.PST
 ‘... and he became afraid, and, having become afraid, he returned.’
 (AK, 1:36)

There is no general andative category in Kharia indicating motion away from a deictic center.³⁹

3.2.7 VoicelVersion

Kharia has two basic voices – active and middle. Most predicates which appear in the active are transitive and most which appear in the middle are intransitive, that is,

TABLE 9.10: PREDICATING LEXEMES WHICH APPEAR IN BOTH THE MIDDLE AND ACTIVE

	Middle	Active
<i>ga?</i>	'rip (of paper)' (ITR)	'rip' (TR)
<i>pa?</i>	'break' (ITR)	'break' (TR)
<i>ru?</i>	'open' (ITR)	'open' (TR)
<i>sului</i>	'warm up' (ITR)	'warm up' (TR)

lexemes which can appear in either voice are generally transitive in the active and intransitive in the middle. Table 9.10 presents a few examples.

There are a number of other criteria which combine to determine basic voice. These include aspect/*Aktionsart*, 'thoroughness', and volitionality, among others.

Aspect/Aktionsart. When the stem is a (potentially) free-standing form (section 3.2.8), the predicate obligatorily appears in the middle voice⁴⁰ and denotes an action that occurs over and over for a long time or, with telic predicates, takes a very long time to happen. In the past, this form can also denote a remote past and with respect to future time either a distant future or an event whose time of action is uncertain. My data strongly suggest that any lexeme can participate in this alternation, provided that a proper context can be found. For illustration, here are a few predicates which in the unmarked construction always appear in the active voice.

- (95) *Unmarked form (active)* *Marked form (middle)*
absi²b 'begin' *absi²b* 'take a long time beginning'; 'began a long time ago'
bu? 'beat (a drum)' *bu?-bu?* 'beat (a drum) regularly (e.g. as a drummer)'
hakɽe 'grunt (of oxen)' *hakɽe* 'grunt a long time ago'
lam 'search for' *lam-lam* 'search for a long time; hunt'

Example: *bi²d*

- (96) *ij* *da?* *biɽh=o²j*. (*bi²d* > *biɽh* before =*o?*) *Unmarked construction*
 1SG water pour.out=A.PST.1SG
 'I poured the water out.'
- (97) *ij* *da?* *bi²d-bi²d=ki=j*. *Marked construction*
 1SG water pour.out-REDPL=M.PST=1SG
 'I poured the water out over and over (e.g. that was my job, so I did it constantly).'

'*Reflexivity*'. With a number of predicates, the use of the middle denotes an action which the subject either does to him-/herself or which she/he performs on him-/herself. Table 9.11 provides a few examples.

This criterion is closely related to valency, in that the 'reflexive' form is often intransitive. However, as *gujan* (as well as several others in this class) is also transitive in the middle voice, this category must be considered distinct from that of valency. '*Thoroughness*'. In this somewhat heterogeneous class, predicates are in some way less thorough when they appear in the middle than when they appear in the active. Table 9.12 gives a few examples.

TABLE 9.11: 'REFLEXIVITY' AND BASIC VOICE

	Middle	Active
<i>gujuŋ</i>	'wash one's own feet' (TR)	'wash someone else's feet' (TR)
<i>khuray</i>	'shave oneself'	'shave someone else'
<i>uwa?</i>	'bathe' (ITR)	'bathe' (TR)

TABLE 9.12: 'THOROUGHNESS' AND BASIC VOICE

	Middle	Active
<i>ju²d</i>	'sprout roots' (ITR)	'take root firmly' (ITR)
<i>kanⁱ'j</i>	'believe, have faith in (more or less)'	'believe in (strongly)'
<i>lebui</i>	'love (somewhat)'	'love (strongly)'

TABLE 9.13: SUDDENNESS/(NON-)VOLITIONALITY AND BASIC VOICE

	Middle	Active
<i>kui</i>	'stumble upon'; 'appear', 'be found'	'(look for and) find'
<i>malum</i>	'come to know accidentally'	'find out (by searching)'
<i>onɔor</i>	'hear accidentally'; 'be heard'	'(listen and) hear'
<i>yo</i>	'catch a glimpse of'; 'be seen'	'(look and) see'

Suddenness/(Non-)volitionality. There are a number of predicates which are not only intransitive when they appear in the middle, but which also denote that an action occurred suddenly and strongly imply that it was also performed accidentally, as opposed to the same predicate in the active, which is transitive, not sudden and usually agentive. Table 9.13 provides a few examples. Following this are two examples to illustrate this, both from the same story.

Middle ⇒ *Intransitive / Nonvolitional / Sudden.* In the following story, the soldiers are actually searching for a camp-fire while hunting and quite unexpectedly stumble upon a castle in the middle of the forest:

- (98) *jab ho=ki lam=na col=ki=may la? ho=ki=te kinir*
 CR:TEMP that=PL seek=INF GO=M.PST=3PL then that=PL=OBLQ forest
mo?jhi=te moŋ koleŋ=a? mohol kui=ki.
 middle=OBLQ I king=GEN castle find=M.PST
 'When they went to look for [for the source of the smoke], they [unexpectedly] found a king's castle in the middle of the forest.' (Pinnow 1965a:40)
 [literally: 'a king's castle appeared to them.']

Active ⇒ *Transitive/Volitional/Not sudden.* Taking another example from Pinnow as follows:

- (99) ... *la? iŋ=te alsɪ lam=na buŋ saŋghar=e=pe oɔo? ber*
 then 1SG=OBLQ axe seek=INF INS help=A.IRR=2PL and who
kuy=e ho-kaŋ=te iŋ=a? beŋi=ŋ=te
 find=A.IRR that-SG.HUM=OBLQ 1SG=GEN daughter=1SG=OBLQ

ter[=e]=jn.
 give[=A.IRR]=1SG
 ‘... then [please] help me with looking for my axe, and whoever finds [it, I]
 will give him my daughter [in marriage]’ (Pinnow 1965a:39)

There are also a number of other uses, some of which are more or less speaker-specific, which we cannot go into here. See Peterson (2006) for a more detailed discussion.

3.2.8 (Non-)finiteness

There are both non-finite and semi-finite predicates in Kharia. Fully finite predicates are marked for PERS/NUM/HON, whereas non-finite and semi-finite forms lack this marking.

Semi-finite forms. Semi-finite forms may be considered predicates with multiple stems but with a slight difference to the standard case. Normally, the various stems of a predicate with multiple stems are directly juxtaposed to one another with no intervening inflectional marking, as in (100).

- (100) *tay ađi=te=jo batay=o? no ‘dhaỹ[=e] am, han*
 then ANAPH=OBLQ=ADD say=A.PST COMP hurry=A.IRR 2SG that
mařa bo?te đa[?] [a]yi?j. uđ qe=na=m.’
 cave LOC⁴¹ water PRS.COP drink come=M.IRR=2SG
 ‘Then [he] said to him, too, “You hurry, at the cave there is water. *Drink* [and]
come [back].”’ (AK, 1:38)

Similarly, with semi-finite forms, we find a single predicate with multiple stems, but here each stem contains a certain amount of inflectional marking without being fully finite, except for the final element, which is fully finite (101). The two lexical stems may or may not be joined by a conjunction such as *ro* or *ođo*², both of which mean ‘and’.

- (101) *konon no maha jhaři=ga tomne tomne lutui ikuđ*
 small or big all=FOC new REP clothes very
kelom-bo? su=yo? aňku=yo?=ki.
 beautiful-INTENS put.ON=A.PST wear.a.chaddar=A.PST=PL
 ‘Whether big or small, all *put on* [and] *wrapped around* [themselves] the very
 beautiful new clothes.’ (Pinnow 1965a:75)

This type of semi-finiteness is highly restricted and is typical of certain fixed expressions. It is much more common in the past tense, where it is more or less obligatory for certain combinations, than in, for example, the irrealis, and at least in the modern language is restricted to the active voice.⁴² It is especially common with certain everyday actions which tend to be done together, such as eating and drinking. Consider, for example, the following, adapted from Roy and Roy (1937:180f.).

- (102) *u=kiyar taň el=a? sori=ga nog=e uđ=e=kiyar.*
 this=DL now 1PL.EXCL=GEN with=FOC eat=A.IRR drink=A.IRR=DL
 ‘They two will now eat and drink with us.’

Non-finite forms & Free-standing forms. Every finite lexical predicate has a corresponding form which I refer to as the free-standing form. It lacks all TAM

and person marking, although it retains valency-related marking such as causative and passive/reflexive marking. It may govern an object and adjuncts, which take the same marking they would have with a finite predicate, but what corresponds to the subject of the finite predicate appears in the genitive. The free-standing form may also mark for case and number. This form is used in various types of subordination (section 4.6) but is most commonly found in attributive function, corresponding more or less to relative clauses (section 4.4).

The free-standing form is formed as follows: Polysyllabic stems (= root and derivational markers) undergo no changes, while monosyllabic stems are reduplicated. Thus this type of reduplication is purely phonologically motivated. Some examples:

- (103) *Simple root* *Free-standing form*
borol ‘live’ *borol*
ru? ‘open’ *ru?-ru?*
yo ‘see’ *yo-yo*

As noted, the ‘subject’ of the free-standing form appears in the genitive:

- (104) *lekin* *kharīya* *laɟ* *iswar=a?* *ter-ter* *heke*.
 but Kharia language Lord=GEN give-REDPL PRS.COP
 ‘But the Kharia language has been given [to us] by God (= is a *God’s giving*).’ (MS, 1:267)

Free-standing forms have no inherent orientation with respect to subject, object, instrument, locative, etc., nor with respect to tense. Also note that *ɟn=a?* in (105) appears in the genitive as it is the ‘subject’ of the free-standing form, not because of what would appear to be the head nominal (*lebu*). The structure of (105) is thus parallel to that of (104).

- (105) *ɟn=a?* *yo-yo* **lebu**
 1SG=GEN see-REDPL person
 ‘**the person** *I saw/see/will see*’
- (106) *ɟn=te* *yo-yo* **lebu**
 1SG=OBLQ see-REDPL person
 ‘**the person** *who saw/sees/will see me.*’
- (107) *ɟn=a?* *dura=te* *ru?-ru?* **kupji**
 1SG=GEN door=OBLQ open-REDPL key
 ‘**the key** *I opened/open/will open the door with*’

Unlike the infinitive, which is discussed in the following section, the free-standing form is not used with auxiliaries.

The semantics of the free-standing form are entirely predictable from the basic meaning of the lexical morpheme: For lexical morphemes which denote an event, the meaning of the free-standing form is the act of performing this event; with change-of-state lexemes, it denotes the property itself.

Finally, recall that the potentially free-standing form may also serve as the stem of a finite predicate (section 3.2.7).

Infinitives. Kharia is one of the few Munda languages – perhaps the only one – which possess a general infinitive form, marked by =*na*. What corresponds to the subject of a finite predicate appears in the genitive (108), while objects of an infinitive have the

same marking as with finite predicates (109). The infinitive may serve as a nominal, as in (108) or in forms such as *noʔ=na* ‘food’ (literally: ‘(to) eat’), *uʔq=na* ‘drink’, in attribution (109), and is used with a number of auxiliaries (section 3.2.12).

- (108) *am=aʔ jib=na=te thik um hoy=ki.*
 2SG=GEN touch=INF=OBLQ good NEG become=M.PST
 ‘Through your touching [it, it] has become spoiled (= did not become good).’
 (MS, 1:247)

- (109) *musnij u-ghay hoy=ki no doli=te dʊʊq=na*
 one.day this-way happen=M.PST COMP sedan=OBLQ take=INF
bhere u kharīya pahan=te haḍa laʔ=ki.
 time this Kharia priest=OBLQ pee EMOT=M.PST
 ‘One day it happened this way that, **at the time** [*they were*] to take away the sedan chair, the Kharia priest had to pee.’ (AK, 2:8)

Sequential converbs (‘*Conjunctive participles*’). There are three sequential converb markers in Kharia: *ker*, *ke* and *kon*. These forms are traditionally referred to as ‘conjunctive participles’ in South Asian linguistics. The first two are direct borrowings from Sadani (See Jordan-Horstmann 1969:96f.).

kon, the more common form, appears to be a calque of the Sadani form *-ker*: Like the cognate form *-kar* in Hindi, *-ker* in Sadani appears to derive from the root *kar*-‘do’. Similarly, the sequential converbal marker *kon* in Kharia apparently derives from *ikon* ‘make, do’.

The core function of these forms is to denote the completion of one action before another begins.

- (110) ... *kaʔ kom=ki ḍhoʔ ke muʔ goʔq=ki=may.*
 bow arrow=PL grab SEQ set.off C:TEL=M.PST=3PL
 ‘... they took their bows and arrows and set off (i.e. “*having grabbed* their bows ...”).’ (AK, 1:11)

These converbs can also be used to denote the manner in which an action is carried out. In the following example, the action denoted by the converb (with two stems) effectively amounts to that of the morphologically finite predicate and the converb simply denotes the manner in which this action was carried out.

- (111) ... *lay koj kon goʔjuj bay=siʔ=may.*
 dig scrape SEQ path make=PRF=3PL
 ‘...they have built the path *by digging and scraping* [the dirt away].’ (MT, 1:45)

Imperfective converbs. There are three constructions which function as imperfective converbs. To the stem is added one of the following forms:

- a form which is homophonous with the present general imperfective, middle marker =*ta*
- the focal marker =*ga*
- the infinitival marker =*na*

This form is then reduplicated. Only =*ta* may unambiguously be considered an imperfective converb marker, whereas the others primarily fulfil other functions.⁴³

As the following examples show, in addition to simultaneity, the imperfective converbs can also express an additional ‘adverbial’ meaning, generally that of manner:

- (112) *ele am=pe=te goʔ=ta goʔta han-tiʔj u-tiʔj*
 1PL.EXCL 2=2PL=OBLQ carry.on.shoulders=CV REP that-side this-side
qor=e=le.
 take=A.IRR=1PL.EX
 ‘We will carry you around on our shoulders (= *carrying*, we will take).’
 (AK, 2:34)

- (113) *muda ekle aw=ga awga muruk ansa laʔ=ki.*
 but alone COP=FOC REP very unhappiness EMOT=M.PST
 ‘But [He (= God)] became unhappy being alone (= *being* alone, he became unhappy).’ (AK, 3:6)

- (114) ... *lekin lam=na lamna souʔb=te ikudʒ jughay qaʔ piyas*
 but search=INF REP all=OBLQ very much water thirst
laʔ=ki.
 EMOT=M.PST
 ‘But *searching and searching*, [they] all became very thirsty.’ (AK, 1:17)

Participles. There are three participial markers in Kharia, two of which have been borrowed from Indo-Aryan. They are only used in nominal attribution (section 4.4).

- The first ‘participle’ consists of the simple predicate stem plus the ending *-quʔ*, which has a number of other functions.⁴⁴ It has no inherent temporal orientation and may refer to past, present or future events.
- The second participle has been borrowed from Hindi and consists of the infinitive in *=na* followed by the form *=wala*. This form is used to denote iterativity and habituality.
- Finally, there is the suffix *-l*, which attaches only to lexical predicates of Sadani origin which end in *-a* or *-ay*. This suffix has simply been borrowed along with the root. For example, *ʔayay* ‘hang (TR)’ *ʔaya-l* ‘hung’, *biha* ‘marry’ *biha-l* ‘married’. It seems to be restricted to past events only and is not acceptable to all speakers.

3.2.9 Negation

Non-modal (i.e. indicative) sentential negation is expressed through the negative particle *um*. When the predicate is negated, PERS/NUM/HON marking attaches to the negative morpheme *um*, which precedes the now semi-finite predicate, as Box 9.8 shows. The negative morpheme *um* is not a verb: It cannot mark for any of the TAM categories given in Box 9.8. Hence, NEG=PERS/NUM/HON should not be mistaken for a negative auxiliary.

BOX 9.8: A SCHEMATIC OVERVIEW OF THE NEGATED MORPHOLOGICALLY FULLY FINITE PREDICATE IN KHARIA

NEG=PERS/NUM/HON RECIP CAUS-**Lexeme** -<CAUS> AKT BEN PASS AUTOPOES TEL=PRF=TAM/VOICE

- (115) *ter*[=*e*]=*ijn* *um*=*ijn* *ter*=*e*
 give=A.IRR=1SG NEG=1SG give=A.IRR
 'I will give.' 'I will not give.'

The only (partial) exception to this is the second person, singular, where person marking may either attach to the predicate or to *um*. The predicate-final position of person marking is by far the more common of the two in the second person, singular. (cf. from Malhotra 1982:285):

- (116) *ubhroy* *um*=*em* *qe*=*na* or *ubhroy* *um* *qe*=*na*=*m*.
 these.days NEG=2S come=M.IRR these.days NEG come=M.IRR=2S
 'These days you do not come.'

Recall also from section 3.2.1 that the form **um*=*ki* 'NEG=PL' is ungrammatical for the third person, plural. Instead, only the form *umay* from **um*=*may* 'NEG=3PL' is found.

- (117) *juj*=*o*?=*may* *umay* *juj*=*o*? *kuy*=*o*?=*ki* *umay* *kuy*=*o*?
 ask=A.PST=3PL NEG.3PL ask=A.PST find=A.PST=PL NEG.3PL find=A.PST
 'they asked' 'they didn't ask' 'they found' 'they didn't find'

As noted in section 3.2.1 above, this (highly limited) freedom of movement is in line with the analysis of the person marker on the predicate as the subject of the clause, a kind of incorporated pronominal, while the overt NP is actually in apposition to this subject. See also section 5.

To negate the imperative or the optative, the form *abu* is used. In the second person, it marks for person and number and has the forms *abu* '2SG', *a?bar* '2DL/2HON' and *a?pe* '2PL'. The form *abu* is also used for the third person but may not occur with a first person subject.

- (118) *muda* *u* *gharana*=? *jimi* *u²bne* *to?* *purlu²d* *purlu²d*
 but this family=GEN name so.much day spotless REP
ayi²j *je?* *abu* *koqil* *gu?u?*
 PRS.COP SO NEG.MOD stain OPT
 'But this family's name has been spotless for such a long time, so [he] should not stain [it].' (Kerkeṭṭā 1990:12)

For many speakers the form is always *abu* in the second person as well and PERS/NUM/HON marking appears at the end of the predicate. Compare the following two examples, where the second person is marked on the negative particle (119) or predicate-finally (120).

- (119) *kay*=*e*=*bar!* *kimin* *kūṛu?*, *a?*=*bar*=*ga* *ro?*
 lift=A.IRR=2HON daughter.in.law child NEG.MOD=2HON=FOC drop
melay=*e!*
 leave=A.IRR
 'Pick [them] up (HON)! Daughter-in-law, don't drop [and] leave [any]!' (MT, 1:42)
- (120) *i* *lam*=*te*=*m?* *la?*, *mane*, *ijn* *ter*[=*e*]=*ijn* *lekin*
 what want=A.PRS=2SG then umh 1SG give=A.IRR=1SG but
ijn=*te* *jan* *abu* *tar*=*e*=*m*.
 1SG=OBLQ life NEG.MOD kill=A.IRR=2SG

'What do you want? Then, umh, I will give [it to you], but don't kill me.'
(MS, 1:64)

For negation of the copula, see section 4.1.

3.2.10 Derivation

In this section I deal with categories which may loosely be termed derivation, such as the causative. I also include here those categories in Box 9.7 which appear between the stem and the perfect marker, that is, the 'V2s' or 'explicator verbs'.⁴⁵ The status of these 'V2s' as derivational markers is, however, questionable. For example, with the exception of the causative, all other categories discussed in this section are expressed by (grammatical) words, not affixes. Nevertheless, their semantic contribution to the stem is often somewhat unpredictable and many of them are limited with respect to productivity, occurring only with a relatively small number of lexemes. Their status as derivation or inflection will not be dealt with further here.

Causative: //O?B// – Allomorphs: lol, loʔl, lol, loʔbl, lbl, lʔbl, lʔl. The causative increases the valency of the lexical predicate by one by introducing a higher or superordinate agent, that is, causing someone to carry out an action or causing an otherwise nonagentive event to happen. With monosyllabic roots, it appears as a prefix, with disyllabic roots, it is generally an infix. The exact distribution of the various allomorphs is partially idiosyncratic and cannot be presented here in detail, but there are clear tendencies:

- As an infix, the form is either [b] or [ʔb]. Before labials and occasionally elsewhere it is realized as [ʔ].
- The usual form of the prefix is [ʔʔb] before most consonants and [ʔb] before vowels. Some speakers prefer the form [ʔʔ] over [ʔʔb].
- [ʔ] is restricted to a small number of lexemes.

The following presents a few examples:

(121)	Underlying lexeme	Causative stem
	<i>ajoʔd</i> 'dry up (TR)'	<i>a<ʔb>joʔd</i> 'dry up (TR)'
	<i>dɛʔb</i> 'rise; climb'	<i>o-dɛʔb</i> 'raise; offer up, sacrifice'
	<i>nimi</i> 'become named'	<i>ni<ʔ>mi</i> 'name' (TR)
	<i>yar</i> 'flee'	<i>oʔb-yar</i> 'chase away, drive off'

There are also double causatives. In this construction, the second causative marker always appears as a prefix, despite the fact that it is added to a bisyllabic stem.

(122)	Underlying lexeme	Simple causative	Double causative
	<i>aloŋ</i> 'sing'	<i>a<ʔb>loŋ</i> 'have s.o. sing'	<i>ob-a<ʔb>loŋ</i> 'have s.o. make s.o. sing'
	<i>dɛʔb</i> 'rise; climb'	<i>o-dɛʔb</i> 'raise; offer up; sacrifice'	<i>oʔb-dɛʔb</i> 'have s.o. sacrifice'
	<i>lemeʔd</i> 'go to bed'	<i>le<ʔ>meʔd</i> 'put s.o. to bed'	<i>oʔb-le<ʔ>meʔd</i> 'have s.o. put s.o. to bed'
	<i>sore</i> 'become ready'	<i>so<ʔb>re</i> 'prepare'	<i>ob-so<ʔb>re</i> 'have s.o. prepare'

Except for the temporo-aspectual function of the middle voice (section 3.2.7) and unless the predicate is marked for the reciprocal or the passive/reflexive, a predicate marked as causative appears in the active voice.

'Reciprocal': *kol*. The marker *kol* generally has a reciprocal interpretation, that is, two entities mutually affect one another in some manner. However, with many predicates this marker can also denote that the two entities did something together or helped one another, as the following forms show.

- (123) *kol ba²j=ki=kiyar*
 RECIP like=M.PST=DL
 'they (DL) liked each other.'
- (124) *kol aloy=ki=kiyar*
 RECIP sing=M.PST=DL
 'they (DL) sang together.'
- (125) *kol absi²b=ki=kiyar*
 RECIP start=M.PST=DL
 'they (DL) helped each other start.'

The basic meaning of this category appears to be that of mutual affectedness in general, and not merely reciprocity. That is, both (or more) entities equally carry out and are equally affected by an action. *kol* has no independent lexical meaning.

A predicate marked for the reciprocal always appears in the middle voice. With respect to the 'mobility' of *kol*, see 'Semi-productive incorporation' in section 3.2.11.

Passive / Reflexive: qom This category has two main functions:

- It denotes a backgrounding passive, in which the agent of the action is suppressed.
- It denotes reflexivity, that is, the subject's action affects the subject itself.

Thus, the following form can have two very different meanings, depending on the context:

- (126) *yo qom=ki=kiyar.*
 see PASS/REFLXV=M.PST=DL
 'They (DL) were seen (by someone else) / They (DL) saw themselves (e.g. in the mirror).'

This category can also express an 'indirect' reflexive meaning:

- (127) ... *ta²j qom=ki=may komay=te.*
 distribute REFLXV=M.PST=3PL meat=OBLQ
 '... they distributed the meat amongst themselves.' (AK, 1:56)

In the passive use of this morpheme, the undergoer may appear either in the unmarked form (= direct case), in which case it may be considered the 'subject NP' and is cross-referenced on the predicate (128), or it may appear in the oblique case, if definite, in which case the predicate has default marking (third person, singular) (129).⁴⁶

- (128) *pothi=ki ter qom=ki=may.*
 book=PL give PASS=M.PST=PL
 'Books were given.'

- (129) *pothi=ki=te ter dom=ki.*
 book=PL=OBLQ give PASS=M.PST
 ‘The books were given.’

dom has no independent lexical meaning, although it is homophonous with the marker of third person, inalienable possession (section 3.1.3).

‘Durativity’. There are three markers which may all be subsumed under the general heading ‘durativity’. Although there are subtle differences between them, they are largely interchangeable and it is often difficult to distinguish between them semantically.

In two of the following three categories, the semel-iterative and the iterative, the stem may be a (potentially) free-standing form or, less commonly, a simple stem. When the stem is a free-standing form and the predicate is marked by either of these two categories, the predicate appears in the same basic voice as it would in the unmarked form, that is, in the same basic voice as when the stem is a simple lexeme and not a free-standing form. These are the only predicates which may have a stem which is a free-standing form but which need not appear in the middle voice (sections 3.2.7).

None of these three markers has an independent lexical meaning.

Continuative: kan. This category denotes that an action continues for a longer time than expected (‘keeps on’). It is also often used to express that the subject continues to carry on with one action while performing another:

- (130) *ho-kar kayom=ta ro kinbhar=jo jo? kan=te.*
 that-SG.HUM talk=M.PRS and courtyard=ADD sweep CONT=A.PRS
 ‘She talks and keeps sweeping the courtyard.’

With some change-of-state predicates, particularly those of motion but also transitives such as *do²d* ‘take (away)’, the notion of ‘continuation’ denotes that the change is more or less permanent:

- (131) *in am=bar=a? pothi=te do²d kan=o²j.*
 1SG 2=2HON=GEN book=OBLQ take CONT=A.PST.1SG
 ‘I took your book (and I’m not giving it back!).’

Marked for a person other than the first person, this could also mean that the person who took the book lives far away and will not be able to bring it back for some time.

Semel-Iterative: lo? This category is used primarily by the older generation. For those speakers who use this form, it marks an event which occurs over and over, but generally only on one occasion, in contradistinction to *khor* (see below), although these two are often interchangeable.

With *lo?* the stem is often the free-standing form discussed in section 3.2.8, although the use of the simple lexical morpheme instead of the free-standing form is quite common.

- (132) *bero²d lo?=ki=may ro doko lo?=ki=may.*
 stand.up S:ITER=M.PST=3PL and sit.down S:ITER=M.PST=3PL
 ‘They kept on standing up and sitting down again (e.g. at a meeting).’

Iterative: khor. Like *lo?*, *khor* denotes iterativity but generally with an action which comes to a complete stop and then begins again. Here as well, the free-standing

form generally serves as the stem, although the simple root may also be used. Some examples:

- (133) *qubhni=te pe? yo kon kowa? han-ti²j u-ti²j ley-ley*
 cup=OBLQ rice see SEQ crow that-side this-side fly-REDPL
khor=ki muda solo?=te yo kon pe? no?=na um
 ITER=M.PST but dog=OBLQ see SEQ rice eat=INF NEG
are=ki.
 descend=M.PST
 ‘Having seen rice in the cup, the crow kept flying back and forth, but having seen the dog, it didn’t come down to feed.’

According to my data, the use of *ley-ley lo?=ki* here instead of *ley-ley khor=ki* would denote that the crow kept flying back and forth without landing, whereas the use of *khor* makes it clear that the crow landed here and there and then flew again.

The notion ‘here and there’ is quite common with *khor*, primarily with predicates of motion (134) but not exclusively with these predicates (135).

- (134) *u bo?=ki^{A7} ti²j=ga ciŋna=ki buli khor=na*
 this place=PL direction=FOC chick=PL wander ITER=INF
la?=ki=may.
 IPFV=M.PST=3PL
 ‘The chicks were running here and there in all directions.’

- (135) *no?-no? khor=o?*
 eat-REDPL ITER=A.PST
 ‘[She/he] ate a little of this and then a little of that (i.e. here and there from the plate).’

The telicizers. In this section I discuss the three most common telicity markers. ‘Anticipatory telic’ *qo²d* and ‘culminatory telic’ *go²d*. These two markers are largely synonymous and serve to mark an action or change-of-state as being completed. The difference between the two is one of narrative structure.

qo²d (with the allomorph *qor*) indicates that another event follows directly upon the event denoted by the predicate that it marks;

go²d (with the allomorph *gor*) marks a turning point or culmination in a narrative.

To get an idea of the difference between these two markers, consider their use with the lexeme *aw* ‘stay, live’: *aw qo²d=ki=may* (stay A:TEL=M.PST=3PL), with the anticipatory telic marker, would mean ‘they stayed there (and then ...),’ implying that their stay at a particular place was short, as they were on their way to somewhere else. *aw go²d=ki=may*, on the other hand, with the culminatory telic marker, would mean that they settled down somewhere to live, and the narrative would then continue on a different topic. Here is an example for *qo²d*:

- (136) *tay raja jha?i mudh kalo=ki=te ro yahudi*
 then king all chief priest=PL=OBLQ and Jew
jait=a? etoŋdaq=a? guru=ki=te qoklo? rema?
 ethnic.group=GEN receiving.orders=GEN teacher=PL=OBLQ meeting call

dəʔh=oʔ *ro* *ho=ki=te* “*masih=te* *ate* *jorme=na*
 A:TEL=A.PST and that=PL=OBLQ messiah=OBLQ where be.born=INF
ayiʔjʔ” *gam=oʔ* *ro* *juʔ=oʔ*.
 PRS.COP say=A.PST and ask=A.PST

‘The king *called* all the chief priests and teachers of the Jewish people who received orders [from the king] to a meeting and asked (= said and asked) them “Where is the Messiah to be born?”’ (Kullū 1992:2 = Luke, 2:4)

goʔd is much more common than *dəʔd* and signals not only that the action has reached its conclusion, but also that the story is probably also to continue in a different direction.

In the next example, the Kharia priest has dallied too long and another person, a Brahman, has taken his place and is performing the sacrifice. The Kharia priest then finally arrives and sees this. The text reads:

(137) *yo=te* *laʔ* *pujapaʔh=ko* *suru* *goʔd=siʔ*, *absiʔb*
 see=A.PRS then sacrifice=CNTR begin C:TEL=PRF begin
goʔd=siʔ=may.
 C:TEL=PRF=3PL

‘He sees, then, that the sacrifice *has* already *begun*, [i.e. they] *have* already *started* (AK, 2:29)

The story then continues: Seeing that he has been replaced, the Kharia priest then gives up his priesthood entirely and even begins carrying the new priest around in his sedan chair.

Totality: may. *may* denotes the totality of an action or change-of-state with respect to a particular participant in the scenario. For example, in the following example this is with respect to the tea, the object:

(138) *sori* *sori* *caha* *uʔd* *may=oʔ=ki* (**ro* *tam* *jou*
 together REP tea drink TOTAL=A.PST=PL and now up.to
uʔd=teʔj=ki).
 drink=A.PROG=PL

‘They drank up all the tea together (*and are still drinking).’

With transitive activity predicates, the affected entity may also be a plural subject. Here, the clause whose predicate is marked by *may* has a gradual interpretation, that is, ‘one by one’ until all have performed the action:

(139) *ho-kaʔ=te* (*paʔi paʔi* / *muʔu muʔu* / **sori sori*) *gil*
 that-SG.HUM=OBLQ turn REP one.CLSSFR REP together REP hit
may=oʔ=ki.
 TOTAL=A.PST=PL

‘They all hit him (in turn / one after the other / *together).’

With predicates denoting a simple change-of-state or a telic activity, the affected entity is usually either an intransitive subject or an object. If the affected entity is plural, *may* denotes that all members of the group have been affected. If the predicate is scalar, *may* then has the interpretation ‘really, very, completely’. Both a gradual and a simultaneous interpretation are then possible.

- (140) *ka²b_{to} saŋgo²q̄ qom may=ki.*
 door close PASS TOTAL=M.PST
 ‘The door was shut completely.’

The affected entity may also be a locative adjunct:

- (141) *ho-kaṛ sou²b palaŋ=te gita² may=ki.*
 that-SG.HUM all bed=OBLQ lie TOTAL=M.PST
 ‘He took up the whole bed (e.g. by tossing and turning) (= he lay in the whole bed).’

For further discussion, see Peterson (in press).

Ambulative. san ‘while going’. This marker derives from an older lexical predicate **san* ‘go’ (cf. Santali *sen* ‘go’ and Kharia *sango²q̄* ‘walk’ (< **san* and the culminatory telic marker *go²q̄* (see above and also footnote (56)).

- (142) *ol=na olna kayar beloŋ san=ki.*
 bring=INF REP mango ripen AMB=M.PST
 ‘While being brought [to market], the mangoes ripened along the way.’

With *ol* ‘bring’, *qo²q̄* ‘take’, *col* ‘go’ and *qel* ‘come’, *san* denotes comitativity and is best translated as ‘along with’ (Kullū 1981:49). Note that Kullū also writes that *san* is only compatible with these four lexemes and in fact, some speakers I questioned rejected example (142) as ungrammatical.

- (143) *baysali tay ... ho=ki moŋ baysali koŋtaŋ ro baysali sāṛ*
 Baisali ABL that=PL 1 Baisali cow and Baisali ox
ol san=o²=ki.
 bring AMB=A.PST=PL
 ‘From Baisali ... they brought one Baisali cow and Baisali ox along with them.’ (MT, 120)

Suddenness: bha², hamba² and dha²b. These three markers denote that an action occurred suddenly or quickly. None of the three may be used as free lexemes.

- (144) *ki²jta²=te col bha²=ki u-tay hontay?*
 what.time=OBLQ go SUD=M.PST this-ABL then
 ‘When did [he] suddenly leave from here, then?’ (Kerkettjā 1990:2)
- (145) *akil=a² kayom no musniŋ a-ti²j tay=ko qel*
 mind=GEN opinion COMP one.day Q-side ABL=CNTR come
hamba²=ki moŋ maha dano ro po²da=ya² jhaṛi
 SUD=M.PST 1 big demon and village=GEN all
lebu=ki=te diyoga muḍu muḍu=te jo²=na māṛe=yo².
 person=PL=OBLQ daily 1.CLSSFR REP=OBLQ eat=INF begin=A.PST
 ‘It is believed (= opinion of the mind) that one day a great demon suddenly appeared from somewhere and began to eat all the people of the village, one by one.’ (MT, 1:64)

- (146) *pa²ṭopur=te=gā ho=ki ikuḍ=gā kisro=sikh=o²=ki hin=a²*
 Patna=OBLQ=FOC that=PL much=FOC wealthy=PRF=A.PST=PL that=GEN
ca²qom ho=ki rohtaspur hinte qoko dha²b=na pal=o²=ki.
 for that=PL Rohitasgarh LOC sit.down SUD=INF be.able=A.PST=PL

'In Patna they had become very wealthy, therefore they were able to settle (= sit) down quickly in Rohitasgarh.' (MT, 1:188)

Departive: *tu*. This marker denotes that the subject performed some action and then departed. *tu* cannot serve as the lexical root of a predicate.

- (147) *maha kulam paʔtopur tay sej muʔ=ki=may. ho=ki*
 big brother Patna ABL first go.out=M.PST=3PL that=PL
kuḍaʔb aw=ki=may ho=ki=te u gam tu=yoʔ=ki: "...
 behind COP=M.PST=3PL that=PL=OBLQ this say DPT=A.PST=PL
 'The elder brothers set out from Patna first. They said to those who were behind (before they left): "...' (MT, 1:160)

Conative: *dakhallakha*. Conativity is expressed through a number of means in Kharia, one of which is *dakha*. The form *lakha* is also occasionally encountered. Neither has an independent lexical meaning.

- (148) *aḍ=ki=yaʔ ompay paro=na boʔ=te=jo mudui=ki kuday*
 ANAPH=PL=GEN river cross=INF place=OBLQ=ADD enemy=PL chase
dakha=sikh=oʔ=ki, ...
 CONAT=PRF=A.PST=PL
 'The enemies_i also tried to chase [them_j away] at the place where they_j had crossed the river, ...' (MT, 1:167)

Excessive: *bay*. *bay* is homophonous with, and undoubtedly derives from, the lexical predicate *bay* 'make, do'. As a marker on the predicate it serves to denote that the action was 'more intense' than the unmarked form, or even 'excessive', somewhat similar to the function served in English by the adverbs 'extremely', 'strongly' or 'intensely'.

- (149) *muʔjḍaʔ pheinga=te leʔj bay=oʔ, 'u cini iḅ=aʔ*
 ant grasshopper=OBLQ curse EXCES=A.PST this sugar 1SG=GEN
bohin=na=ḅ=ki thuray=siʔ=may.'
 sister=1POSS=1SG=PL gather=PRF=3PL
 'The ant gave the grasshopper a bad scolding, "This sugar my sisters have gathered."' (TK, 2:42)

The compatibility of *bay* with lexemes is highly idiosyncratic and its semantic contribution is often unpredictable. For example, it is compatible with *aʔj* 'splash' and *gil* 'beat' with an 'excessive' or 'intensifying' meaning but also with *kayom* 'talk' but here with the resultant meaning 'deceive'. On the other hand, it is not compatible with *kamu* 'work' or *geʔb* 'burn', to name just two examples.

Autopoiesis: *jom*. With avolitional predicates, this marker denotes that something happened on its own, that is, there was no outside force which caused it to happen (150). With potentially volitional predicates it denotes that the actor simply performed the action because she/he wanted to and was under no obligation to do so (151). It is best translated by the English *just*. It is not used as an independent lexical morpheme, although, as Pinnow (1966:112f.) notes, it derives from an earlier lexeme meaning 'eat'.⁴⁸

- (150) (**beɾo buŋ*) *belom jom=ki*.
 sun INS ripen AUTOPOES=M.PST
 'It (e.g. fruit) just ripened (i.e. all by itself) (*through / *because of the sun).'
- (151) *ɟn koʃhri=te badli jom=oʔj*.
 1SG room=OBLQ change AUTOPOES=A.PST.1SG
 'I just changed the room around (for no particular reason).'

Benefactive: kay. *kay* denotes that an action was carried out on behalf of someone other than the subject. It has no independent lexical meaning.

In the following example from a children's story, a boy is in a tree in which loaves of bread grow. An elderly woman is waiting at the bottom of the tree and, thinking she is waiting for something to eat, he throws her down two loaves.

- (152) *hobne=te=ga kongher kãɽayboʔ=te yo=yoʔ ro*
 that.much=OBLQ=FOC boy old.woman=OBLQ see=A.PST and
ubar koloŋ ho-kar=aʔ mugamte ob-gur kay=oʔ.
 2 bread that-SG.HUM=GEN in.front.of CAUS-fall BEN=A.PST
 'After a while, the boy saw the old woman and threw down (= caused to fall) in front of her two loaves of bread for her.' (BB, 1:40)

3.2.11 Nominal incorporation

While relatively seldom, Kharia does allow a certain amount of incorporation of arguments and adjuncts into the predicate. Incorporation in general is limited to a small number of commonly occurring combinations. There is both non-productive (i.e. lexicalized) incorporation and semi-productive incorporation.

Non-productive incorporation. In non-productive incorporation the incorporated element appears directly after the predicative root and precedes all TAM markers. If the incorporated element is polysyllabic, one syllable, usually the last, is dropped. In the incorporation of monosyllabic free lexemes, no phonological changes occur (* denotes a reconstructed form).

- (153) *ajoʔd* 'dry up' (ITR) *ajoʔd-qaʔ* 'dry up (of water)(ITR)' < *qaʔ* 'water'
si 'plow' *si-loʔ* 'plow' < **loʔ*⁴⁹

This process creates a new (compound) stem, and what appears to be an incorporated subject or object is in fact not an argument but merely a part of this new stem, as the following examples show. That is, this can only be considered 'incorporation' from a historical point of view.

- (154) *qaʔ ajoʔd-qaʔ=ki*
 water dry.up-water=M.PST
 'The water dried up.'
- (155) ... *hoɽom=ki=yaʔ goʔjloʔ=te si-loʔ=na ayiʔj laʔ*
 other=PL=GEN rice.field=OBLQ plow-earth=INF PRS.COP then
madet remaʔ=te=ki
 help call.for=A.PRS=PL
 '... the rice fields of the others must be plowed, then [they] call for help.'
 (AK 5,13–14)

Below I present a few more examples.

- (156) Incorporation of ‘subject’
iʔj ‘defecate’ *iʔj-thaŋ* ‘defecate (of cows)’ < **taŋ* ‘cow’⁵⁰
- (157) Incorporation of ‘object’
ol ‘take’ *ol-ɬay* ‘marry’ < *kanday* ‘wife, woman’
- (158) *Other types*
betoʔɬ ‘be(come) hungry’ *betoʔɬ-ɬaʔ* ‘be(come) thirsty’ < *ɬaʔ* ‘water’
boŋ ‘come to an end, finish’ (ITR) *boŋ-sor* ‘petrify’ < *soreŋ* ‘stone’

Semi-productive incorporation. Some speakers will accept the incorporation of non-specific and non-referential objects into the predicate itself. The incorporated element in this construction appears directly after the reciprocal marker *kol*, if present, and before the root. All speakers who accepted these forms, however, preferred forms in which the object was not incorporated. All forms given here were elicited in interviews. Only minor optional phonetic changes, such as the loss of /ʔ/ in the coda, occur.

Acceptable (to some)

Preferred

- (159) *kol* *ulaʔ* *likha=ki=kiyar* *ulaʔ kol likha=ki=kiyar*
 RECIP letter write=M.PST=DL
 ‘they wrote each other letters.’
- (160) *kol* *ɬa* *uʔɬ=ki=kiyar* *ɬaʔ kol uʔɬ=ki=kiyar*
 RECIP water drink=M.PST=DL
 ‘they gave each other water to drink.’

This type of incorporation is limited to certain commonly occurring combinations, such as *letters* and *writing*, *water* and *drinking*, etc., but is not possible in other combinations:

- (161) **kol* *tomliŋ* *uʔɬ=ki=kiyar*
 RECIP milk drink=M.PST=DL
 ‘they gave each other milk to drink.’

3.2.12 *Auxiliary verb constructions*

There are a number of auxiliaries which occur with a lexical predicate in the infinitive. These cover a variety of functions, including aspect in its most general sense (‘phase verbs’ and general imperfectivity) as well as modals.

Aspect, ‘phasal verbs’

laʔ ‘IMPERFECTIVITY’. This is by far the most common auxiliary. It is usually found in the past tense and is also compatible with the irrealis or perfect, but not with the general present nor with the progressive. It denotes imperfectivity in general, including iterativity, habituality and non-present progressivity. It is always marked for middle voice in those categories in which an active/middle opposition is found (i.e. past, irrealis).

- (162) ... *khaṛiya lebu=ki pujaṣaṭh karay=na laʔ=ki=may*, ...
 Kharia man=PL sacrifice do=INF IPFV=M.PST=3PL
 ‘... the Kharia men used to perform sacrifices ...’ (AK, 2:6)
māṛe, suru ‘begin’
- (163) *jeʔ botoṅ buṅ thartharay=na māṛe=yoʔ=ki*.
 so fear INS tremble=INF begin=A.PST=PL
 ‘So [they] began to tremble with fear.’ (MT, 1:32)
melay ‘stop’
- (164) *hin=aʔ ghaʔd̥=ga ho=ki=te oʔb-koda=na melay*
 that=GEN for=FOC that=PL=OBLQ CAUS-paint.oneself=INF leave
goṭh=oʔ=ki
 C:TEL=A.PST=PL
 ‘Therefore, [they] stopped tattooing (=painting) them.’ (MT, 1:173)

Modality *lam* ‘seek; want’. As an independent predicate, *lam* has the meaning ‘look for’. As an auxiliary with an animate subject, it means ‘want’:

- (165) ... *ij̄n u ikuʔd̥ sundar kontheʔd̥=ki=te bajhay kon*
 1SG this very beautiful bird=PL=OBLQ trap SEQ
satay=na um=ij̄n lam=te.
 torment=INF NEG=1SG want=A.PRS
 ‘... I don’t want to trap these beautiful birds and torment them.’ (BB, 2:43)

ter ‘allow’. *ter* as a lexical predicate means ‘give’ but as an auxiliary it has a permissive meaning.

- (166) *umboʔ, jiyom aw=na bheir=ko ho-ghay=na um=niṅ*
 NEG life COP=INF up.to=CNTR that-way=INF NEG=1PL.INCL
ter=e.
 allow=A.IRR
 ‘No! As long as there is life [in us], we will not allow that to happen (= we will not allow [it] to become that way).’ (Kerkeffā, 1990:7)

pal ‘can; be able’. With an infinitive, *pal* denotes capability. It has no independent lexical meaning. See examples (65), (146), (184), (209).

The copula obligation, necessity. The person who is under obligation appears in the oblique case and the infinitive is the ‘subject NP’.

- (167) ... *ij̄n=te rayga bhere=yaʔ ghaʔd̥ terom thuray=na ayiʔj̄*.
 1SG=OBLQ cold time=GEN for honey gather=INF PRS.COP
 ‘... I have to gather honey for the cold season.’ (TK, 2:20)

3.3 Expressives

Kharia has a construction which is often termed the ‘echo-word’ formation in South Asian linguistics, in which a lexical stem is followed by an element which has no independent meaning but which modifies the meaning of the first element, generally indicating something akin to the English ‘etc.’, for example, *raja rajwaṛ* ‘king ECHO’ = ‘the king etc., the king and his court’. The ‘echo-word’ construction in Kharia is

actually a collection of forms from at least three different sources, which I now discuss separately.

- 1 The first consists of forms in which the second element seems to have once had an independent meaning, which is however no longer in common use. In all cases it appears that the original meaning of this element was very close to that of the first element, a common strategy in ‘compounding’ in Kharia, cf. *ompay khirom* ‘big.river river’ = ‘rivers’, *ba? rumku?b* ‘unhusked.rice husked.rice’ = ‘rice’ or *paysa dhebua* ‘money money’ = ‘wealth’ (adapted from Malhotra 1982:72f.).

Consider now *kinir jhaṅkoy*, which Malhotra (1982:73) translates as ‘forest, wilderness’ and analyses as a compound of *kinir* ‘forest’ and *jhaṅkoy* [also realized as *jhaṅkor*, JP] ‘scrubland’. Some speakers I questioned were not familiar with this second term in isolation, although it would seem to have lost its independent meaning only recently, since it is still attested in Roy and Roy (1937:165) with the form *jhankor* and the meaning ‘forest’ or ‘*sarna* (holy forest)’.

Other expressions which would seem to belong to this category are *kunḍu? hakon* ‘family’ from *kunḍu?* ‘child’ and the echo-word *hakon*, which Pinnow (1965a:92, ln.10) translates as ‘illegitimate child’ (*Kegel*) but which is no longer used as an independent morpheme, and *khe?d.jol* ‘bite, etc.’ from *khe?d* and the echo-word *jol* (?).

- 2 The second group is quite large and consists of loans from Sadani. Here just a few examples: *laṛe bhiṛe / laṛai bhiṛai* ‘fighting, etc.’, *main marjad* ‘honour, etc.’, *culha cawka* ‘stove, etc.’, *kado walo* ‘mud, etc.’, *khoj puchar* ‘investigation, etc.’.
- 3 Finally, there is a highly productive construction in Kharia in which a telic action can be denoted through a type of reduplication which is similar to the echo-word formation. The meaning is usually that of a durative or gradual telic action, although there are some irregularities. In the following I present a few examples.

(168)	<i>Unmarked stem</i>	<i>Gradual telic</i>
	<i>col</i> ‘go’	<i>col cila</i> ‘go away’
	<i>ḍubay</i> ‘press down, deplete’	<i>ḍubi ḍubay, ḍubay ḍubi</i> ‘completely deplete’
	<i>dho?</i> ‘grab, take’	<i>dho? dhiga</i> ‘take, grab (e.g. an electric wire which one is not supposed to touch, or steal s.th.)’
	<i>melay</i> ‘leave’	<i>meli melay</i> ‘leave behind’
	<i>jo?</i> ‘eat’	<i>jo? jiga</i> ‘eat up’
	<i>uwa?</i> ‘bathe, wash’	<i>uwa? uwi?</i> ‘finish washing up, bathing’

4 SYNTAX

4.1 Syntax of the simple sentence

Kharia has the unmarked constituent order Subject–Object–Predicate, although word order is quite flexible: Topical elements tend to appear first in the sentence, with new information appearing further to the right but before the predicate, as in the following example:

- (169) *tama iṅ anjṅ=a? gotar=a? kahni batay=na col=ta?jḍ=iṅ.*
 now 1SG 1PL.INCL=GEN clan=GEN story tell=INF GO=M.PROG=1SG
 ‘Now I am going to tell the story of our clan.’ (AK, 1:1)

As the following example shows (= second half of (149) above), the subject can follow the object (or other entities) if the object is considered more topical. Here, the ant is telling the grasshopper that she cannot give him any sugar, since her sisters have worked hard gathering it. The sisters are mentioned here for the first time in the story and are hence less topical than the sugar, which the grasshopper has just asked for.

- (170) *u cini jn=a? bohina=na=j=ki thuray=si?=may.*
 this sugar 1SG=GEN sister=1POSS=1SG=PL gather=PRF=3PL
 ‘This sugar my sisters have gathered.’ (TK, 2:42)

Although the predicate is usually in clause-final position, this is not obligatory. Especially common is the placement of the goal with predicates of motion after the predicate:

- (171) *...ro dhirom dhirom ho=ki dam=ki=may raylogarh.*
 and slowly REP that=PL arrive=M.PST=3PL Raylogarh.
 ‘... and they gradually arrived at Raylogarh.’ (MT, 1:82)

Sentence types. The majority of examples presented so far have been for what Bloomfield (1933[1984]:173) refers to as ‘narrative predication’, in which, for example, a narrative is being continued, such as when the subject performs an action. This is in contradistinction to what Maas (2004) refers to as qualitative predication, in which an inherent property of an entity is given (e.g. *I am a man, That is a tree*). ‘Inherent’ here should be taken to mean something akin to ‘central or most important characteristic’ and will undoubtedly be subject at least to some extent to a particular speaker’s judgement.

There is also a third type, which I will refer to as non-inherent qualitative predication, in which the position of an entity is given or a statement is made about a temporary or non-essential characteristic of this entity. As the preceding text contains abundant examples for narrative predication, I will only discuss inherent and non-inherent qualitative predication in the following text.

Table 9.14 presents the stems of the present copula. Forms other than the present are derived from *aw*, which as a lexical morpheme has the meaning ‘stay, remain; live’: Here the inherent / non-inherent distinction is neutralized. The present forms do not mark for basic voice but do mark for PERS/NUM/HON.

Inherent qualities. Except in the present, the expression of inherent qualities obligatorily employs a copula alongside the predicate NP. An example:

- (172) *raja=ya? jimi aw=ki sembho odo rani=ya? jimi*
 king=GEN name COP=M.PST Sembho and queen=GEN name
aw=ki dakay.
 COP=M.PST Dakay
 ‘The king’s name was Sembho and the queen’s name was Dakay.’ (AK, 1:4)

TABLE 9.14: THE PRESENT COPULAE, NON-NEGATIVE AND NEGATIVE

Type of copula	Non-negated form	Negated form
Inherent qualities	<i>heke</i>	<i>nalage</i>
Non-inherent qualities	<i>ayi²j(d)</i>	<i>umbori²j(d)</i>

In the present, the copula is also generally used, although not always. Both the non-negated form *heke* and the negated form *nalage* have been borrowed from Sadani.

- (173) *u-je?*=ga *heke* *khariya* *gotar*=a? *cho?ka* *kahani*.
 this-SG.NHUM=FOC PRS.COP Kharia clan=GEN small story
 ‘This is a short history of the Kharia clan.’ (AK, 1:70)

The copula often precedes the predicate NP, as in the last two examples. As the following shows, however, this is not necessarily the case.

- (174) *u* *kahani* *u* *go?a* *duniya*=te *lebu*=ki=ya? *kahani* *heke*.
 this story this entire world=OBLQ person=PL=GEN story PRS.COP
 ‘This is the story of the people on the entire world.’ (AK, 3:1)

Finally, the copula is often omitted in the present.

- (175) ...*ro* *u*=ga *ho* *jinis*=a? *koma?a*.
 and this=FOC that animal=GEN meat
 ‘... and this [is] the meat of that animal.’ (AK, 1:57)

Non-inherent qualities. This type of predication makes use of the copula *ayiⁱjd* (*ayiⁱj* with all persons other than the first and second, singular) when non-negated and *umbo?iⁱjd* (*umbo?iⁱj* with all persons other than the first and second, singular) when negated. *ayiⁱjd* appears to be an Indo-Aryan loanword: Although it is not entirely clear from which language, it is clearly not from (modern) Sadani, where the corresponding form is *āh-*. However, the presence of the third person, non-honorific form *əich* ‘he/she/it is’ in Maithili (Yadav 1996:217ff.) would seem to indicate that it has in fact been borrowed from some Indo-Aryan source.

- (176) *ho*=kiyar=ya? *tomo?d*=te *dā?a?* *ayiⁱj*.
 that=DL=GEN mouth=OBLQ stick PRS.COP
 ‘There is a stick in their mouths.’ (TK, 1:44)
- (177) *u* *khori*=te *be?a* *umbo?iⁱj*=may.
 this village.section=OBLQ boy NEG.PRS.COP=3PL
 ‘In this village section there are no boys.’ (RK, 5:3)

ayiⁱjd is generally present but optional:

- (178) *samudar* *lekhe* *ompay*=ki *bo?e* *bo?e*.
 ocean like river=PL be(come).big REP
 ‘The river [is] very big, like an ocean.’ (MS, 1:157)

4.2 Complex sentence structure

Kharia has a number of conjunctions and disjunctions. Their presence has no effect on word order. Here I present a few of the most common ones, followed by a simple example.

- (179) *ro*, *o?o*(?) ‘and’ *hina?* *tho?*, *hina?* *ghaⁱd* ‘therefore’
muda, *lekin*, *magar* ‘but’ *tewjo*, *tewa?jo* ‘nevertheless’
la? ‘(and) then’ *ina no* (why COMP) ‘because’
la?ko ‘but’ *um la?ko* ‘otherwise’
agar, *yadi* ‘if’

- (180) *musniḡ dinu bheir lam=oʔ=ki muda umay kuy=oʔ.*
 once day entire search=A.PST=PL but NEG.3PL find=A.PST
 ‘Once [they] searched all day but didn’t find [the demon]’ (MT, 1:68)

4.3 Complement clauses

In subject function. Complement clauses in subject function are rare. The predicate may be an infinitive (181) or a finite predicate in which the entire complement clause may then be marked by the postposition *guʔd* ‘like’ (182), although this is not necessary (183).

- (181) *oreʔj koḡtaḡ bui=na=ko ho=ki=yaʔ dhatam aw=ki.*
 ox cow raise=INF=CNTR that=PL=GEN custom COP=M.PST
 ‘But raising oxen and cows was their custom.’ (MT, 1:9)

- (182) *menson=ko um dḡaḡ=e=m gudḡ laʔ=ta.*
 once=CNTR NEG send=A.IRR=2SG like EMOT=M.PRS
 ‘It looks like you won’t repell (= send) [them] all at once.’ (Kerkeṭṭā 1990:11)

- (183) *am um co=na=m laʔ=ta.*
 2SG NEG go=M.IRR=2SG EMOT=M.PRS
 ‘It looks like you won’t be going.’

In object function. Complement clauses in object function are quite common, especially with predicates of speech and thought. These clauses have a finite predicate and usually begin with the complementizer *no*.

Speech and thought are generally presented as they would have been uttered by the person who said or thought them. *no* can be omitted.

- (184) *ap=dom raṭa=te remakh=oʔ ro gam=oʔ no ‘babu*
 father=3POSS Rata=OBLQ call=A.PST and say=A.PST COMP child(VOC)
musa iḡ kinir co=na um=iḡ pal=e.’
 today 1SG forest go=INF NEG=1SG be.able=A.IRR
 ‘His father called Rata and said [to him] “Son, today I will be unable to go to the forest.”’ (BB, 2:21)

Instead of *no*, *gam kon*, the sequential converb of *gam* ‘say’, is occasionally found clause-finally as a kind of quotative particle, although this is not common:

- (185) *‘je janwar tar=e=ki ho janwar=yaʔ ghos ol=e=ki’*
 CR animal kill=A.IRR=PL that animal=GEN meat take=A.IRR=PL
gam kon gam=oʔ, swapan, munu buḡ
 say SEQ say=A.PST dream dream INS
 ‘“Whatever animal they kill, that animal’s meat they should bring [back to you]” [he] said, in (= through) a dream.’ (MS, 2:8)

4.4 ‘Relative clauses’

Kharia is especially rich in strategies for expressing predicative nominal attribution, which I will simply refer to here as ‘relative clauses’. I have found 13 productive strategies, including pre-, post- and circumnominals as well as two types of correlatives.

The prenominals show the greatest amount of variation, including prenominal participial constructions, prenominal infinitives and prenominally juxtaposed (fully, semi- and non-) finite predicates. As space will not allow for a full discussion of these forms, I merely present an example for each construction below and refer the reader to my (2006) grammar.⁵¹

4.4.1 Correlatives

There are two types of correlative constructions in Kharia:

- *The 'je'-class.* The correlative marker has the invariable form *je*, which has been borrowed from Indo-Aryan (cf. the Sadani correlative pronominal *je*).
- *The 'ali'- or 'Q'-class.* This construction makes use of correlative forms which are homophonous with interrogatives, generally either the bound interrogative marker *a-* 'Q' followed by a lexical nominal, such as *a-ti²j* 'Q-side' = 'where', or its free form *ata* 'Q'. There is also a small class of interrogatives which are based on the morpheme *i* 'what?', such as *i-ghay* (what-way) 'how' or *i-bhere* (what-time) 'when'. Other forms found here include *be(ha)r* 'who'.

With both of these constructions, the head nominal is typically present in both the main and subordinate clauses, and often appears in the main clause with a demonstrative. Alternatively, it is indicated only by a pronominal in the main clause.

- (186) *in jelata phon²ten buj likha=sikh=o²j ho-je? kuy=o²j.*
 1SG CR pen INS write=PRF=A.PST.1SG that-SG.NHUM find=A.PST.1SG
 'I found *the pen I had written with*.'

In a very similar construction, the 'correlative' can be used postnominally as a simple relative pronominal:

- (187) *ho q²ho²hi=te=ga moj kinir=a? jantu a-je? no*
 that hollow=OBLQ=FOC I forest=GEN animal Q-SG.NHUM COMP
cho²tanagpur=a? lebu 'ba²lo' gam=te=ki aw=na la²=ki.
 Chotanagpur=GEN person wildcat say=A.PRS=PL live=INF IPFV=M.PST
 'In just that [tree] hollow lived **an animal of the forest** which the people of Chotanagpur call "Wildcat".' (Pinnow 1965a:62)

4.4.2 Prenominals

Participles (section 3.2.8)

Infinitive + wala

- (188) *jharkhand²=te aw=na=wala lebu=ki iku² jughay milansar*
 Jharkhand=OBLQ live=INF=PRTCPL person=PL very much friendly
aw=ta=ki.
 COP=M.PRS=PL
 'The people who live in Jharkhand are very friendly.'

Participle in -l

- (189) *muda moj brahman ho-ka²=a? daru=te ta²na-l*
 but 1 brahman that-SG.HUM=GEN tree=OBLQ hang-PRTCPL

janew=te *yo=yo?*
 holy.thread=OBLQ see=A.PST

‘But a Brahman_i saw **the holy thread** which he_j had hung on the tree.’ (AK, 2:14)

Participle in -qu?

- (190) *ho=ki* *ho-kaq=te* *yo-qu?* *dinu* *somto?* *aw=ki*.
 that=PL that-SG.HUM=OBLQ see-PRTCPL day Monday COP=M.PST
 ‘**The day** they saw him on was a Monday.’

-nV- ‘nominalization’ (section 3.1.10)

- (191) *khariya=ki* *ber=a?* *ji<ni>b* *pe?=jo* *umay* *no?=na*
 Kharia=PL who=GEN touch-<NML> rice=ADD NEG.3PL eat=INF
la?=ki, ...
 IPFV=M.PST
 ‘The Kharia did not use to eat **rice** which had been touched by anyone,’
 (Pinnow 1965a:122, ln. 57)

Infinitives (section 3.2.8)

Genitive case

- (192) *kabhi kabhi*, *moka soka*, *kole²j=ta=may* *lekin*
 sometimes sometimes fight=M.PRS=3PL but
mel *prem* *se=ga* *aw=na=ya?* *kornis* *karay=te=ki*.
 harmony love ABL=FOC live=INF=GEN attempt do=A.PRS=PL
 ‘Sometimes they fight, but they try to live in peace and harmony.’ (AK, 5:33)

Unmarked form (= direct case): (= (109))

- (193) *musnij* *u-ghay* *hoy=ki* *no* *qoli=te* *qo²q=na*
 one.day this-way happen=M.PST COMP sedan=OBLQ take=INF
bhere *u* *khariya* *pahan=te* *haqa* *la?=ki*.
 time this Kharia priest=OBLQ pee EMOT=M.PST
 ‘One day it happened this way that, **at the time** [they were] to take away the sedan chair, the Kharia priest had to pee.’ (AK, 2:8)

Differing levels of finiteness of the predicate in the attributive clause (section 3.2.8)

Fully finite

- (194) *ijn* *yo=yo²j* *lebu=ki* *ijn=a?* *hotel=te* *aw=ta=ki*.
 1SG see=A.PST.1SG person=PL 1SG=GEN hotel=OBLQ live=M.PRS=PL
 ‘**The people** I saw live in my hotel.’

Semi-finite (*aw=ta* for the fully finite *aw=ta=ki*)

- (195) *jharkhand=te* *aw=ta* *lebu=ki* *ikuq* *jughay* *milansar*
 Jharkhand=OBLQ live=M.PRS person=PL very much friendly
aw=ta=ki.
 COP=M. PRS=PL
 ‘**The people** who live in Jharkhand are very friendly.’

Free-standing forms

- (196) *ijn=a?* *likha* ***phonten=te*** *kuy=o?*²*j.*
 1SG=GEN write pen=OBJ find=A.PST.1SG
 'I found **the pen** I had written with.'
- (197) *am=pe* *leḡ-leḡ* ***jāut*** *tar* *ol=e=pe.*
 2=2PL fly-REDPL animal kill bring=A.IRR=2PL
 'You kill and bring [back] **flying animals** (= birds).' (MT, 1:131)

4.4.3 *Circumnominals*

- (198) *biru* *ḡeb=na* *ḡebna* ***ba?*** ***ruḡku?***²*b* *ro?*²*ki*
 mountain ascend=INF REP 'rice' spill=M.PST
kay=e=bar!
 lift=A.IRR=2HON
 'Pick up (HON) **the rice** which spilled while you were walking up the hill!
 [MT, 1:40]

4.5 'Adverbial clauses'

Purpose clauses. Purpose clauses are constructed with the infinitive, often in the genitive case, and either of the following postpositions: *gha?*²*q* 'for; PURP' and *thon* 'for; PURP'.

- (199) *la?* *sou?*²*b=ga* *bhai=ki* *ho* *khajar* *tar=na* *ghaḡ* *juda juda*
 then all=FOC brother=PL that deer kill=INF PURP separate REP
*mu?*²*ki=may,* *kinir=te.*
 emerge=M.PST=3PL forest=OBLQ
 'Then all the brothers set out separately to kill that deer, into the forest.'
 (AK, 1:15)

Causal clauses. The most common means of expressing a causal clause is through the conjunction *ina no* 'because', which consists of *ina* 'why?' and the complementizer *no*. In these clauses, the predicate is always fully finite.

- (200) *kunḡab* *aw=ki* *tomliḡ* *khariya* *gam* *ḡom=na* *la?*²*ki=may*
 behind COP=A.PST milk Kharia say PASS=INF IPFV=M.PST=3PL
ina no *u=ki* *tomliḡ* *u?*²*ḡ=ga* *ḡel=ki=may.*
 why COMP this=PL milk drink=FOC come=M.PST=3PL
 'Those who were behind were called "Milk Kharia" because they came drinking milk.' (MT, 1:180)

Temporal clauses. To express that one event took place before or after another event, either the postposition *seḡ* 'before' or *lo?*²*ḡho* 'after' is used. The predicate of the temporal clause appears in the infinitive, generally in the genitive case.

- (201) *ijn=a?* *gam=na* *no* *co=na=?* *seḡ* *sosreir=ga?*
 1SG=GEN say=INF COMP go=INF=GEN before husband's.parents'.home=FOC
rema? *ḡaḡ=si?*²*may,...*
 call send=PRF=3PL
 'It's my opinion (= saying) that, before he goes, your parents⁵² should call [him] and send [him back], ... (= have called [and] sent).' (Kerketta 1991:11)

- (202) *uɔ=na loʔdho aɔi ho maha daru tuta=te=ga*
 drink=INF after ANAPH that big tree bottom=OBLQ=FOC
ɔel=ki.
 come=M.PST
 'After drinking [the water], he came to the bottom of that big tree.' (AK, 1:23)

A common means of expressing a temporal clause is with a 'relative clause' modifying the head nominal *bhere* 'time'. The predicate of the subordinate clause can be an infinitive, as in (109)/(193), although it is probably more often fully finite, as in the following example:

- (203) *ho-kaɔ ʒhik biha=na laʔ=ki=may bhere ɔam=ki.*
 that-SG.HUM just marry=INF IPFV=M.PST=3PL time arrive-M.PST
 'He arrived just as they were celebrating the wedding (= **at the they-were-just-getting-married time**).' (Pinnow 1965a:41)

Conditionals and counterfactuals. In both conditional and counterfactual clauses, the apodosis is always marked by the conjunction *laʔ* 'then'. The Indo-Aryan loanword *agar* 'if' can appear in the protasis, although it is generally not used. Very rarely, *yadi* 'if', a borrowing from modern Indo-Aryan (where it is a borrowing from Sanskrit), is found.

In conditionals, the predicate of both the protasis and apodosis appears in the irrealis, regardless of the degree of probability. In counterfactuals, the predicate of the protasis appears either in the irrealis or in the irrealis perfect, while that of the apodosis is marked as irrealis.

Conditional

- (204) *am am=aʔ tiʔ buɔ ob-dhog=e=m laʔ sebol*
 2SG 2SG=GEN hand INS CAUS-grab=A.IRR=2SG then delicious
laʔ=na.
 EMOT=M.IRR
 'If you give me it (= cause [me] to grab [it]) with your [own] hands, then [it] will taste delicious.' (BB, 1:43)

Counterfactual

- (205) *agar iɲ hante aw=siʔ=na=iɲ laʔ ho-kaɔ=aʔ*
 if 1SG there COP=PRF=M.IRR=1SG then that-SG.HUM=GEN
kayom onɔor[=e]=iɲ.
 speech hear[=A.IRR]=1SG
 'If I had been there I would have listened to His [= God's] words.' (Pinnow 1965a:103)

Concessives

Concessive clauses are formed by the correlative construction *jawbhi ... tawbhi* 'although ... nevertheless', which has been borrowed from Indo-Aryan. *tewjo* or *tewaʔjo* 'nevertheless' is also found instead of *tawbhi*. The predicate of the subordinate clause is fully finite.

- (206) *jawbhi ho-kaṭ dhaṅgar ayij, tawbhi ho-kaṭ=aʔ=te*
 although that-SG.HUM servant COP.PRS still that-SG.HUM=GEN=OBLQ
laij besu peʔ um kui=taj.
 stomach full rice NEG find=M.PROG
 'Although he is a servant, still he goes hungry (= a stomach-full of rice isn't
 being found to him).' (Pinnow 1965a:60)

4.6 Same subject and different subject in subordination

There is a strong tendency in Kharia to restrict the sequential converb to clauses which have the same subject as the main clause and to use a free-standing form (section 3.2.8) in the oblique case when the subject of the subordinate clause is different from that of the main clause.⁵³

Same subject. (from a story on the mythological pre-historical wanderings of the Kharia in Egypt)

- (207) *dhirom dhirom khaṭiya maha rokeʔdʒoʔ paro kon rusuŋ*
 slowly REP Kharia big desert cross SEQ red
samuder dʒam=ki=may
 ocean arrive=M.PST=3PL
 'Slowly but surely, the Kharia crossed the great desert and arrived at the
 Red Sea.' (MT, 1:21)

As with many other South Asian languages, a so-called dative subject may also be the controller of the sequential converb, although it is marked for oblique case:

- (208) *etwa=te u=ki=yaʔ haleit yo kon lebui laʔ=ki.*
 Etwa=OBLQ this=PL=GEN condition see SEQ love EMOT=M.PST
 'Etwa, seeing their condition, felt compassion.' (RD, 2:108)

Different subject

(Free-standing form = *aw-aw=te*)

- (209) *muda dada=dʒom=ki=yaʔ aw-aw=te ber*
 but elder.brother=3POSS=PL=GEN COP-REDPL=OBLQ who
beʒi=ki=te um uʒuŋ=na pal=e.
 daughter=PL=OBLQ NEG make.flee=INF be.able=A.IRR
 'But as long as their elder brothers are present, no one will be able to drive
 the daughters away.' (MT, 1:169)

This is, however, for most speakers only a tendency, and the use of the sequential converb is not categorically restricted to same subject contexts. In the following example, it is the porcupine (*jijray*) who has stayed a few days, before the father sends his daughter off with the porcupine:

- (210) *thor toʔ aw kon jijray=aʔ sori dʒaŋ goth=oʔ.*
 few day stay SEQ porcupine=GEN with send C:TEL=A.PST
 '[Ø:] having stayed a few days, [he:] sent [her_k] off with the porcupine_i.'
 (Pinnow 1965a:40)

The use of the sequential converb with a clause whose subject is different than that of the main clause is, however, uncommon. In fact, for some speakers the sequential converb cannot be used in different-subject environments and the attested example (210) above was considered ungrammatical by one speaker I questioned, while other speakers accepted it as grammatical.

Similarly, there appear to be two different strategies to express ‘without doing’, depending, at least tententially, on whether both clauses have the same or different subjects.

Same subject. In sentences where both clauses have the same subject, *enem* ‘without’ is used and the lexical predicate appears in its base form, as in (211), or it may appear as a sequential converb (212).

(211) *musa=ko in enem daichna- dʔoʔd um=in melay[=e].*
 today=CNTR 1SG without fee-take NEG=1SG leave=A.IRR
 ‘Today, without receiving my fee, I will not leave.’ (Kerkettjā 1990:3)

(212) *jaisan musa in enem yo kon saʔak-tij baʔke=na*
 CR today 1SG without see SEQ road-side hurry=INF
laʔ=ki=in no?
 IPFV=M. PST=1SG Q
 ‘Like I started⁵⁴ running towards the street today without looking?’
 (Pinnow 1965a:95, section 4)

Different subject. With different subjects, *enem* is used with the free-standing form, as in example (47), repeated here as (213).

(213) *enem raja ro rani=kiyar=aʔ aw-aw=te khaʔiya=ki*
 without king and queen=DL=GEN COP-REDPL=OBLQ Kharia=PL
kaʔiʔj=ko qher=ga akbakay=ki=may.
 somewhat=CNTR very=FOC be.in.a.flurry=M.PST=3PL
 ‘Without there being a king and queen, the Kharia panicked.’ (MT, 1:104)

5 SEMANTICS

Here I briefly address a phenomenon which is typical for Kharia, that of ‘semantic agreement’. Consider the following example:

(214) *botoj=ta=pe ho=ki lutui su kon peʔ cakhnaʔ iʔjthaj*
 fear=M.PRS=2PL that=PL clothes put.on SEQ rice curry cow.dung
kinbhar=na=pe, in=ko laʔe=na=in.
 courtyard=M.IRR=2PL 1SG=CNTR fight=M.IRR=1SG
 ‘Those of you who are afraid, you put on your [house] clothes and see to your house work like cooking and cleaning the courtyard with cowdung, but I will fight.’ (Kerkettjā 1990:7)

Consider the apparent contradiction between the ‘subject’ pronominal and the ‘agreement’ marker on the predicate. The subject would appear to be *hoki* ‘they’, but the predicate is clearly marked as having a second person, plural subject, ‘you (PL)’. This is further complicated by the fact that *hoki* is the head of a ‘relative’ clause which consists of the predicate *botoj=ta=pe* ‘you (PL) are afraid’.

As this utterance is directed toward a large group of people, the subject is the second person, plural, although only a part of this group is meant, that is, ‘those of you’, not ‘(all of) you’. In cases such as these, Kharia invariably chooses the second person over the third person in terms of predicate marking, as the group whose members are intended is being directly addressed. This also entails ‘overriding’ what would seem to be the explicitly mentioned grammatical subject, *hoki* ‘they’. Here I present two further, similar examples:

- (215) *koro²b=si[?]=na=pe. ber=jo i=jo a[?]=pe gam=e.*
 silent=PRF=M.IRR=2PL who=ADD what=ADD NEG.MOD=2PL say=A.IRR
 ‘Be quiet! Don’t any of you say anything.’ (Kerkeṭṭā 1990:2)

behar bujhay=e=pe?
 who explain=A.IRR=2PL
 ‘Which of you will explain?’

As noted in sections 3.2.1 and 3.2.9, these data strongly favour considering the ‘subject NP’ to be a kind of apposition to the person marker on the predicate, which would then be the actual subject of the clause.

6 LEXICON

Although the status of Kharia as a Munda language cannot be called into question, as core areas of the lexicon such as the pronominals, grammatical morphemes and many other areas have cognate forms in other Munda languages, the language has been enormously affected by its neighbouring languages, such as Sadani, Mundari, and in more recent times (standard) Hindi, Persian and English (the last two most likely via Hindi and Sadani), and even Biblical Hebrew through contact with Christian missionaries.

6.1 Austroasiatic/Munda components

The personal pronominals have probably been the most resistant area of the lexicon against borrowing, although even here it is possible (although in my opinion highly unlikely) that there is one loan, *adi*, the anaphoric pronominal of the third person, which may be a loan from Kurukh (Grierson 1906 [1994]:194). All other pronominals are clearly of Munda origin. Similarly, most grammatical markers such as TAM markers, case markers, etc., can easily be shown to be of Munda origin.

Other areas of the lexicon have been more receptive to loans – referential lexemes, predicating lexemes and postpositions of Indo-Aryan origin are found in abundance. Even areas such as kinship terms, household appliances, cardinal and ordinal numerals and religious life now abound in loanwords from Indo-Aryan. Although there are no exact figures, easily one-third – but more likely closer to one-half – of the lexemes now in current use have been borrowed, and many of the native lexemes found, for example, in the texts Pinnow (1965a, b) collected are no longer understood by younger speakers I consulted (i.e. up to about 30 years of age) and often not even by older speakers.⁵⁵

6.2 Loan strata

6.2.1 Loans from North Munda

Basing their work on the Kharia words given in Pinnow (1959), Zide and Stampe (1968) present a list of 17 lexemes in Kharia which may have been borrowed from North Munda, most likely from Mundari, which is still spoken today in many of the same villages as Kharia and with which it has undoubtedly been in close contact for centuries.

Some of these putative loans are indeed likely candidates, such as *rutu* ‘small fife’, whose form can only be explained through recourse to Mundari grammar, and *duray* ‘song’, whose form (initial /d/) would certainly argue for a loan from Mundari, as one would expect a /dʒ/ in the onset in words of Kharia origin (Zide and Stampe 1968:174).

On the other hand, some of the forms the authors list require more research, such as *gonoy* ‘bride price’, which is to my knowledge actually *giniy-tay*, an archaic compound composed of *giniy* ‘price’ and the bound morpheme *-tay* ‘cow’. Others, such as Kharia *haḍa*, Santali *aḍo* ‘urine’, are probably cognate forms deriving from the same Proto-Munda forms.

Others are more doubtful, such as *san* ‘go’, Santali *sen* ‘go’: In Kharia there is a lexeme *sangoʔq* ‘walk’, which has lexicalized from the combination **san* ‘go’ and *goʔq* ‘C:TELIC’. *san* is no longer used as a free lexeme, but is still found in its ‘ambulative’ function (section 3.2.10). Its use as a grammatical marker in (Dudh) Kharia⁵⁶ would favour an analysis of this form as being of common Munda descent, not a loan word from North Munda, for which there is no evidence. The different vowels (*sen* (Santali) vs. *san* (Kharia)) also point in this direction.

6.2.2 Loans from Sadani

By far the largest number of loanwords comes from Sadani. Loans from Sadani can be found in almost all areas of the lexicon. For example, Sadani loans have now completely ousted the native numerals (section 3.1.8) and the large majority of kinship terms now in use are from Sadani, even when the Kharia words are still widely known, for example, *bhai* ‘brother’ and *bahin* ‘sister’. The native expressions are *kulam* and *kulam ḍay*, respectively, but these are now virtually restricted to addressing large groups of people, for example, *khaṛiya kulam*, *kulam ḍay!* ‘Kharia brothers and sisters!’.

The days of the week, months of the year, and cardinal directions have suffered the same fate. For example, *munuʔsiy* ‘east’ (literally: ‘where the sun rises’), of Kharia origin, is now virtually unknown, and *purab*, from Indo-Aryan, is used. The same applies to many common terms, such as *dinu* ‘day’ (IA) (cf. Kharia *toʔ*), *mas l māsu*, *ghos* ‘meat’ (IA) (cf. Kharia *komaṅ*), etc.

6.2.3 Loanwords from English and other languages

The large majority of English loanwords have probably made their way into Kharia via Indo-Aryan languages, such as *pablik* from English *public* but with the meaning ‘people’, or *sapoḥ* ‘support’, *risaj* ‘research’, etc. Others would seem to have taken a

different, perhaps more direct route, such as *tebul* ‘table’, perhaps through contact with the missionaries.

As Pinnow (1959:9) notes there are also a number of loanwords in Kharia from (Biblical) Hebrew, which entered the language at the time of the activities of the missionaries. This influence is largely restricted to names from the Bible, such as *yaphat* ‘Japhat’, *melki sedek* ‘Melchizedek’, etc. Similarly, the many Christian personal names such as *mariyanus*, *paulus* or *ursula*, originating in a variety of languages, were introduced by the (mostly Belgian and German) missionaries.

One name deserves special mention here due to the semantic shift it has undergone: *beljam* from English *Belgium*. The majority of Christian missionaries who worked with the Kharia were Belgians. As a result, the name *beljam* in older texts has several connotations, but apparently none of which refers to Belgium proper. *beljam* refers either to Europe, the Garden of Eden or to the ‘city of Belgium’ (in Palestine!) (cf. Pinnow 1965a:118, 125).

7 BRIEF ANALYSED TEXT

The following text⁵⁷ was chosen as it is very similar to a story which is also told by the Kurukh in Jharkhand in a slightly different version and refers to a (presumably real) battle, the details of which are no longer known.⁵⁸ Interestingly, in the Kurukh version it is the Kurukh women who fight, while in this Kharia version it is the Kharia women who fight. The text deals with the Kharia’s legendary pre-historic travels through South Asia, as they were constantly forced to flee further by their enemies, the *daqhiyal*, literally ‘the bearded ones’, whose identity is unknown. The term appears to be a generic term for non-tribal peoples.

The festival referred to here is known as Sarhul in the North Munda languages. The Kharia name is *jaṅkoy*, which means ‘spring’, that is, the ‘spring festival’ or festival of flowers. It is celebrated during the months of March/April.

Speaker: Marianus Tete (29), Thethaitangar, southwest Jharkhand

- (i) *musniṅ jaṅkoy=a? dinu aw=ki.*
 one.day spring.festival=GEN day COP=M.PST
 ‘Once it was the day of the Spring Festival.’
- (ii) *ho dinu purkha=ki osel siṅkoy=te aḍ=ki=ya?*
 that day ancestor=PL white rooster=OBLQ ANAPH=PL=GEN
sarna hinte ḍo²ḍ kon bhagwan=ḍom=ki=te ḍaṛom
 place.of.worship LOC take SEQ god=3POSS=PL=OBLQ sacrifice
o-ḍe²b=na la?²=ki=may.
 CAUS-ascend=INF IPFV=M.PST=3PL
 ‘On that day the ancestors used to take a white rooster to their place of worship and offer it up to their gods.’
- (iii) *ho dinu purkha kinir jāut=a? komay=jo jo?²=na*
 that day ancestor forest animal=GEN meat=ADD eat=INF
la?²=ki=may.
 IPFV=M.PST=3PL
 ‘On that day the ancestors also used to eat the meat of forest animals.’

- (iv) *hin=aʔ ghaʔdʔ=ga jhaʔi kōpuʔuʔ=ki kinir lam-lam tay*
 that=GEN for=FOC all man=PL forest search-REDPL ABL
ey=sikh=oʔ=ki.
 return=PRF=A.PST=PL
 ‘For just this reason all the men returned from the forest hunt.’
- (v) *ho-tiʔj golaʔ=jo purkas kui=sikh=oʔ=ki bul*
 that-side rice.beer=ADD very.much find=PRF=A.PST=PL get.drunk
kon lelem=sikh=oʔ=ki.
 SEQ fall.asleep=PRF=A.PST=PL
 ‘There they found a great deal of rice beer, got drunk and fell asleep.’
- (vi) *mudui=ki ho=ki=yaʔ raydaʔ=te yo=yoʔ=ki, ho=ki*
 enemy=PL that=PL=GEN drunkenness=OBLQ see=A.PST=PL that=PL
ghaʔ kuy=oʔ=ki.
 opportunity find=A.PST=PL
 ‘The enemies saw their drunkenness, they found an opportunity [for attacking].’
- (vii) *ho=ki gam=oʔ=ki ‘el=aʔ ghaʔdʔ ... [the rest in Sadani].*
 that=PL say=A.PST=PL 1PL.EXCL=GEN for
 ‘They said, “For us ..., if we don’t surround them now, we’ll never be able to.”’
- (viii) *jhaʔi mudui=ki hathiyar dhoʔ=ta dhoʔta aʔ=ki=yaʔ poʔda*
 all enemy=PL weapon take=CV REP ANAPH=PL=GEN village
tay muʔ=ki=may.
 ABL emerge=M.PST=3PL
 ‘All the enemies grabbed their weapons and set out from their village.’
- (ix) *khaʔiya ʔay=ki=ko buʔjhi goʔdʔ=siʔ=ki.*
 Kharia woman=PL=CNTR understand C:TEL=PRF=PL
 ‘The Kharia women understood [the situation].’
- (x) *ho=ki turthe=ga apan apan oʔiyay dhaʔb=na suru=yoʔ=ki.*
 that=PL fast=FOC RFLXV REP put.in.order SUD=INF begin=A.PST=PL
 ‘They quickly began to put themselves in order.’
- (xi) *mudui=ki=yaʔ⁵⁹ lutui su dhaph=oʔ=ki no jhaʔi=ga*
 enemy=PL=GEN clothing put.on SUD=A.PST=PL and all=FOC
kaʔ kom dhokh=oʔ=ki ro loʔimaʔe=ga
 bow arrow grab=A.PST=PL and run.here.and.there=FOC
mudui=ki=te kuday=na māʔe=yoʔ=ki.
 enemy=PL=OBLQ chase=INF begin=A.PST=PL
 ‘They put on the [men’s] clothing and all grabbed bow and arrow and, running to and fro, they began to drive back (= chase) the enemy.’
- (xii) *mudui=ki botoʔ buʔ u-tiʔj, a-tiʔj pal=oʔ=ki,*
 enemy=PL fear INS this-side Q-side be.able=A.PST=PL
hin-tiʔj=ga botoʔ-son yar=oʔ=ki.
 that-side=FOC fear-INTENS flee=A.PST=PL
 ‘The enemies fled in fear to wherever they could.’

- (xiii) *khariya day=ki ho=ki=te kuday kon hoqom siman*
 Kharia woman=PL that=PL=OBLQ chase SEQ other border
ti²j o²b-yar=na ma²re=yo²=ki.
 side CAUS=flee=INF begin=A.PST=PL
 ‘The Kharia women began to drive them off to a different region, chasing them.’
- (xiv) *jhari po²da hinte mu²du la² mu²du⁶⁰ ga²ysi aw=ta=ki=ga.*
 all village LOC 1.CLSSFR then 1.CLSSFR spy COP=M.PRS=PL=FOC
 ‘In all villages there are some spies.’
- (xv) *ho po²da=te=jo mu²du mahara day ga²ysi aw=ki.*
 that village=OBLQ=ADD 1.CLSSFR Mahara woman spy COP=M.PST
 ‘In that village as well there was a woman Mahara spy.’
- (xvi) *ho-ka² kho²i buli=na modhe bu² khariya=ki=ya²*
 that-SG.HUM village.section wander=INF means INS Kharia=PL=GEN
jhari habhaw=te erikhudi ko² may=sikh=o².
 all mannerisms=OBLQ from.bottom.to.top find.out TOTAL=PRF=A.PST
 ‘She had learned through wandering through the village all of the mannerisms of the Kharia inside and out.’
- (xvii) *ho-ka²=a² mudui=ki=ya² bo²=te col kon ho=ki=te*
 that-SG.HUM=FOC enemy=PL=GEN place=OBLQ go SEQ that=PL=OBLQ
utun go²h=o².
 speak C:TEL=A.PST
 ‘She went to the enemies’ place and spoke to them.’⁶¹
- (xviii) *ho-ka² ho=ki=te gam=o²:*
 that-SG.HUM that=PL=OBLQ say=A.PST
 ‘She said to them.’

[The spy speaks in Sadani] ‘You people, running out of fear of the women until your *dhotis* and *lungis* fell off! Those people were women! You watch – they are now going to wash their hands and legs. Now, they will collect water with both hands and wash their hands and legs. Men wash their hands and legs with only one hand.’]

- (xix) *mudui=ki yo=yo²=may la² khariya day=ki ubar ti²*
 enemy=PL see=A.PST=3PL then Kharia woman=PL 2 hand
bu²=ga guju²=na gu²the=na ma²re=yo²=ki.
 INS=FOC wash.feet=INF wash.hands=INF begin=A.PST=PL
 ‘The enemies saw then that the Kharia women began to wash their hands and legs with both hands.’
- (xx) *mudui=ki hathiyar dho²=ta dho²ta khariya=ki=te ku²day=na*
 enemy=PL weapon take=CV REP Kharia=PL=OBLQ chase=INF
ma²re=yo²=ki.
 begin=A.PST=PL
 ‘The enemies, grabbing their weapons, began to chase the Kharia away.’

- (xxi) *khariya* *ḍay=ki* *yar=na* *gu?juj* *umay* *kuy=o?* *ho=ki*
 Kharia woman=PL flee=INF path NEG.3PL find=A.PST that=PL
u-ti²j, *a-ti²j* *pal=o?=ki*, *hinte=ga* *yar=o?=ki*.
 this-side Q-side be.able=A.PST=PL there=FOC flee=A.PST=PL
u-ti²j=ko *kõpuṣu?=ki=ya?* *matwari* *chadke=ki*.
 this-side=CNTR men=PL=GEN drunkenness be.finished=M.PST
 ‘The Kharia women didn’t find the path to flee on. They fled wherever they could. On this side [i.e. over here], the men’s drunkenness came to an end.’
- (xxii) *ho=ki* *po?da=qom=ki=ya?* *halet* *onḍor=o?=ki*, *kunḍu?*
 that=PL village=3POSS=PL=GEN condition hear=A.PST=PL child
hakon=qom=ki=te *sumtay=o?=ki* *ro* *kuda kudi* *khariya*
 ECHO=3POSS=PL=OBLQ assemble=A.PST=PL and hurry REP Kharia
ḍay=ki=ya? *kunḍa²b* *kunḍa²b* *yar=o?=ki*.
 woman=PL=GEN back REP flee=A.PST=PL
 ‘They heard what was happening in their village (= the condition of their village), assembled their children and hurriedly fled after the Kharia women.’

ACKNOWLEDGEMENTS

The present study is based on the results of approximately eight months of field work conducted during five trips to Jharkhand, India. I would like to express my gratitude to the German Research Council (*Deutsche Forschungsgemeinschaft*) for two generous grants which made two of these trips possible (PE 872/1–1, 2).

I would also like to thank the many individuals in Jharkhand who made these trips both pleasant and productive. Special thanks go to Ganesh Murmu, Ram Dayal Munda, Basil Baa, Father Joachim Dungdung, Rayem Olem Dungdung, Rose Dungdung, Rose Kerketta, Saroj Kerketta, Silvester Kerketta, Anil Virendra Kullu, Anugrah Kullu, Stanislas Kullu, Tarkeleng Kullu, Anita Soreng, Dilip Soreng, Father Soreng, Mary Soreng, Pratap Soreng, Marianus Tete, Father Nicholas Tete, the students and faculty of the Department of Tribal and Regional Languages at Ranchi University, the wonderful people of Saldega, and finally to the ‘Kharia Boys’ of St Alois Minor Seminary in Ranchi: Sanjay Dung Dung, Edwin R. Dung Dung, Ajit Kerketta, Alexzendar Kerketta, Pavan Prakash Kerketta, Ravi Kujur, Ranjit Kullu, Samir Kullu, Vinay Kullu, Alok Soreng, Anil Soreng, Shilanan Tete, and of course our ‘Kurukh supporter’ Pascal Toppo.

A final word of thanks goes to Greg Anderson, Christoph Gabriel, Arlo Griffiths, Utz Maas, Ulrich Mehlem, Tatiana Oranskaia and Toshiki Osada, who somehow managed to find the time to read earlier versions of this paper and make many valuable comments and suggestions. Of course, I alone am responsible for any remaining errors and misconceptions.

NOTES

- 1 The present districts of Simdega and Gumla together formed the district of Gumla during the first year of the new state of Jharkhand, but are now separate districts, hence literature from this earlier period refers to them both together as Gumla District. These two present-day districts were also formerly a part of the Ranchi District of

- Bihar, from which Jharkhand gained independence on 15 November 2000. In older texts this area is therefore referred to as Ranchi District.
- 2 www.ethnologue.com/show_language.asp?code=KHR (May, 2005), accessed on 5 November 2007.
 - 3 Or 'Sadri', as it is referred to in the region.
 - 4 Kurukh is also often referred to as 'Oraon'. The language is also spoken in portions of southern Nepal, where it is known as 'Dhangar'.
 - 5 See note 2 above.
 - 6 Note also that the number of speakers given in the 15th edition of the *Ethnologue* for all speakers in both India and Nepal is only slightly higher than the number of speakers in India: 293,575.
 - 7 See the discussion of the free-standing form in section 3.2.8.
 - 8 *-n-* here serves to avoid a hiatus.
 - 9 The reader at this point may justifiably wonder why I have chosen to write the glottal stop at all, as it is not – or at best only marginally – phonemic. There is only one argument against this, but I feel it is important enough to override all other factors: Glottalization is highly characteristic of spoken Kharia and when Kharia is written, authors will generally indicate in some manner that a consonant in the coda is pre-glottalized and will ALWAYS use an unambiguous marker for the glottal stop – /g/ is NEVER used to represent [ʔ] and I do not believe that a transcription which does not take this into account would be accepted by speakers of the language. For this reason, I consistently indicate glottalization.
 - 10 Loanwords from Indo-Aryan generally retain their syllable structure, resulting in consonant clusters in both the onset and the coda as well as diphthongs in closed syllables.
 - 11 Similarly, where the glide *-y-* appears before other enclitics, such as *=e* '=_A.IRR' I will also consider the *-y-* to belong to the enclitic for ease of glossing.
 - 12 Sources are generally given only for examples consisting of full-length sentences. Unless otherwise noted, all examples have been adapted to the system of transliteration and glossing used here.
 Examples from my own corpus are abbreviated as follows: For example, [RD, 1:25], where RD refers to the name of the speaker, and 1:25 to the text and text-line in my corpus, respectively. These texts will be published in Peterson (in preparation). Full-length examples for which no source is given are all from my own interviews. Glosses of Sadani lexemes are based on Blain (1975), Jordan-Horstmann (1969) and Nowrangi (1956).
 - 13 In section 3.2 we see that a similar situation holds for the markers of the finite predicate (Box 9.7). In fact, a phrase such as that in (14) could also serve as the lexical base of a predicate. Cf. for example, examples (66)–(69).
 - 14 This is different in Peterson (2007). I have since been able to gather data which clearly indicate that possessive and number marking are not unique to 'NPs' but may also occur in predicates. In Peterson (2006), this topic is dealt with in considerably more depth.
 Note also that I do not assume the presence of nouns, verbs and adjectives in Kharia. For ease of presentation I will merely refer here to 'nominals'/'NPs', 'predicates' and 'modifiers'/'attributes' and will not discuss this topic in any great depth in this chapter. The reader is referred here to the brief discussion of predicate types in section 3.2 and to Peterson (2006) for more details.
 - 15 They can, however, also serve as the lexical base of a predicate.
 - 16 On the relationship between *=ki* and *=kiyar* (< **ki=bar* 'PL=2?') see Pinnow (1966:164).
 - 17 From a historical perspective, *=a?* is the original form of this suffix, which is still found in most north Munda languages either as a genitive marker or as a nominalizer/attributivizer. The form *=ya?* results from the insertion of a non-phonemic glide in

certain phonological environments which has become more or less mandatory and spread to other environments (see section 2.6 above). Other allophones are =*waʔ* (after /o/ and /u/), =ʔ after /a/ and =*naʔ* with some demonstratives, but, with the exception of the demonstratives, in all these environments as well the form =*yaʔ* is the most common form.

- 18 Malhotra (1982:110) provides an example in which the oblique marker =*te* attaches to the genitive in a so-called dative-subject environment: *ʝi=aʔ=te lemeq laʔ=ta* (1S=GEN=OBLQ sleep EMOT=M.PRS) ‘I feel sleepy’ Similarly, see example (206). Forms such as these are exceedingly rare and I have no such forms in my own corpus. Most important for the dialect I describe here however is the fact that these examples were rejected by speakers I consulted. Hence I will consider them dialectal or perhaps idiolectal forms.
- 19 Note the similarity between many of these forms and the free-standing pronominals, discussed in section 3.1.6.
- 20 Although *beʔq* would be more in line with the system I am using, ‘son’ marked for inalienable possession is consistently written *beʔt*, as it derives from Indo-Aryan *beta*, which is also its free form (i.e. when not marked for possession).
- 21 With respect to the [±animate] distinction made by Biligiri (1965): Although it is true that there is a stronger tendency to observe number agreement with an animate subject than with an inanimate subject, number agreement with inanimate subjects is quite common and its use or non-use would appear to be dependent upon other criteria, such as speaker preference, perhaps also definiteness, etc., and not formal noun classes.
Similarly, although the classifier *j(h)an* is restricted to human referents (section 3.1.8), it can always be replaced by the more common general classifier =*o*, and classifiers are never obligatory. Hence, the evidence for the opposition [±animate] is at best marginal.
- 22 For some speakers *qay* ‘woman’ is still used as a free lexeme, for others it is restricted to compounds.
- 23 Anaphoric pronouns are also found elsewhere in South Asia, for example, in Mundari (Osada, 1992:66f.), in Sinhalese (Gair, 2003:782) and in at least some Dardic languages (Bashir, 2003:845; 858).
- 24 I deal in this section only with cardinals and ordinals, as I only have significant data for these two groups. There are, however, also traces in the data in Pinnow, 1965a of native distributives, involving the ‘nominalizer’ -*nV-* (section 3.1.10), such as *mo<ne>n* ‘one each’ and *u<nu>ʔphe* ‘three each’.
- 25 The data for this second group were provided by a group of young men (all aged 17), who read them to me from their school notes. All confirmed that they had been taught these numerals in school. However, they themselves use the numerals borrowed from Sadani.
- 26 On the form *boriya*, see the discussion of the collectives.
- 27 See section 3.2.8. The use of the genitive here (*rani=kiyar=aʔ*) is not related to the preposition *enem* but marks the ‘subject’ of a free-standing form.
- 28 There are also a number of borderline cases which I cannot discuss further here, such as the two highly speaker-specific locative markers *hinte* and *boʔte*, which show traits of nominals, postpositions and case markers. See Peterson (2006) for details.
- 29 For a discussion of this marker from a broader Munda perspective, see Zide (1968).
- 30 It is interesting to note here that the three terms which have not been replaced by the corresponding Sadani terms are the three left-most colour terms in the two hierarchies of basic colour terms found in Berlin and Kay (1969:104) and the term for ‘yellow’, which has only recently been borrowed, immediately follows these three in one of their two hierarchies.
- 31 To express stative predication, either a nominal or the free-standing form is used together with the copula. See section 4.1 on this use and the respective forms of the copula.

- 32 As Arlo Griffiths (p.c.) has pointed out to me, there is a similar alternation with many borrowed elements in Gutob (Zide, 1985:97ff.) between *-ei* and *-a*, although Sadani is not spoken in this area. As I am not familiar with the developments of earlier forms of the eastern Indo-Aryan languages, I will not pursue this topic further here, but it seems evident that the marker *-e/-y* in Kharia has been borrowed from some Indo-Aryan source, along with the respective lexemes.
- 33 According to Jordan-Horstmann (1969:56f), *-ā* is a causative marker, in addition to its functions as a verbalizer, causative passive and reflexive marker. According to Nowrangi (1956:115), the same suffix is also a marker of the simple passive. Cf. (from Nowrangi, 1956:114f.) *sun-* ‘hear’, *sun-ā-* ‘be heard’, *piy-* ‘drink’, *piy-ā-* ‘cause to drink’ and *nāc-* ‘dance’, *nac-ā-* ‘be made to dance’. While not especially common, there are a number of languages where predicates marked by a specific morpheme may be interpreted either as passive or causative, such as Korean (cf. Keenan, 1985a:262, citing another source) as well as a number of Siberian languages (Greg Anderson, p.c.). The Sadani data appear more complex, since *-ā* can be interpreted as both passive and causative, but also as causative passive. This topic requires further study.
- 34 This is not true of all predicates here, though. Cf. Kharia *coray* ‘steal’ from Sadani *cor-ā-* ‘steal’, which derives from the noun *cor* ‘thief’ (cf. Jordan-Horstmann, 1969:56).
- 35 With regard to the perfect, the present perfect is morphologically unmarked for active and middle voice, as well as tense. However, the past and irrealis perfect are marked for both basic voice and tense. The optative and ‘Past II’ never mark for basic voice.
- 36 Recall from section 2.6 that a stem-final plosive is devoiced and aspirated before the past active marker *=oʔ*. If the stem-final is the glottal stop, this becomes /kh/ before /oʔ/. Hence /siʔ/ + /oʔ/ = /sikhoʔ/. Note, however, that for this analysis to hold, we must assume that the perfect marker has the underlying form /siʔ/ and not /siq/.
- 37 Alternatively, one could analyse this as the ‘Past II’ marker /kh/ and the past active marker /oʔ/. I will retain the analysis here of this form as /khoʔ/, unmarked for basic voice, as there is no corresponding middle form.
- 38 As noted in note 36 above, the fact that the perfect has the form *=sikh* before the active past ending *=oʔ* suggests that the underlying form is *siʔ* (< /sig/) and not *siʔd* (= /siq/). This topic is dealt with in Peterson (2006).
- 39 But see the closely related departive *tu* and the ambulative *san* in section 3.2.10.
- 40 Unless it is marked explicitly for iterativity, see section 3.2.10.
- 41 See note 28 above on *boʔte*.
- 42 This would seem to be a very old Munda construction. For similar data in Santali, see Neukom (2001:176f.). See also under the discussion of root types in Remo (Fernandez, 1968:37) stems of the type *Root-oʔ-Root*, where *-oʔ* is homophonous with – if not identical to – the marker of the active past, referred to by Fernandez as ‘Class I’, past. It is perhaps significant that the past active form appears to have been generalized in Remo and that this construction in Kharia is much more common in the past and does not occur in the middle voice.
- 43 *=ga* otherwise functions as a focus particle. Similarly, an ‘imperfective converb’ marked by *=na* can appear in the oblique case, suggesting that this marker is in fact the infinitival marker.
- The status of *=ta* as a converbal marker is much clearer: Since the present active marker *=te* can never be used in this function and since the form marked by *=ta* cannot mark for person, this *=ta* is clearly not the same marker as the marker of the present middle, although I assume that is what it derives from.
- 44 For example, as a non-productive derivational suffix of indeterminate meaning, e.g., with nominals: *konsel* ‘girl’ vs. *konselquʔ* ‘woman’. With finite predicates its meaning is more predictable and denotes a prolonged action: *yo* ‘see’, *yoquʔ* ‘look at, stare’, although it is not compatible with many lexemes with this meaning. With typical property lexemes, it is highly productive and its meaning is entirely

- predicable: ‘-ish’, for example, *rusuŋ* ‘red’, *rusuŋ-đu?* ‘reddish’. Its status as a participial marker is hence very uncertain.
- 45 Due to the large number of these categories, not all can be dealt with here.
- 46 On the status of the overtly expressed NP as the subject of the clause, see sections 3.2.1, 3.2.9 and 5.
- 47 The plural marker =*ki* is also used to indicate approximation, that is, ‘here and there; more or less these places’. Other examples include *ida?*=*ki* ‘yesterday or so’, *tuɖa*=*ki* ‘tomorrow or so’.
- 48 However, Pinnow also writes that *jom* still means ‘eat’ in Kharia, whereas all speakers I consulted rejected this, and for them the form no longer has any independent meaning. Note also that Pinnow considers it an alternant form of the passive marker *ɖom*, although the two have very distinct functions.
- 49 No longer used as a free morpheme but found in a number of forms related to ‘earth’: *lo?kha* ‘earth, dirt’, *uslo?* ‘land’, *go?jlo?* ‘rice field’, *roke?ɖlo?* ‘desert’ (cf. *roke?ɖ* ‘sand’).
- 50 While not a free lexeme nor productive, *-taŋ* appears in a number of words related to cows: *koŋtaŋ* ‘cow’, *dimtaŋ* ‘stall’, *giniŋ-taŋ* ‘bride price’ (formerly paid in cattle), etc. A comparison with other South Munda languages also shows that it must have meant ‘cow’ at some earlier time.
- 51 I follow here the terminology in Keenan (1985b [1995]) in which the ‘relative pronoun’ in the subordinate clause is referred to as the co[r]relative. The pronominal element in the main clause in Kharia is either a simple pronominal or a demonstrative of some type and will be glossed as such. This is in distinction to most studies on South Asian languages, where what I refer to as the correlative is termed the ‘relative pronoun’ and the demonstrative element in the main clause the ‘correlative’.
- 52 The wife is speaking to her husband about his parents.
- 53 Or rather, the use of the converb is restricted to subordinate clauses when *the finite clause corresponding to the subordinate clause* would have the same subject as the main clause. The same holds *mutatis mutandis* for different subjects: If the finite clause corresponding to the subordinate clause would have a different subject than the main clause, then the free-standing form is preferred, and the element corresponding to the subject of the corresponding finite version of the subordinate clause appears in the genitive.
- 54 Very rarely, *la?* can also denote that an action began instead of imperfectivity. This topic is discussed in greater detail in Peterson (2006).
- 55 In a brief (non-representative) text consisting of five pages, 54 out of 121 lexemes (= 44.6%) were Indo-Aryan borrowings. In fact, this figure would seem to be too low, as the text was written especially for me and contained a number of forms no longer in current use in the spoken language, such as the native cardinal numbers with which most speakers are not even familiar. However the actual *frequency* of loanwords in texts appears to be much lower than this, as the most common lexemes are usually of Kharia origin, but as yet I have no figures for frequency.
- 56 Roy and Roy (1937:38) quote a Dhelki Kharia speaker who claims that this is one of the differences between Dudh and Dhelki Kharia. According to him, the Dudh Kharia say *cholla-ki* for ‘he went’ [translated by the authors as ‘they come’, JP], whereas the Dhelki Kharia say *sanning*. The speaker seems to be referring to the past imperfective form *cona la?ki* ‘he was going / used to go’ in Dudh Kharia. There is no form **col la-ki*, and the simple past form is *colki*.
- The Dhelki Kharia form *sanning* seems to have been misrecorded. *san* undoubtedly means ‘go’, and *-niŋ* is the ending for the first person, plural, exclusive. What would seem to be missing here is a TAM-marker.
- 57 This text consists of lines 194–218 of a text which appeared in full in Peterson (2006). I present it here in a normalized orthography.

- 58 It is also referred to in passing in Roy and Roy (1937:220f.). Cf. also Tirkey (1998: 28f.), concerning the title of a Kurukh-language periodical, *Sinagi Dai*, and the story behind its name.
- 59 This is undoubtedly a slip of the tongue, as the women put on their men's clothing, not that of the enemy.
- 60 *muḍu la? muḍu* is an idiomatic expression meaning 'one or another, some'. Note also that the present general imperfective form *aw=ta=ki* is used instead of the present copula *ayi?j=may*, as a habitual situation is being portrayed, not a present state.
- 61 The use of *=a?* as a focus marker is quite common, although speakers I consulted, even many who themselves use it, denied using it or even being familiar with it and corrected their recorded texts. I was not able to question this speaker on this topic. Like *=ga, =a?* appears to derive from the 'full' particle form *=ga?* 'FOC'.

BIBLIOGRAPHY

- Abbi, Anvita (1993) 'Language contraction, language shrink, and language conflation: a case study of Kharia', in M. Miri (ed.) *Continuity and Change in Tribal Society*, Simla: Indian Institute of Advanced Study, 542–553.
- (1997) 'Languages in contact in Jharkhand', in A. Abbi (ed.) *Languages of Tribal and Indigenous Peoples of India. The Ethnic Space*, Delhi: Motilal Banarsidass, 131–148.
- Anderson, Gregory D.S. (2001) 'A new classification of South Munda: evidence from comparative verb morphology', *Indian Linguistics*, 62:21–36.
- Anderson, Gregory D.S. and Norman H. Zide (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: the Bimoraic constraint', in M.A. Macken (ed.) *Papers from the 10th annual meeting of the Southeast Asian Linguistics Society*, Tempe, AZ: Arizona State University, South East Asian Studies Program. (Monograph Series Press), 55–74.
- Archer, W.G. (ed.) (1942) *khaṛiyā aloṅ. A Kharia Song Book*, Ranchi: Catholic Press.
- Ba?, Julyus (1983) *khaṛiyā dhvaniśāstra. Kharia Phonetics*, Ranchi: Merā Silām. [In Hindi]
- (2001) *khaṛiyā dhvani sārācnā kā vaijnānik viśleṣan*, [In Hindi: *A Scientific Analysis of the Structure of Kharia Sounds*], Ranchi: Dāso Khaṛiyā.
- Banerjee, Gagan Chandra (1894) *Introduction to the Khariā Language*, [Reprint: New Delhi/Chandigarh: Bahri Publications, 1982].
- Bashir, Elena (2003) 'Dardic', in Cardona and Jain (eds), *The Indo-Aryan Languages*, London/New York: Routledge (Routledge Language Family Series), 818–894.
- Berlin, Brent and Paul Kay (1969) *Basic Color Terms: Their Universality and Evolution*, Berkeley: University of California.
- Bhagat, Sureś (2001) 'hil-khaṛiyā [In Hindi: 'Hill Kharia']', *Bulletin of Jharkhand Tribal Welfare Research Institute, Ranchi*, 39: 21–26.
- Biligiri, H.S. (1965) *Kharia. Phonology, Grammar and Vocabulary* (Building Centenary and Silver Jubilee Series:3), Poona: Deccan College, Postgraduate and Research Institute.
- Blain, Father Edgar (1975) *English-Sadri Dictionary*, Jharsuguda: Society of the Divine Word.
- Bloomfield, Leonard (1933) *Language*, New York: Holt, Rinehart, and Winston, [Reprint: Chicago: University of Chicago, 1984].
- Cardona, George and Dhanesh Jain (eds) (2003) *The Indo-Aryan Languages*, London/ New York: Routledge (Routledge Language Family Series).

- Duṇḍuṇ, Phā. Jovākim. (1999) *khariyā jīvan aur paramparāē* [In Hindi: *Kharia Life and Traditions*], Ranchi: Ranchi Society of Jesus.
- Duṇḍuṇ, Mathiyas (1986) *hindī aur khariyā (tulnātmak aur viślesṇātmak adhyāyan)* [In Hindi: *Hindi and Khariya. A Comparative and Analytical Study.*], Ranchi: Satya Bhārafi.
- Fernandez, Frank (1968) *A Grammatical Sketch of Remo: A Munda Language*, Unpublished PhD dissertation, Chapel Hill, North Carolina.
- Floor, H., S.J., V. Gheysens, S.J. and G. Druart, S.J. (1934) *Dictionary of the Kharia Language English-Kharia, Kharia-English*. (Tea Districts Labour Association), Calcutta: Catholic Orphan Press.
- Gair, James W. (2003) 'Sinhala', in Cardona and Jain (eds), *The Indo-Aryan Languages*, London/New York: Routledge (Routledge Language Family Series), 766–817.
- Grierson, G.A. (1906) *Linguistic Survey of India. Vol. IV. Munda and Dravidian Languages*, Prepared by Sten Konow. [Reprint: Delhi: Low Price Publications, 1994].
- Grimes, Barbara F. (ed.) (1988) *Ethnologue. Languages of the World*, 11th edition, Dallas: Summer Institute of Linguistics.
- Hoffmann, Revd J., S.J. (1903) *Mundari Grammar and Exercises* [Reprint: New Delhi: Gyan Publishing House, 2001].
- Jordan-Horstmann, Monika (1969) *Sadani. A Bhojpuri Dialect Spoken in Chotanagpur*, Wiesbaden: Harrassowitz (Indologia Berolinensis, 1).
- Keenan, Edward L. (1985a) 'Passive in the world's languages', in T. Shopen (ed.) *Language typology and syntactic description. Volume I: Clause Structure*, Cambridge: Cambridge University Press, 243–281.
- (1985b) 'Relative clauses', T. Shopen (ed.) *Language Typology and Syntactic Description. Volume II: Complex Constructions*, Cambridge: Cambridge University [Reprint: 1995]. 141–170.
- Kerkeṭṭā, Khrist Pyāra. (1990) *Jujhair dāṛ (khariyā nāṭak)* [In Kharia: *The Battle Field (A Kharia Drama)*], Ranchi: Tribal Language Academy, Government of Bihar.
- Kullū, Paulus (1981) *khariyā-vyākaraṇ evā sākskṛipt śabdakoś* [In Hindi: *Kharia Grammar and Abridged Dictionary*], Ranchi: Religious Literature Society.
- (translator) (1992) *tonme kol-ajre dharam pothī* [In Kharia: *The New Testament*], Ranchi: Catholic Press.
- (2000) *khariyā dharam aur sāskṛti kā viślesan* [In Hindi: *An Analysis of Kharia Religion and Culture*], Ranchi: Satya Bhārafi.
- Maas, Utz (2004) '“Finite” and “nonfinite” from a typological perspective', *Linguistics*, 42(2): 359–385.
- Mahapatra, B.P. (1976) 'Comparative notes on Juang and Kharia finite verbs', P.N. Jenner, Laurence C. Thompson and Stanley Starosta (eds) *Austroasiatic Studies* (Oceanic Linguistics, Special Publication, No. 13), Honolulu: University of Hawaii, 801–814.
- Malhotra, Veena (1982) *The Structure of Kharia: A Study of Linguistic Typology and Language Change*, Unpublished PhD dissertation, New Delhi: Centre for Linguistics and English, School of Languages, Jawaharlal Nehru University.
- Neukom, Lukas (2001) *Santali*, München: Lincom Europa (Languages of the World/Materials, 323).
- Nowrangi, P.S., S.J. (1956) *A Simple Sadani Grammar*, Ranchi: D.S.S. Book Depot.
- Osada, Toshiki (1992) *A Reference Grammar of Mundari*, Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Peterson, John (2002) 'Morphological and semantic finiteness in Kharia: a first look', in M. Bommes, C. Noack and D. Tophinke (eds) *Sprache als Form. Festschrift für Utz Maas*, Wiesbaden: Westdeutscher Verlag, 63–73.

- Peterson, John (2006) *Kharia. Grammar, Texts and Glossary. Three volumes*, Unpublished Habilitationsschrift, Osnabrück: Universität Osnabrück.
- (2007) ‘Languages without nouns and verbs? An alternative to lexical classes in Kharia’, in Colin Masica (ed.) *Old and New Perspectives on South Asian Languages. Grammar and Semantics. Papers growing out of the Fifth International Conference on South Asian Linguistics (ICOSAL-5), held at Moscow, Russia in July 2003*, Delhi: Motilal Banarsidass, (MLBD Series in Linguistics, XVI), 274–303.
- (in press) ‘The “totality” morpheme *-may* in Kharia’, to appear in Ram Dayal Munda (ed.) *Felicitation Volume for Norman Zide*.
- Pfeffer, Georg (1993) ‘Kharia totemism’, *Journal of the Indian Anthropological Society* 28: 221–227.
- Pinnow, Heinz-Jürgen (1959) *Versuch einer historischen Lautlehre der Kharia-Sprache*, Wiesbaden: Harrassowitz.
- (1965a) *Kharia-Texte (Prosa und Poesie)*, Herausgegeben und übersetzt unter Mitwirkung von Paulus Kerketta und Junas Kullu, Wiesbaden: Harrassowitz.
- (1965b) ‘“Der Wertiger” und andere Geschichten in Kharia, aufgezeichnet von Junas Kullu, übersetzt und mitgeteilt unter Mitwirkung von Paulus Kerketta’, *Indo-Iranian Journal*, 9(1): 32–68.
- (1966) ‘A comparative study of the verb in the Munda languages’, N.H. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, London/The Hague/Paris: Mouton and Co. (Indo-Iranian Monographs, V), 96–193.
- Rehberg, Kerstin (2003) ‘Phonologie des Kharia – Prosodische Strukturen und segmentales Inventar’, Unpublished MA thesis, Osnabrück: Universität Osnabrück.
- Roy, Sarat Chandra and Ramesh Chandra Roy (1937) *The Khāriās*, Ranchi: Man in India.
- Sāhu, R.P. (1979/1980) *khaṛiyā vyākaraṇ* [In Hindi: *Kharia Grammar*], Ranchi: Kharia Department, Gossaner College.
- Sinha, Dikshit (1984) *The Hill Kharia of Purulia. A Study on the Impact of Poverty on a Hunting and Gathering Tribe*, Calcutta: Anthropological Survey of India, Government of India, (Memoir, Anthropological Survey of India, 59).
- Tea Districts Labour Association (1929) *Language Hand-Book Kharia. Printed for Private Circulation Only*, Calcutta: Begg, Dunlop and Co.
- Tirkey, L. (1998) *Tribals, their Languages and Literature*, Ranchi: Don Bosco Publishers.
- Vidyarthi, L.P. and V.S. Upadhyay (1980) *The Kharia: Then and Now. A Comparative Study of Hill, Dhelki and Dudh Kharia of the Central-Eastern Region of India*, Delhi: Concept Publishing Company.
- Yadav, Ramawatar (1996) *A Reference Grammar of Maithili*, Berlin/New York: Mouton de Gruyter (Trends in Linguistics, Documentation, 11).
- Zide, Norman H. (1968) ‘Some comparative notes on Gata? pronouns’, in J.C. Heesterman, G.H. Schokker, V.I. Subramoniam (eds) *Pratidānam. Indian, Iranian and Indo-European Studies Presented to Franciscus Bernardus Kuiper on his Sixtieth Birthday*, The Hague/Paris: Mouton, 348–358.
- (1969) ‘Munda and non-Munda Austroasiatic languages’, in T.A. Sebeok (ed.) *Current Trends in Linguistics. Vol. 5: Linguistics in South Asia*, The Hague: Mouton, 411–430.
- (1985) ‘Notes mostly historical on some participant roles in some Munda languages’, in A.R.K. Zide, David Magier, and Eric Schiller (eds) *Proceedings of the Conference on Participant Roles: South-Asia and Adjacent Areas*, Bloomington: Indiana University Linguistics Club, 92–103.
- Zide, Norman H. and D. Stampe (1968) ‘The place of Kharia-Juang in the Munda Family’, in B. Krishnamurti *Studies in Indian Linguistics* (Professor M.B. Emeneau Saṣṭipūrṭi Volume), 370–377.

ANNOTATED BIBLIOGRAPHY

The Kharia are one of the best studied Munda groups and studies on their language and culture date back to the nineteenth century. There are five works in Western languages (English and German) which must be considered standard works on the Kharia and their language: Malhotra (1982), Pinnow (1959; 1965a,b) and Roy and Roy (1937).

Malhotra (1982), the author's Ph.D. dissertation, is a comprehensive grammar of Kharia, dealing extensively with all aspects of the language, except phonology, and including a chapter on language change. Unfortunately, the dissertation is difficult to obtain, as it has not been published.

Pinnow (1959) deals with Kharia phonology from a predominantly historical perspective and is standard literature for anyone interested in historical Munda and Austroasiatic linguistics. Pinnow (1965a) is a large collection of texts in Kharia, accompanied by a word-for-word translation and a free translation and followed by detailed notes. Pinnow (1965b) is a shorter collection of the same type but without the word-for-word translation. These last two works provide a valuable resource for anyone wishing to study almost any aspect of Kharia, including language change to a certain extent, as these texts are now over 40 years old.

Roy and Roy (1937) is a treasure-house of information on the Kharia from the first half of the twentieth century and is based on twelve years of research among all three Kharia groups (Dudh, Dhelki and Hill Kharia) by S.C. Roy, often referred to as the 'father of Indian ethnography', and his son. This work (two volumes) contains information on virtually all aspects of Kharia life at that time, with the exception of the language. It stands alone in its scope in Munda ethnography and is so detailed that it has even led one modern scholar to write of it that 'Later ethnographic accounts have nothing significant to add' (Pfeffer 1993:222).

Other studies which should be mentioned in this context are primarily of interest from a historical perspective, such as Banerjee (1894 [1982]), the first attempt to describe the language systematically, Tea Districts Labour Association (1929), a short but remarkably precise description of the language, Floor *et al.* (1934), also from the Tea Districts Labour Association, a highly useful and compact English–Kharia, Kharia–English dictionary and to date the only one published, and Biligiri (1965), the first full-length study of the language. While none of these is up-to-date, they have played an important role in the history of Kharia linguistics.

There are also a number of more recent works on the Kharia people and their language in English and German, such as Abbi (1993; 1997), Mahapatra (1976), Peterson (2002, 2006, 2007, in press), Pfeffer (1993), Rehberg (2003), Sinha (1984), Vidyarthi and Upadhyay (1980) and Zide and Stampe (1968).

The number of works in Hindi on the Kharia and their language is also significant: Ba? (1983; 2001) deal with Kharia phonology; Bhagat (2001) is a brief, six-page overview of the Hill Kharia; Đuᅇᅇ (1986) contains a brief grammar, a large number of Kharia stories and songs, all accompanied by a Hindi translation, as well as a Hindi–Kharia and Kharia–Hindi glossary; Đuᅇᅇ (1999), based on the author's own interviews with Dudh and Dhelki Kharia, contains a large amount of information on all aspects of Kharia life; Sāhu (1979/80) and Kullū (1981) are concise grammars of Kharia, the latter also with a glossary, and finally Kullū (2000) deals extensively with Kharia religion and culture.

Most of these authors who have written in Hindi are themselves Kharia, so that one can also speak of an indigenous scholarly tradition among the Kharia.

In fact, Kharia (culture, language, literature and history) can also be studied at the Department of Tribal and Regional Languages at Ranchi University, where many of the courses are taught by Kharia.

There is also a growing amount of literature in Kharia, beginning with Archer (1942), a collection of songs and riddles in Kharia, although unfortunately only some of these have been translated (see the indices in Pinnow 1965a:284, 286). There are also a number of Bible translations, or rather partial translations, song-books for the church service and other religious texts, as well as a translation of some of Premchand's works. Original work in Kharia is also on the increase, ranging from poetry and short stories to drama (e.g. Kerkeṭṭā 1990, from which many examples in this work were taken).

For a more complete list, see: www.SouthAsiaBibliography.de/Bibliography/Austroasiatic/Munda/Kharia/kharia.html

JUANG*

Manideepa Patnaik

1 INTRODUCTION

Juang,¹ belongs to the South Munda subgroup of the Western Austroasiatic language family. The other names of Juang are Patua, Puttoas, Patra and Juango. According to an analysis by Zide (1969) Juang was paired with Kharia to form the Central Munda language subgroup.

Juang is spoken in southern Keonjhar, northern Angul and eastern Dhenkanal districts of Orissa state by about 30,875 speakers (2001 census). According to the *Ethnologue* version of 2000, the number of speakers is 50,000.

Juang is a close cognate of Kharia. As per the *Ethnologue*, Juang shares lexical similarity of about 20–22% with Kharia. There is still a lot of work pending on Juang and its relative position in the Munda subgroup. One can conjecture at this point, that Juang may be an isolate language of the Munda group (as suggested by Anderson 2001). The neighbouring Indo-Aryan language Oriya has heavily influenced the Juang language. The influence has resulted in the loss of glottal stop, Oriya versions of relative clause formation, quotative constructions and anaphora system among others. Juang still has some idiosyncratic patterns which are very interesting and some other patterns which are very typically Munda.

Juang doesn't have a script on its own. When written, it uses the Oriya script. The earliest reference to Juang can be found in Grierson's Linguistic Survey of India and Sten Konow's brief sketch of the language. Other work on various aspects of Juang is found in Pinnow (1960), Mahapatra (1962a, b, c), Matson (1962), Dasgupta (1978) and Patnaik (2000). The Academy of Tribal Dialects and Culture has prepared language-learning materials and short documents on almost all Munda languages that cover in brief the phonology and morphology of the languages along with some useful lexicon, songs and folk tales. The one on Juang is authored by Sashmita Mohapatra (1991). Much has been done on the cultural, anthropological and sociological aspects of this tribe some of which have been published and a lot in dissertations. A magazine called 'Banakua' edited by Upendra Mishra has published a brief sketch of the language too which is currently unavailable. The sketches available provide pronominal tables and verb paradigms with few examples of simple sentences and complex sentences that run to a couple of pages with a number of un glossed folk tales, folk songs and relatively good lexica.

Juangs can broadly be divided into Plains Juangs and Hill Juangs. Hill Juangs live in Gonasika and Pallara hills whereas the Plains Juangs live in about 147 villages in and around Keonjhar and Dhenkanal districts. Speakers of this language use their mother tongue not only at home but also with their peers in school and other places in the community. They look up to their culture and have a very positive attitude towards their language. Language proficiency is pretty impressive even in the case of bilingual or multilingual Juangs though it is higher among the older generation.

The medium of education is Oriya though the government has prepared a few Juang primers for children and older people.

2 PHONOLOGY

2.1 Vowel inventory

Juang has a six-vowel inventory system. The vowels in Juang are /i/, /e/, /a/, /ɔ/, /o/ and /u/. /e/ has a variant of /ɛ/. Though not many contrastive pairs are found between /ɔ/ and /o/ there are cases where /o/ is used and not /ɔ/ and vice versa. /a/ is sometimes realized as a schwa.

- (1) *Vowel inventory*
- | | |
|---|---|
| i | u |
| e | o |
| a | ɔ |

Glide formation. The two vowels /u/ and /i/ have a tendency towards glide formation when they precede another vowel. Some examples are given below:

- (2)
- | | |
|-------------------------------|------------|
| <i>arɔkia</i> – <i>arɔkya</i> | ‘they two’ |
| <i>uaʎi</i> – <i>waʎi</i> | ‘child’ |
| <i>kui</i> – <i>kwi</i> | ‘get’ |
| <i>kua</i> – <i>kwa</i> | ‘handful’ |
| <i>dui</i> – <i>dwi</i> | ‘two’ |

Diphthongs. All vowels can become diphthongized when they are followed by any palatal stops or palatal nasals. For example:

- (3)
- | | |
|--|------------|
| <i>aiŋl</i> <i>ayny</i> | ‘I’ |
| <i>deŋceɔ</i> / <i>deyŋceiɔ</i> | ‘had come’ |
| <i>niŋ</i> / <i>niyŋy</i> | ‘we’ |
| <i>puɕuka</i> / <i>puicwika</i> | ‘boil’ |
| <i>ɔbgɔj</i> / <i>ɔbgɔy²j</i> | ‘kill’ |

2.2 Suprasegmental phenomena

Stress, juncture, tones, and registers are not phonemic in Juang though they are present in the language to be used for various contextual and stylistic purposes. Actually these have not yet been investigated to any degree.

Vowel length. There are very few instances where vowel length is noticed to be phonemic. Vowels often get lengthened in tense, aspect, mood categories (Patnaik (in preparation) has details). In the past tense the vowel is always lengthened irrespective of which subject it takes.² Sometimes borrowing encourages vowel length. Most of the time lengthening is used for the formation of imperative construction as the following examples exhibit:

- (4)
- | | |
|--------------|--|
| <i>jɔyɔ-</i> | ‘saw’ (past tense where vowel is lengthened) |
| <i>jeel-</i> | ‘jail’ (borrowing where vowel is lengthened) |

- (5) *kɔ-* ‘shave’ vs. *kɔɔ!* ‘shave!’
unɔ- ‘keep’ vs. *unɔɔ!* ‘keep!’

Vowel nasalization. All vowels can be nasalized in Juang. There are some cases where even nasalization is seen to have phonemic status.

- (6) *tɔɔɔ* ‘I fastened’ vs. *tɔ̃ɔ̃ɔ* ‘elephant’s trunk’
kɔnia ‘marriageable’ vs. *kɔ̃ɲia* ‘safe’

Nasal consonants of the preceding syllable allow vowel nasalization.

- (7) *ɡuɲtia* / *ɡũɲtia* / *ɡuɲtiã* / *ɡuɲtiã̃*

2.3 Consonants

The consonant system in Juang follows the typical South Munda pattern. The genuine consonant inventory of Juang doesn’t have voiced stops and aspirated stops. With increasing borrowing it is now possible to find aspirations in Juang. Unlike most Munda languages, glottal stops are infrequently noticed in Juang.

- (8) Consonant inventory

Bilabial	Dental	Retroflex	Palatal	Velar
<i>p</i>	<i>t</i>	<i>ʈ</i>	<i>c</i>	<i>k</i>
<i>b</i>	<i>d</i>	<i>ɖ</i>	<i>j</i>	<i>g</i>
<i>m</i>	<i>n</i>	<i>ɳ</i>	<i>ɲ</i>	<i>ŋ</i>
	<i>s, l</i>	<i>ʂ</i>		
	<i>r</i>		<i>y</i>	

Contrastive pairs. The following examples exhibit some contrastive pairs indicating the different phonemic status of the phonemes described above.

- (9) *k – g* *kɔŋ* ‘know’ *ɡɔn* ‘weave’
k – c *kɔke* ‘shaves’ *cake* ‘taste’
c – j *cake* ‘taste’ *jɔge* ‘watch’
t – ṭ *tɔke* ‘cheat’ *tɔke* ‘tired’
t – d *tɔke* ‘tired’ *dɔkɔ* ‘sit’
p – b *page* ‘break’ *bɔge* ‘command’
b – m *bɔge* ‘command’ *mɔtej* ‘nose’
r – l *rɔge* ‘pour’ *lage* ‘continue’

Retroflex stops are rarely found in word-final position. Word-medial //l/ is often replaced by /ʈ/ as in *cɔncɔʈɔ* → *cɔncɔʈɔ* ‘quickly’ or *kiʈɔg* → *kiʈɔg* ‘tiger’. However, *ɔʈpɔ* which is an Oriya loan is never heard as **ɔʈpɔ* ‘few’ nor *turaʈi* as **turaʈi* ‘sword’ nor *buʈug* as **buʈug* ‘spade with a straight narrow blade’.

Sound alternations. In certain other instances some sounds alternate with other sounds in a particular environment and at other times, there appear to be free or context-less variations. Note that there is considerable inter-speaker variation in the realization of many sounds in Juang: both consonants and vowels. This variation is faithfully represented in the transcriptions throughout this chapter. Apart from the ones enumerated below, this includes variation between palatal and

dental-alveolar and retroflex and dental–alveolar nasals and obstruents. Examples of some of the often–found morphophonological changes and alternations are given below:

- (10) (a) /d/ changes to /r/
Example – the definite marker /dɛ/ > /rɛ/
(b) /j/ changes to /y/
Example – the negative /jenal > /yenal
(c) /e/ alternates with /ɛ/
Example – /dake/ meaning ‘calling’ > /dake/
(d) /o/ alternates with /ɔ/
Example- /arol meaning ‘he’ > /arɔ/

Gemination of consonants is normally found word-medially in disyllabic words. If the word has more than two syllables it is found between the ultimate and penultimate syllables of the word.

- (11) *ape* / *appe* ‘you pl.’ (*apperiki* ‘you all’)
lubbɔ / *lubbɔɔ* ‘greed’
kəkəm / *kəkəkəm* ‘arrow’
dukɔ / *dukkɔ* ‘sorrow’

Sibilants and the liquid /r/ are excluded from gemination:

- (12) *dusɔ* / **dussɔ*
bɔɛɔ / **bɔɛɛɔ*

Homorganic nasal formation is very often noticed in Juang. Examples where homorganic nasal formation is followed by the language are given below:

- (13) [ŋg] *ɔŋgi* ‘shirt’
ekɔlɔŋg ‘now’
[ŋc] *pəŋcgɔtɔ* ‘five sections’
kɔŋcelan ‘girl’
[ŋj] *kɔŋje* ‘arrange’
[ŋt] *eŋtej* ‘small’
[ŋd] *aŋdara* ‘white’

2.4 Syllable structure

There are monosyllabic, dissyllabic as well as multisyllabic words in Juang. The minimal syllable structure of a word could be monosyllabic which would just be a vowel. A monosyllabic word that is just a vowel is often a marker having grammatical relevance which is always hosted by a word. I provide some examples below of words with various syllable structure where V stands for ‘vowel’ and C stands for ‘consonants’.

- (14) Monosyllabic words
a ‘of, at’ (as in *ina* ‘of the house’)
u ‘but’
i ‘be’

(15) Disyllabic words

V.CV	<i>ɔle</i>	'mango'
V.CVV	<i>ɔrei</i>	'mango'
V.CVC	<i>ɔmɔr</i>	'eye'
VV.VC	<i>ɔek</i>	'yes'
VV.CV	<i>aera</i>	'not blossoming'
VV.CVC	<i>alilaŋ</i>	'above'
VC.CV	<i>aŋtu</i>	'knee'
VC.CVC	<i>ɔnsɔm</i>	'jack fruit'

(16) Trisyllabic words

V.CV.CV	<i>ariga</i>	'curry'
V.CV.CVC	<i>asimɔr</i>	'eye lash'
V.CVC.CV	<i>aciŋka</i>	'an event three days old'
VC.CV.CV	<i>ɔmbaɬɔ</i>	'two'
CV.CV.CV	<i>kanati</i>	'mop the floor'
CV.CV.CVC	<i>barosiŋ</i>	'wall'
CV.CVV.V	<i>sɔmɔɔ</i>	'tolerate'
CV.CVV.CV	<i>kaniare</i>	'bride'
CV.CVC.CVC	<i>leraŋtɔr</i>	'moon light'
CV.CVC.CCV	<i>sumundrɔ</i>	'sea'
CVV.V.CV	<i>bɔɔsɔ</i>	'age'
CVV.CV.CV	<i>bɔisakɔ</i>	'old'
CVV.CV.CVC	<i>kuirɔsoŋ</i>	'smoke'
CVC.CV.CV	<i>gɔndɔŋɔ</i>	'dirty water'
CVC.CV.CVV	<i>laŋkurei</i>	'touch-me-not plant'
CVC.CV.CVC	<i>kɔŋkulaŋ</i>	'butterfly'
CVC.CVC.CVC	<i>kɔŋjimɔak</i>	'star'
CCV.CV.CV	<i>sraɔɔnɔ</i>	'rain'
CCVC.CVC.CVC	<i>grɔŋdaksɔr</i>	'sweating'

The contrast between voiced and voiceless stops is not found word-finally. When the word ends with a glottalized consonant, both voiced and voiceless stops occur in free variation. Glottalization is no longer phonemic in Juang. It looks like a sociolinguistic variable the details of which remain to be worked out. Glottal stops occur before voiceless stops. Consider the examples in which the distinction between voiced and voiceless stop is neutralized and there is a glottal stop before the voiceless stop or a pre-glottalized release as is found in many other Munda languages.

(17) *selɔ²k / selɔŋ* 'dog'

2.5 Intonation/stress

Not much work has been done on intonation/stress of Juang. My study which largely focuses on morphological and syntactic aspects of the language shows that speakers of this language often express interrogation, surprise, exclamation, and sadness with the use of intonation, but this has not been systematically studied. Intonation falls on the last word of the sentence which is a verb.

- (18) *ju lukɔ-re-te aɪn abujɔ arɔ am-te jɔyɔ*
 who man-DEF-ACC I see:PST he you-ACC see:PST
 ‘The man who (m) I hit saw you’
- (19) *ju lukɔ-re-te aɪn abujɔ arɔ am-te jɔyɔ?*
 ‘The man who(m) I hit saw you?’

Stress is always on the second syllable in this language.

2.6 Morphophonology

Morphophonological alternations seem to happen in case of verb morphology more often than nominal morphology in Juang. Matson (1964: 16–19) has cited a list of environments where morpho-phonemic alternations are noticed. Some of them are given below:

- (20) (a) *lol* → *loi-l* // - {-o} (Past tense)
 {*gito*-} {-o}
gitoi-o ‘he sang’
- (b) *lel* → *li-l* // - {-an}
 {*ɖaki*-} {-an}
ɖakean ‘he called’
- (c) *lsl* → *lcl* // - {Perfect Aspect}
 {*gar*-} {-se} {-re}
garcere ‘he will have begged’
- (d) *lɖl* → *lr-l* // V-
kako-ɖo → *kako-ro* ‘his armpit’

The perfect marker *-se/* exhibits several morpho-phonemic processes. If the tense suffix has an initial vowel, *-t-* is inserted.

- (21) *lon+se +o* → *lonseɔ*
 ‘he was looking’

Adding *-se* to the root causes a devoicing of the final root consonant.

- (22) *jib+se+ke* → *jipseke* ‘he has touched’
jeg+se+ke → *jekseke* ‘he has cried’
kij+se+ke → *kiceke* ‘he has danced’

The plural suffix *-ki/* may change the vowel of the suffix to *i/* (as it is the case for the tense suffixes):

- (23) *jɔi+se+ke+ki* → *jɔi-si-ki-ki* ‘They have seen.’

This assimilation is optional: *domsekiki* ‘they are waiting’

- (24) *dɔm+se+re+ki* → *dɔmseriki*

3 MORPHOLOGY

Munda languages show some unusual patterns for the South Asian *Sprachbund* in terms of their morphological pattern. Nominal morphology is equally as interesting

as verbal morphology. As some of the typical features in these languages adjectives can take tense markers, there are multiple-agreement markers on various constituents in a sentence; conjunctive participles can carry the agreement features of main subject and many more. As this is just a sketch grammar, the following sections describe some of those features briefly.

3.1 Nominal morphology

Nouns in Juang carry person and number agreement markers which occur either as suffixes or prefixes. They carry definiteness markers, classifiers, and case markers. A noun phrase may consist of a single noun, pronoun, adjective or numeral, or a combination of these; determiners precede the head. Adjectives, numerals, and quantifiers precede the noun.

Examples of various combinations are:

- (25) Numeral+adjective+noun
muıntɔ kãʎa lukɔ
 one blind person
 ‘a blind man’
- (26) Determiner phrase+noun
uaʎi-a gata-rɔ
 child-GEN voice-DEF
 ‘A child’s voice’
- (27) (Interrogative) pronoun+noun
maŋɖi kɔnter-ɖe
 which bird-DEF
 ‘Which bird?’

3.1.1 Number

Juang has a three-way number-distinction system, namely singular, dual, and plural. The singular noun is unmarked. Dual number is marked by the suffixes *-bal* and *-kial*. The free morpheme [*banog*] is also used to mark dual in case of honorificity. Plural number is marked by *-kil*.

- (28) *uaʎi* ‘child’ *uaʎi-kia* ‘two children’ *uaʎi-ki* ‘many children’
 (29) *nij-ba* ‘we two’ *apa banog* ‘you two’
 (30) *arɔ-kia* ‘they two’ *arɔ-ki* ‘they all’

Honorificity in number system. Honorificity plays a role in the number system in Juang in the case of ‘one’ and ‘two’. They use *muınto* for non-honorific ‘one’ and *minog* for honorific ‘one’; *ombaɖo* for non-honorific ‘two’ and *banog* for honorific ‘two’.

The numbers can be used either pre-nominally or post-nominally.

- (31) (a) *muınto luko* ‘one person’ vs. *luko muınto* ‘one person’
 (b) *ombaɖo olekia* ‘two mangoes’ vs. *apa banog* ‘two people’

Classifiers with numbers. Number markers are affixes which occur as per the rules of number affixation as described above. Numerals take classifiers like *goʔa* when they occur prenominally.

- (32) *tini goʔa uaʔi-qi-ki*
 three CLSSFR child-DEF-PL
 ‘Three children’

3.1.2 Case

Juang is a postpositional language. Nouns carry lexical case markers as well as postpositions. Case markers can take postpositions too. There can be instances of double case marking as well.

- (33) *Case markers*
 (a) Nominative: Ø
 (b) Accusative/dative case: *-tel/*; sometimes *-bol/* or *-bɔ/*
 (c) Genitive case: *-al/ -rɔ/* or *-ral/ -kal* (+pronominal clitic of the possessors)
 (d) Locative case: *-bɔ/ -tel*
 (e) Ablative case: *-tal/ -tasuml*

Nominative. Nominative case markers in Juang is null. Definite markers or diminutive markers like *-qel/ -qol/ -rel/* and *-rol/* occur with subject nouns or nouns that occur in nominative case positions.

- (34) (a) *ber-ke-kia*
 sleep-PRS -DL
 ‘Both of them are sleeping.’
 (b) *arɔ-kia am-te ʔiɔ buji-kia*
 he-DL you-ACC well do-DL
 ‘Both of them love you.’

Accusative/dative. The accusative marker in Juang is *-tel/*. The dative case marker is *-tel/* too. The accusative case marker is homophonous with the dative case marker in most Indo-Aryan languages. In the case of ditransitive verbs, the direct object gets deleted when it can. When the direct object doesn’t occur independently, the number/person marker of the direct object occurs with the verb. In case both the objects occur in one sentence both can take *-tel/*.

- (35) *arɔ-kia am-te aijn-te ab-mo-jo-oo-m-kia*
 3-DL 2-ACC 1-DAT CAUS-2-see-1-FUT-DL
 ‘Both of them will show me to you’

When the direct object is [-animate], the accusative case marker may not surface. When the case marker does not surface, a classifier or a definite marker normally occurs with the direct object. The genitive/possessive precedes the governing noun. In the possessive/genitive construction, pronominal clitics are copied onto the possessor or the governing noun. The genitive marker is different from the possessive marker in this language.

- (36) *aiŋ-te juang-ka gata-ro-ki ab-soim-ijn*
 I-DAT Juang-GEN language-DEF-PL CAUS-teach
 ‘Teach me the language of Juang’ (Matson 1964)

Genitive. The genitive case marker in Juang is */-a/*. The Oriya loans */-ra/* and */-rɔ/* are often used as genitive case markers. The genitive case markers can be accompanied by other case markers or postpositions. Juang makes a difference between possessor and genitive. In the possessor–possessed relationship the agreement is between the possessor and the possessed.

- (37) *inja+rɔ+a* → *inja-rɔ-a* ‘in his house’
selɔg+dɔ+a → *selɔg-da* ‘of the dog’
iti+pe+a → *iti-p-a* ‘in your hand’
babuŋ+ki+a → *babuŋ-k-a* ‘of the snakes’
dinɔ+rɔ → *dinɔ-rɔ* ‘of the day’

In the possessive construction, pronominal clitics are copied onto the possessor or the governing noun. The genitive case marker may or may not be present in such a construction type. The noun complex with genitive construction can take other case markers. The following examples show the possessor construction that carries accusative case markers.

- (38) *aba-ijn-ate*
 father-1SG-ACC
 ‘to my father’

- (39) *aba-ijn-a-te*
 father-1SG-GEN-DAT
 ‘To my father’

Inanimate nouns do not take pronominal copy of the possessor. Occasional counterexamples are found which is neither systematic nor a pattern. When the pronominal copy is not found, the genitive case marker */-a/* is compulsorily added to the possessor that is preposed to the possessed noun.

- (40) *selog-a bokob-d-o*
 dog-GEN head-DEF-3POSS
 ‘The dog’s head’

Sometimes the suffix */-ka/* occurs in a genitive/possessive construction. It is not clear whether it should be taken as a genuine genitive case marker of Juang or as a deviation such as a Hindi loan.³

- (41) *uaŋi-qi-ki-ka*
 child-DEF-PL-GEN
 ‘Of the children’

- (42) *ar-a-ka tikɔna*
 3SG-GEN-GEN address
 ‘His addresses’

- (43) *konter-a-ka qāṅa-ro-ki asi-ke-ri?*
 bird-GEN-GEN wing-3-1 be-PRS-Q
 ‘Do the birds have wings?’

Since this occurs with a genitive case marker */-a/* most of the times, some speakers prefer to recognize this as an emphatic rather than a case of double genitive. The younger generation prefers to recognize this marker as a genitive case marker.

Genitive case vs. possessor construction. Juangs prefer to use possessor–possessed relationship with kinship terms or nouns indicating body parts. In other cases they use the genitive.

- (44) *inja- a*
 house-GEN
 ‘Of the house’
- (45) *kouṅṅi-ti-m*
 your elbow-hand-2
 ‘Your hand’s elbow’
- (46) *kaka-ij*
 brother-1
 ‘My elder brother’
- (47) *aj*-*a* *aba-ij* *am-te* *di* *buji-ke*
 I-GEN father-1 you-ACC well do-PRS
 ‘My father loves you’

Locative. There are several suffixes that mark direction or location of space and time in Juang. The most often used locative case marker is */-bɔ/*. Other suffixes which are used to mark location are: */-al/*, */-yal/*, */-ɖal/*, */-ral/*, */-el/* and */-rel/*. The following examples are illustrative.

- (48) *nin-a nɔdi-bɔ n-ɔn-ɖe.*
 we.PL-GEN river-LOC 1PL-go-PRS
 ‘We are going to the river.’
- (49) *aj*-*a* *betɔj* *bode-rag* *nij-a* *gāō muji-bo* *dulmuli-na*
 I-GEN fear flood-water 1PL-GEN village mid-LOC enter/3FUT/-FUT
 ‘I fear the flood water will get into our village.’
- (50) *arɔ-kia banɔg-bɔ di* *a-ku-buji-ri-kia.*
 3DL two-LOC good NEG-RECIP-like-PRS-DL
 ‘They don’t like each other.’
- (51) *aj* *ar-a-bɔ* *loj-se-ke.*
 I 3SG-OBLQ-LOC see-PRF-PRS
 ‘I’m looking at him.’
- (52) *ere lukɔ-re kutia-bɔ susumusij-ɖe-te* *dej-ke.*
 that person-DEF axe-INS REDPL:tree-DEF-ACC cut-PRS
 ‘That man is cutting the tree with an axe.’

- (53) *sumusiŋ-a*
tree-LOC
'On the tree'
- (54) *inja-ya*
house-LOC
'In the house'
- (55) *sumusiŋ-ɖa*
tree-LOC
'On the tree'
- (56) *ɔŋʃi-ra*
throat-LOC
'Inside the throat'
- (57) *aʃike-re*
pot-LOC
'In the pot'

Ablative. The ablative case markers are */-tasun/* and */-ta/*. The suffixes may be added either directly to the noun or to the genitive marker */-a/*.

- (58) *bojarɔ-ta*
market-ABL
'From the market'
- (59) *ɔlej-a-ta*
COW-GEN-ABL
'Than the cow'

The suffix *-ta* (*-sun*) indicates the starting point.

- (60) *subudinɔ dikɔɔ-te nin skulu-ta ni-duŋ-ke.*
every day afternoon-LOC 1PL school-ABL 1PL-return-PRS
'Every afternoon we come back from school.'
- (61) *ajŋ sɔhɔɔ-tasun laŋka rɔhe-jena.*
I town-ABL far live-NEG
'I don't live far from town.'
- (62) *ajŋ kebe:le ar-aka-ta ruŋɔ ki-kib-jena.*
I ever 3-GEN-ABL loan REDPL-make-NEG
'I never borrow money from him.'
- (63) *am-a aba-m sekaɔkia-ta kiti sɔmɔɔ jirai-re?*
you-GEN father-2 lunch-ABL how much time rest-PRS
'For how long does your father rest after lunch?'
- (64) *nin ariga-ta sugei tɔnɔ n-ur-ke.*
1PL curry-ABL very rice 1PL-eat-PRS
'We eat more rice than curry.'
- (65) *nin-a ɔlej ape-a ɔlej-ata ɔʃɔɔ dudɔ ɖin-ke.*
1PL-GEN cow 2PL-GEN COW-ABL little milk give-PRS
'Our cow gives less milk than yours.'

- (66) *suruma jiarɔ- alij-ta dɔkɔ-se-ran.*
 Suruma chair-on-ABL sit down-PRF-PST
 ‘Suruma was sitting on a chair.’
- (67) *rasɔta sag-iti bakɔʔ-ta dajij -pe.*
 street left-hand side-ABL keep-2PL
 ‘Keep to the left.’

Comparison. The ablative case marker /*ta*/ is used to mark comparison too.

- (68) *ajɪ-a inja-n-ɖe am-a inja-m-ta kuba.*
 I-GEN house-1-DEF you-GEN house-2-ABL big
 ‘My house is bigger than your house.’

Dative subject. Like many South Asian languages, the subject NP of predicate expressing emotions, attitude and volition carry dative case marker.

- (69) *am-te jɔɔ ian-ri?*
 you-DAT fever happen-Q
 ‘Do you have fever?’

Like other South Asian languages, even in Juang when the subject argument is marked for dative the object argument is marked for nominative. When there is a dative subject along with a nominative object, the verb agrees with the object because the dative case marker blocks subject verb agreement. When the dative subject occurs in an embedded clause the clause is often in an infinitive or in a gerund form. In this case the dative case, which is also called an experiencer subject, exhibits impersonal agreement with the matrix verb.

Genitive subject. In the genitive subject construction the subject carries genitive case marker and the object carries nominative case. Even here the verb agrees with the object and not with the subject.

- (70) *ramɔ-a [ajɪ-a inja-te ɔngɔr-te] beʔɔ a-sian*
 ram-GEN I-GEN house-LOC go-INF time NEG-be-PST
 ‘Ram doesn’t have time to go to my house.’

3.1.3 Person

There are three persons and three numbers in Juang, normally, first person, second person and third person. Every person has a singular form, a dual form and a plural form. Person markers do not necessarily precede roots unless it is non-third.

The pronouns may have case suffixes. The possessive pronouns are formed by the addition of the case /-*al*/.

- (71) *ajɪ-a uɲi-n-ɖe kɔsɔkɔ-ra, ele ape-a sɔpa-rɔ*
 I-GEN shirt-1-DEF dirty-DEF but 2PL-GEN clean-DEF
 ‘My shirt is dirty, but yours is clean.’

Possessor agreement. Juang, like most Munda languages, exhibit possessor agreement in which the possessed agrees with the possessor. The following table illustrates the possessor agreement markers.

(72)	Pronouns	Person marker	Agreement marker
	1SG	<i>ajn</i>	<i>-ijn</i>
	1DL	<i>nijn-ba</i>	<i>-nijnba</i>
	1PL	<i>nijn</i>	<i>-nejnin</i>
	2SG	<i>am</i>	<i>-nom</i>
	2DL	<i>apa</i>	<i>-pa</i>
	2PL	<i>ape</i>	<i>-pe</i>
	3SG	<i>arɔ</i>	<i>-rɔ</i>
	3DL	<i>arɔ-kia</i>	<i>-kia</i>
	3PL	<i>arɔ-ki</i>	<i>-ki</i>

- (73) *oba-ijn* 'my father'
oba-rɔ 'his father'
oba-m 'your father'

Possessor agreement is compulsory in case of kinship terms and inalienable possession. It's not compulsory otherwise as the following examples illustrate.

- (74) *ajn-a injaj-ijn* vs. *ajn-a inja* 'My house'
arɔ-kia selog-do vs. *arɔ-kia selog* 'Their dog'

3.1.4 Definiteness

Definite markers behave in a very interesting way in Juang. The definiteness markers in Juang are: *-de/ -re/* and *-rol/*. The definite markers occur with nominative subject and genitive cases quite often. They are seen to occur with verbs too in non-finite clause constructions. They are occasionally seen with finite verbs too. The following sections describe the occurrence and importance of definite markers as and when they occur.

We provide below a few examples in which the definite markers occur most often.

- (75) *ualji-de* 'The child'
aba-n-de 'My father'
ijij-n-re 'My foot'

Emphatic. The suffix *-ka/* which is seen as a genitive occurs in some other cases too where there are no genitives. In this case they are used to add some type of emphasis of definiteness.

- (76) *arɔ-ki curɔ-ka dagan jɔ-re-ki.*
 3-PL thief-EMPH like look-PRS-PL
 'They look like thieves.'

3.1.5 Gender/nominal class

Gender in Juang is marked by several affixes as well as full words. Full words are borrowed from Oriya whereas the affixes are genuine to Juang. Bhattacharya (1975) makes a full list of gender-forming affixes in many Munda languages. I provide below gender-forming affixes in Juang some of which are given in Bhattacharya and some are my own.

- (77) Class I gender-forming suffixes
/-ser/
/-mail
- (78) Class II gender-forming suffixes indicating feminine
/-i /
/-ai /
/-ḍey/
/-rael
/-ray/
/-jey/
/-jei /
/je /
/-rma/
/-nday/
/-nṛay/
- (79) Class III gender-forming suffixes denoting masculine
/-nger/
/-ger/
/-er/
/-al
/-oal
/-ail
/-ul
/-ull

I provide few examples below where addition of these markers changes the gender of the noun.

- (80) *juan* ‘Juang man’ *juan-dai* ‘Juang woman’
bəkə ‘brother’ *bəkə-ai* ‘sister’
buḍa ‘old man’ *buḍi* ‘old woman’
andjira selɔg ‘male dog’ *mai selɔg* ‘female dog’

3.1.6 Reflexives and reciprocal pronouns

Juang has separate forms for reflexives, reciprocals, and personal pronouns. Reflexives and reciprocals always find their antecedents either in the smaller domain or in the larger domain inside the full sentence. Pronouns are often used in place of reflexives and they too find their antecedents inside the full sentences as well as outside thus creating ambiguity of reference. Patnaik (2000) discusses the details of the anaphoric structure with binding possibilities. I provide a very short description of the categories in this chapter.

Reflexives. Juang has a nominal reflexive. The element */-dom/* which is referred to as a verbal reflexive in other Munda languages is present in Juang, but in Juang it behaves like a passive marker only and not as a reflexive. The simple forms of the reflexives in Juang are *aapein* and *aapaṇaa*. The reflexives can be followed by all

lexical case markers like full nouns. Juang also has composite as well as reduplicated forms of the reflexive.

There can be two ways of reduplicating the reflexives. The structure of the reduplicated reflexive can be either of the two examples provided in Box 10.1.

BOX 10.1: REFLEXIVES IN JUANG I

Reflexive+reflexive+case marker
Or
Reflexive+case marker+reflexive

The structure of the composite form of the reflexive can be either of the two examples provided in Box 10.2:

BOX 10.2: REFLEXIVES IN JUANG II

Pronoun+reflexive+case marker
Or
Pronoun+case marker+reflexive

The lexical case markers can occur either to the right of the entire reduplicated form as well as the composite form or to the right of the first part of the complex.

- (81) *ramo apein-te dʒɔ buji-ke*
Ram self-ACC well like-PRS
'Ram loves himself.'

Reciprocals. There are two types of reciprocals in Juang, namely nominal reciprocal and verbal reciprocal. The nominal reciprocal in Juang is either a reduplicated form of a reflexive or a pronoun. There is also a composite form of reciprocal which consists of a reflexive and a pronoun.

The four forms of the reduplicated reciprocals in Juang are as given in Box 10.3.

BOX 10.3: RECIPROCAL IN JUANG I

- (a) *apein*+case marker+*apein*
(b) *apein*+*apein*+case marker
(c) *araka*+case marker+*araka*
(d) *araka*+*araka*+case marker

The composite form also has four forms provided in Box 10.4.

BOX 10.4: RECIPROCAL IN JUANG II

- (a) *araka*+case marker+*apein*
(b) *araka*+*apein*+case marker
(c) *apein*+case marker+*araka*
(d) *apein*+*araka*+case marker

Note that the form *araka* (*arɔ+ka*) is used in all these cases and not *larɔ*. The suffix *-kal* here may have emphatic definitive connotation if not genitive.

In addition to these forms, there is *paraspara* a borrowed form from Oriya which is used as a reciprocal in Juang. This form can also be reduplicated. In case the lexical case marker occurs to the right of each of the constituents then the form imparts only the distributive meaning.

Apart from the nominal reciprocal, Juang has a verbal reciprocal too which is *lkul*. This verbal reciprocal is always prefixed to the verb.

- (82) *araki araka araka diɔ ku-buji-ri-ki*
 3PL each other well RECIP-DO-DEF-PL
 ‘They love each other.’

The verbal reciprocal may occur with the nominal reciprocal or it may occur in cases where there is no nominal reciprocal. The *-kul* reciprocal is discussed in sections on derivation (section 3.1.10) too.

Juang has three persons and three numbers. All of them can take lexical case markers like full nouns.

- (83) *Personal pronouns*
- | | | | | | |
|------|--------------|------|------------|------|---------------|
| 1 SG | <i>aɪŋ</i> | 2 SG | <i>am</i> | 3 SG | <i>arɔ</i> |
| 1 DL | <i>ninba</i> | 2 DL | <i>apa</i> | 3 DL | <i>arɔ-ki</i> |
| 1 PL | <i>nɪŋ</i> | 2 PL | <i>ape</i> | 3 PL | <i>arɔ-ki</i> |

The pronouns may have case suffixes. The possessive pronouns are formed by the addition of the case suffix *-al*.

Juang has several interrogative pronouns that can also take lexical case markers and occur in various case-marked positions. Some of the interrogative forms are as follows:

- (84) *adi* ‘who’ *biri* ‘what’
adi-te ‘whom’ *adi-a* ‘whose’ *adi-aliŋa* ‘where’

3.1.7 Demonstratives

There are several demonstrative pronouns in Juang. The demonstrative pronouns can also take all lexical case markers like pronouns and anaphors. Postpositions do not occur with demonstratives.

Some of the forms of demonstrative are given below:

- (85) *ni* ‘this’ *ni-ta* ‘from here’
ini ‘this’
auri ‘that’ *au-ra* ‘there’ *au-bɔ* ‘there’
ere ‘that’ *era* ‘there’ *eta-bɔ* ‘here’
uil-a ‘that’
nan-te ‘here’ *enaŋ* ‘here’

eta which means ‘thus’ otherwise has the meaning ‘here’ when it occurs with a locative case marker *lbɔ*, whereas *au* which is otherwise a coordinator has the meaning ‘there’ when it occurs with the same locative case marker *lbɔ*. *au* has the distal meaning like ‘that’ too.

3.1.8 Numerals and classifiers

The numeral system in Juang is pretty mixed. The first four numbers and few other numbers belong to the genuine Juang system whereas the rest are borrowed from Oriya or rather they are of the Indo-Aryan pattern.

After checking it with a large number of speakers and texts I found out that only the numbers till four are of genuine Juang pattern and numbers higher than four are borrowed from Oriya.

- (86) *Number list*
muintɔ ‘one’
ombatɔ ‘two’ *banɔg* ‘two’
egɔta ‘three’ *ganda-min* ‘four’

The original Juang numbers don’t take classifiers. The borrowed forms take a classifier such as *gɔta*.

- (87) *pancɔ-gɔta* ‘five’ *cha-gɔta* ‘six’

The classifier *gɔta* is borrowed from Oriya. *gɔta* is also used after the interrogative and indefinite pronoun *kiti* also.

- (88) *buli-ra kiti-gɔta ɔlei asi-ki-ki.*
 Buli-GEN how class cow be-PRS-PL
 ‘How many cows does Buli have?’

muinto is the only number that *occurs* pre-nominally as well as post-nominally whereas all other numbers have post-nominal occurrence.

- (89) *inja agila-ta juayɔɔai lukɔ-re muintɔ asi-ke.*
 house in front of-ABL woman person-DEF one be-PRS
 ‘There is a woman in front of the house.’

Other important numbers in Juang are as follows:

- (90) *kuɔi-mui* ‘twenty’ *kurie-ganda-mi* ‘twenty-four’
sumui ‘hundred’ *ɔjɔɔ-mui* ‘thousand’
muikua ‘a handful’ *muɔami* ‘a handful’

3.1.9 Postpositions

Nouns in Juang carry postpositions. Most of the postpositions are borrowed from the superstratum Oriya though some are genuine to Juang. Most of the postpositions can follow genitive case markers and oblique case markers. The structure of a postpositional noun is provided in Box 10.5:

BOX 10.5: JUANG NOUN STRUCTURE

Noun+genitive case marker / oblique case forms / definite marker+postposition

Some of the commonly found postpositions are given in the following examples:

- (91) *agila* ‘in front of, before’ [Oriya- *agree*] *aguru* ‘before’ [< Oriya]
aliɔ ‘on’ *aluɔ* ‘in’
belɔ ‘in’ (temporal) [Oriya: *belo*] *lagi* ‘for’ [< Oriya]

<i>buluj</i> ‘till’	<i>ian</i> ‘for’ (temporal)
<i>dire</i> ‘by’ [Oriya: <i>dwarā</i>]	<i>kɔra</i> ‘during’
<i>jakɔi</i> ‘until’ [Oriya: <i>jake</i>]	<i>samu</i> ‘near’
<i>ɔɔta</i> ‘behind’ [Oriya: <i>pɔchɔta</i>]	<i>uculita</i> ‘behind’
<i>tulia</i> ‘under, ago’ [Oriya: <i>tɔlia</i>]	

I provide a few sample examples below with postpositions.

- (92) *ajɲ ua|ɲi-n-a belɔ-a ajɲ kiriketɔ ebɔ-nɔm-an.*
 I child-1-GEN time-GEN I cricket play-PRF-PST
 ‘I used to play cricket in my childhood.’
- (93) *arɔ ajɲ-a lagikici ki-kib-jena.*
 3SG I-GEN for anything REDPL-DO-NEG
 ‘He doesn’t do anything for me.’
- (94) *ajɲ minɔ atɔ-bɔ ɔn-ɔ, ajɲ-a tuloyg-ke ɔn-gɔr-jena.*
 I alone market-LOC go-PST I-GEN anyone-ACC go-PFV-NEG
 ‘I went to the market alone, no one went with me.’

The postpositions may be combined with case suffixes.

- (95) *inja agila- ta juayɲai lukɔ-re muintɔ asi-ke.*
 house in front of-ABL woman person-DEF one be-PRS
 ‘There is a woman in front of the house.’
- (96) *gatɔ-r-aluj-a babuyg asi-ke.*
 hole-DEF-in-GEN snake be-PRS
 ‘There is a snake in the hole.’

3.1.10 Derivation

Several derivational processes are present in Juang. In fact, complex derivation is one of the defining features of Munda linguistics that establishes itself as a separate type from other languages of the South Asian *Sprachbund*.

Some of the derivational processes in Juang are realized through the following:

- (a) Reduplication
- (b) Prefixation
- (c) Suffixation
- (d) Ambifixation

Reduplication is used in the formation of infinitives and participles. Suffixation is used for adjectivalizing nouns and verbs and vice versa.

For example infinitives and gerunds are formed by reduplicating the root.

- (97) *ur* ‘drink’ > *urur* ‘in order to drink’
jɔ ‘see’ > *jɔjɔ* ‘seeing’

Negative verbal noun can be formed by prefixing /a-/ and suffixing /-el/.

- (98) *gata* ‘say’ vs. *a-gata- e* ‘untold story’
ɔmɔ ‘saw’ vs. *a-ɔn-jɔmɔ* ‘unseen things’

3.1.11 *Adjectives*

Adjectives in Juang can occur both in pre-nominal attributive position as well as in predicative position. The most fascinating aspect of Juang adjectives are that they can take definiteness markers, case markers as well as tense markers. The following examples are illustrative.

- (99) *ɖiɔ uali*
 Good child
 ‘Good boy’
- (100) *lalaj lɔɔb-de*
 fire hot-DEF
 ‘Hot fire’/ ‘Fire is hot’

Adjectives can take tense markers in which they behave like verbs.

- (101) *enteij* ‘little’
enteij-ke ‘is becoming little’

Some of the commonly used adjectives are as follows:

- (102) *guleij-ke*
 round-PRS
 ‘It’s become round.’
- (103) *ɖiɔ* ‘good’ *jute* ‘few’ *enteij* ‘little’
sugei ‘many’ *anɖarɔ* ‘dark’ *buɕua* ‘strong’
kuɕig ‘some’ *kuba* ‘big’ *sugei* ‘enough’
lɔɔp ‘hot’ *guleij* ‘round’

Comparison. The ablative case marker /*tal*/ is used to mark comparison too.

- (104) *aij-a inja-ŋ-ɖe am-a inja-m-ta kuba*
 I-GEN house-1-DEF you-GEN house-2-ABL big
 ‘My house is bigger that your house.’

3.1.12 *Adverb(ial)s*

Three types of adverbs are found in Juang, namely simple adverbs, complex adverbs, and compound adverbs. Adverbs denote spatial relationships, temporal relationships, manner, and others. Simple adverbs are monomorphemic in nature.

- (105) *Temporal*
ekɔɔŋ ‘now, just’ *tome* ‘now’
aciŋka ‘day before yesterday’ *aŋɖeka* ‘yesterday’
miŋiŋ ‘today’ *tera* ‘tomorrow’
- (106) *Spatial*
tulina ‘down’
aluna ‘inside’
ekeca ‘near’
- (107) *Manner*
cɔŋcɔɔ ‘quickly’

Adverbs can carry case markers in Juang.

- (108) *ai-ta* ‘that side’
alum-ta ‘from inside’
meser-te ‘in the day’

Adverbs get reduplicated to intensify the action of the verb they are modifying. This happens very often in many Indo-Aryan languages, including Oriya too, which may have been borrowed into Juang.

- (109) *ekura ekura* ‘often’

3.2 Verb morphology

Verbs in Juang carry tense, aspect, mood as well as person and number agreement markers. These markers are suffixed, prefixed, infixes and even at times ambifixes. There are subject and object agreement markers on the verb though object agreement is being lost due to superstratum influence. As it is, object agreement in Juang is not as obligatory or as productive as in other Munda languages. In fact, object agreement and multiple object and noun agreement is one of the famous features of Munda languages that doesn't seem to be noticed in Juang as productively though one can force the speakers to construct them without threatening the grammaticality of the construction types.

Juang has finite as well as non-finite verbs. Finite verbs carry subject person by prefixes, object person by suffixes; number markers by prefixes and suffixes as well; tense markers by suffixes as well as by an infixes as in future tense; aspect markers by suffixes; various polarity and negation markers by prefixes, suffixes, infixes, and ambifixes along with reduplication. Non-finite verbs lack tense and normally agreement. Conjunctive participles carry agreement markers. Lack of finiteness is expressed by specific markers.

Verb morphology is the only aspect on which much work has been done on Juang. According to the Juang–Kharia comparative work done by Mahapatra (1976), the Juang chart alone can generate thirty basic verb paradigms with four basic morphological classes.

For the basic verb structure see Box 10.6.

BOX 10.6: JUANG BASIC VERB STRUCTURE

Root+Aspect+Tense+Person/pronominal clitic
--

Person markers precede the root verb in the case of non-third person. Roots in Juang (Konow 1906 (=Grierson 1967), Mahapatra 1976, Matson 1964, Pinnow 1960) have been subdivided into three subclasses – transitive, intransitives, and transitive-intransitive. Transitive verbs take one specific set of tense markers; intransitive verbs take another specific set of tense markers; and transitive–intransitive verbs can take tense markers from both the sets. It has been observed by Mahapatra that in Juang, a number of roots are clearly exempt from the transitive–intransitive opposition. The function of the root can be determined only from its co-occurrence with the particular set of tense markers. For example, verbs such as *pag* meaning ‘to break/to be broken’ or *guj* meaning ‘to wash’ or ‘to be washed’ can belong to both

set I and set II. I will provide the list of set I and set II verbs in the tense sections (section 3.2.3). Matson's PhD thesis and Mahapatra's paper on verb morphology provide lists of those verbs quite elaborately though I have collected my own verb list independent of their works.

Verb reduplication is a very productive morphological as well as syntactic process in Juang. Monosyllabic roots are more often reduplicated than multisyllabic roots. Some of the examples of verb reduplication as given in Mahapatra (1976) are as follows:

(110)	Root	Reduplicated base	Glosses
	<i>i-</i>	<i>ii-</i>	be
	<i>ig-</i>	<i>iig-</i>	open
	<i>ur-</i>	<i>urur-</i>	eat
	<i>jɔ-</i>	<i>jɔjɔ-</i>	see
	<i>ku-</i>	<i>kukui-</i>	get
	<i>sɔb-</i>	<i>sɔbsɔb-</i>	hold
	<i>dijnɔ-</i>	<i>dijnɔdijnɔ-</i>	give
	<i>jim-</i>	<i>jjjim-</i>	eat

3.2.1 Subject

Juang being a pro-drop language, the subject noun often doesn't occur in an independent position. Instead, the subject agreement marker occurs in the verb complex. The subject agreement marker can either be prefixed or suffixed. In case of the third singular it can be infix.

There are three persons and three numbers in Juang. The three persons (1, 2, and 3) and three numbers (singular, dual, and plural) are expressed through nine different agreement markers on verbs. The agreement marker for 1SG and the agreement marker for 3SG are identical which means they are null. I provide below a small table for the person marking here; towards the end of this section, I will also provide a full table that would show all person-number agreement markers in negative construction, possessive construction, and various tense and aspect constructions.

(111) Person and number affixes

	SG	DL	PL
(a)	Ø	<i>ba-</i>	<i>n (V_i)-</i>
(b)	<i>m (V_i)-</i>	<i>a-</i>	<i>V_i-</i>
(c)	Ø	<i>-kia</i>	<i>-ki</i>

(112)	<i>mV_i+ur-ke</i> → <i>m-ur-ke</i>	'you are drinking'
	<i>nV_i+ur-ke</i> → <i>n-ur-ke</i>	'we all are drinking'
	<i>mV_i+gɔg-dɛ</i> → <i>me-gɔg-dɛ</i>	'you are taking'
	<i>nV_i+gɔg-dɛ</i> → <i>ne-gɔg-dɛ</i>	'we all are taking'
	<i>V_i+gɔg-dɛ</i> → <i>e-gɔg-dɛ</i>	'you all are taking'
	<i>mV_i+jelai-ke</i> → <i>me-jelai-ke</i>	'you are fishing'
	<i>nV_i+jelai-ke</i> → <i>ne-jelai-ke</i>	'we all are fishing'
	<i>V_i+jelai-ke</i> → <i>e-jelai-ke</i>	'you all are fishing'
	<i>mV_i+buli-ke</i> → <i>mi-buli-ke</i>	'you are walking'
	<i>nV_i+buli-ke</i> → <i>ni-buli-ke</i>	'we all are walking'
	<i>V_i+jim-e</i> → <i>me-jim-e</i>	'you all are eating'
	<i>mV_i+dijnɔ-ke</i> → <i>mi-dijnɔ-ke</i>	'you all are giving'
	<i>nV_i+dijn-ke</i> → <i>ni-dijn-ke</i>	'we all are giving'

Dative subject. Subjects in Juang can carry a dative case marker. Then the agreement between the subject and verb gets blocked. I found a few instances in which the agreement was not blocked. It was not blocked probably because in the speaker's mind the dative subject was processed as an accusative NP and he went ahead with the agreement the way he would do it in the case of object agreement. I provide those instances below.

- (113) *ajŋ -te bŋɔ-bɔ ɔn-gɔr-te betɔŋ-ki-ŋ*
 I-DAT forest-LOC go-PFV-PRF fear-PRS-1
 'I fear going alone to the forest.'
- am-te bŋɔ-bɔ ɔn-gɔr-te betɔŋ-ko-m*
 you-DAT forest-LOC go-PFV-PRF fear-PRS-2
 'You fear going alone to the forest.'

3.2.2 Object types

In Juang object agreement is not as productive as in some other Munda languages but it is present. There seems to be a difference between the time when Matson had worked on object agreement in the 1960s and when I worked on the same 40 years after. Matson's work gives the impression of Juang having had very productive object agreement and I don't deny it. Matson's Class-A verbs show object concord very productively. The affix classes of object agreement are as follows:

	1	2
SG	-ŋ	-m
DL	-ŋba	-pa
PL	-ŋenijŋ	-pe

I provide below modified glossing versions of some of the examples given in Matson (1964: 61).⁴

- (115) *am ajŋ-te me-jɔ-ki-ŋ*
 you I-ACC 2-see-PST-1
 'You saw me.'
- (116) *kamuluj ŋijba-te gatai-o-ŋba*
 man we.DL-ACC talk-PST-1DL
 'A man talked to both of us.'
- (117) *selɔg ŋij-te kedao-ɔ-ŋenijŋ*
 dog we-ACC bite-PST-1PL
 'A dog bit us all.'
- (118) *ajŋ am-te jɔ-ŋ-kɔ-m*
 I you-ACC see-1-PRS-2SG
 'I see you.'
- (119) *arɔ apa-te jo-kɔ-pa*
 he you-ACC see-PRS-2DL
 'He sees both of you.'
- (120) *arɔ ape-te jɔ-ke-pe*
 he you.PL-ACC see-PRS-2PL
 'He sees you all.'

- (121) *ajɲ ara-te jɔ-o-ke*
 I he-ACC see-1-PRS
 'I see him.'
- (122) *ajɲ inja-te jɔ-o-ke*
 I house-ACC see-1-PRS
 'I see a house.'

During my field trip I realized that the occurrence of the object suffixes have become largely optional. In all the cases cited below, speakers have expressed optionality in the occurrence of object agreement markers. Some speakers have expressed a strong opinion about the obligatoriness of the occurrence of the object agreement marker though.

- (123) *biri ki-kib-te m-i-na ajɲ-te abjɔ-i-nɲɲ.*
 what REDPL-DO-INF FUT-be-FUT I-ACC show-IMP-1
 'Show me what to do.'
- (124) *am-te aɟi-ko mino dake-ke*
 YOU-ACC who-DEF one call-PRS
 'Someone is calling you.'
- (125) *ajɲ-te dag gilaso min oren-ɲɲ dai*
 I-ACC water glass one bring-1-FEM
 'Please get me a glass of water.'

There is some example in which the speakers resisted to object agreement too. The examples are as follows:

- (126) *am ajɲ-te Juang gata m-ab-soj-e.*
 you I-ACC Juang language 2-CAUS-learn-FUT
 'You will teach me Juang.'

The following table illustrates all possibilities of person–number agreement markers in various cases.

(127)	Pronouns	Subject agreement marker	Object agreement marker	Possessive agreement marker
1 SG	<i>ajɲ</i>	<i>V^c-</i>	<i>-ɲɲ/-ɲ</i>	<i>-ɲɲ/-ɲ</i>
1 DL	<i>nɲɲ-ba</i>	<i>bV^m- (ba-)</i>	<i>-ɲba</i>	<i>-nɲɲ</i>
1 PL	<i>nɲɲ</i>	<i>n V^c- / nV-</i>	<i>-nenɲɲ</i>	<i>-enɲɲ</i>
2 SG	<i>am</i>	<i>m V^c- / mV</i>	<i>-nɔml / -ɔml-m</i>	<i>-nɔml / -ɔm</i>
2 DL	<i>apa</i>	<i>V-(a-)</i>	<i>-pa</i>	<i>-pa</i>
	<i>apa banog</i>			
2 PL	<i>ape</i>	<i>V-</i>	<i>-pe</i>	<i>-pe</i>
3		<i>mV^c</i>		
3		<i>/-mV^c -/</i>		
3		<i>-mV^c</i>		
3 SG	<i>arɔ</i>	—		<i>-dɔ</i>
3 DL	<i>arɔ-kia</i>	<i>-kia</i>		<i>-dɔkia</i>
3 PL	<i>arɔ-ki</i>	<i>-ki</i>		<i>-dɔki</i>

The pronominal affixes which are actually the contracted versions of the pronominal can occur as subject agreement markers and object agreement markers in various tenses, negatives, and possessives.

3.2.3 Tense

The Juang tense paradigm is relatively complex. Transitive verbs take one set of tense markers and intransitive verbs take another set. Set I exemplifies tense suffixes for transitive verbs whereas set II exemplifies tense markers for intransitive verbs. Most of the tense markers occur as suffixes. Only the future tense marker occurs as an infix. There are some verbs which take tense markers from both sets.

(128) Tense markers

	Set I	Set II	
Simple present	-ke	-qe	
Simple past	-o	-an	
Simple future	-e	-na	(+ Infix -mV-, see below)

List of verbs of in set II which are mostly intransitive in nature are as follows:

(129) Set-II verbs

Stem	Present	Future	Past	Perfect	
<i>ari</i>	<i>ariqe</i>	<i>arina</i>	<i>arian</i>	<i>arisere</i>	'come out'
<i>budji</i>	<i>budjire</i>	<i>budjina</i>	<i>budjian</i>	<i>budjisena</i>	'set (sun)'
<i>bułur</i>	<i>bułurqe</i>	<i>bułuna</i>	<i>bułuran</i>	<i>bułurceran</i>	'get up'
<i>dokɔ</i>	<i>dokɔre</i>	<i>dokɔna</i>	<i>dokɔan</i>	<i>dokɔsere</i>	'sit'
<i>dui[r]</i>	<i>duide</i>	<i>duina</i>	<i>duiran</i>	<i>duicere</i>	'enter'
<i>gɔg</i>	<i>gɔgde</i>	<i>gɔna</i>	<i>gɔgan</i>	<i>gɔksere</i>	'take'
<i>i</i>	<i>iqe</i>	<i>ina</i>	<i>ian</i>	<i>iseke</i>	'be'
<i>jirai</i>	<i>jiraiqe</i>	<i>jiraina</i>	<i>jiraian</i>	<i>jiraiserɔ</i>	'rest'
<i>ku[r]</i>	<i>kułqe</i>	<i>kuna</i>	<i>kuran</i>	<i>kucera</i>	'put on'
<i>laŋ</i>	<i>laŋqe</i>	<i>laŋna</i>	<i>laŋan</i>	<i>laŋsere</i>	'belch'
<i>nɔn</i>	<i>nɔnqe</i>	<i>nɔna</i>	<i>nɔnan</i>	<i>nɔncerɔ</i>	'get caught'
<i>tɔŋɔn</i>	<i>tɔŋɔnqe</i>	<i>tɔŋɔna</i>	<i>tɔŋɔnan</i>	<i>tɔŋɔncere</i>	'stand'
<i>uđi</i>	<i>uđire</i>	<i>uđina</i>	<i>uđian</i>	<i>uđiseran</i>	'fly'

(130) Set-I verbs

<i>kui</i>	<i>kuike</i>	<i>kuina</i>	<i>kuian</i>	<i>kuisere</i>	'receive'
<i>sub</i>	<i>subke</i>	<i>suna</i>		<i>suseran</i>	'wear'

(131) List of set-I verbs (Matson [1964:23])

<i>absoŋ</i> 'sell'	<i>aitog</i> 'scratch'	<i>bande</i> 'cheat'
<i>bađ</i> 'beat'	<i>botae</i> 'command'	<i>dolae</i> 'shake'
<i>đake</i> 'call'	<i>đijn</i> 'give'	<i>gaj</i> 'fry'
<i>gata</i> 'speak'	<i>gɔla</i> 'cook'	<i>gɔŋe</i> 'count'
<i>gɔse</i> 'rub'	<i>jan</i> 'lick'	<i>jim</i> 'eat'
<i>jij</i> 'ask'	<i>jɔ</i> 'see'	<i>jułi</i> 'copulate'
<i>keđab</i> 'bite'	<i>kep</i> 'cut'	<i>koŋ</i> 'know'
<i>lej</i> 'scold'	<i>lobtor</i> 'burn'	<i>loŋ</i> 'look at'
<i>maje</i> 'rinse'	<i>mane</i> 'obey'	<i>ojae</i> 'lose'
<i>roj</i> 'milk'	<i>sade</i> 'leave'	<i>sesej</i> 'give haircut'
<i>sɔb</i> 'hold'	<i>tałam</i> 'caress'	<i>taŋ</i> 'measure'

<i>taŋe</i> ‘pull’	<i>teke</i> ‘lift’	<i>tele</i> ‘push’
<i>tice</i> ‘carry on head’		<i>tin</i> ‘bury’
<i>tintar</i> ‘carry on shoulder’		<i>tom</i> ‘throw’
<i>uci</i> ‘wipe’		<i>un</i> ‘keep’

Notice in the following examples the behaviours of the verb *lɔen/* that takes tense marker of class-I verb and present tense marker and future tense marker of class-II verbs.

(132) *ɔen ɔeŋɔ ɔena ɔenɔ* ‘come’

There are some more verbs that carry suffixes of either class. The following examples are illustrative.

(133) <i>kib</i>	<i>kibke</i>	<i>kibe</i>	<i>kibɔ</i>	<i>kipseke</i>	‘do’
		<i>kibɔe</i>	<i>kima</i>		
<i>lage</i>	<i>lageke</i>		<i>lagaɔ</i>	<i>lageseke</i>	‘feel’
	<i>lagare</i>			<i>lagian</i>	
<i>sɔb</i>	<i>sɔbke</i>	<i>sɔbe</i>	<i>sɔbɔ</i>	<i>sɔbseke</i>	‘hold’
			<i>sɔban</i>		

In the progressive, reduplication occurs only with monosyllabic stems and only some final consonants get deleted. Nasals do not get deleted but get assimilated.

3.2.4 Aspect

There are two basic aspectual categories in Juang: perfect and progressive. Usually the verb gets reduplicated to denote an imperfect aspectual dimension.

The perfect aspect markers in Juang are *l-sel*, *lce-l*, and *l-tel*. Morphophonemic variants of *l-sel* are *l-cerl* in the environment of *jlnln-V*; *l-cel* in the environment of *jlnln_C*; *l-serl* before V if not following *lj/*; and *l-sel* elsewhere. The continuous aspect markers are *l-nɔml* and *l-ɟl*.

There are two groups of aspect markers. The first group consists of two suffixes that are added directly to the root. The second group consists of several suffixes that are added to the right edge of the verb form.

l-sel most often occur as a perfect marker. *l-sel* has other functions too.

(134) *ajŋ akurɔsɔ-rɔ* *cake-se-ke*
 I sugarcane juice-DEF taste-PRF-PRS
 ‘I have tasted sugarcane juice.’

l-sel doesn’t occur when the perfective aspect is negated. In that case the negative marker *ljenal* occurs and still the negative perfect meaning is understood.

(135) *ajŋ kebe akurɔsɔ-rɔ* *cake-jena*
 I when sugarcane taste-NEG
 ‘I have never tasted sugarcane juice.’

l-sel has a variant *lcel* as the following example illustrates.

(136) *ajŋ-a bui-ŋ* *ar-a-ka* *nokotej* *suna* *beŋke-a* *un-ce-te*
 I-GEN mom-1 3-GEN-GEN little gold bank-LOC! keep-PRF-PST
 ‘My mother has kept most of her gold in the bank.’

There are instances where *l-sel* and *lcel* occur to describe the state of being of an event.

- (137) *rasɔta kɔdɔ-a muintɔ kãla lukɔ dɔkɔ-se-re.*
 road side-LOC one blind man sit.down-PRF-PRS
 ‘A blind man is sitting by the side of the road.’
- (138) *sedag pakɔ-a aɖi tɔŋɔn-ce-re*
 gate near-LOC who stand up-PRF-PRS
 ‘Who is standing at the gate?’
- (139) *arɔ aŋɖeka tɔŋɔn-ce-ran.*
 3 yesterday stand up-PRF-PST
 ‘Yesterday he was standing.’
- (140) *suruma jiarɔ aliŋta dɔkɔ-se-ran.*
 Suruma chair on sit.down-PRF-PST
 ‘Suruma was sitting on a chair.’

There are also instances where */-se/* doesn’t occur with perfect. Simple past is used in case of present perfect in these cases.

- (141) *aiŋ-a kaka-iŋ ekolon baibɔ ɔnɔ*
 I-GEN brother-1 just.now out go-PST
 ‘My brother has just gone out.’

The following examples illustrate instances in which */-se/* is used in case of imperatives or in simple future.

- (142) *ape iatɔ-se-na beɔte gaɖi ecɔ[ao-rɔ*
 2.PL be tired-PRF-FUT time car drive-PROHIB
 ‘When you are tired, don’t drive.’
- (143) *nɪn nɪn-a nɪn-a ni-kɔ-ɔŋ-se-na.*
 we we-GEN we-GEN 1PL-RECIP-see-PRF-FUT
 ‘We will see each other.’

The perfect marker */-tel/*. */-tel/* is another perfect marker in Juang which marks completion of an action in present as well as in the past. Like */se/*, even */tel/* can be taken as a present as well as past perfect marker. The marker */-tel/* is rather tricky in this language because the infinitival clause is also marked with */-tel/*. */tel/* occurs in several other positions too.

- (144) *belo budi-an-te.*
 sun set-PST-CMPLT
 ‘The sun has set.’

Sometimes */-tel/* occurs along with the perfect marker */-se/* or the variant of it.

- (145) *aiŋ soi-ce-ke-te.*
 I learn-PRF-PRS-CMPLT
 ‘I have learnt.’

There are instances where a suffix such as */-ro/* occurs to mark perfect aspect though it is not common. One such instance is the following.

- (146) *nai boi-ke-ro*
 river flow-PRS-PRF
 ‘The river has flown.’

The progressive aspect /-*nɔm/* or /-*nɔman/*. The progressive aspect marker /-*nɔm/* or /-*nɔman/* occurs with reduplicated verb bases only when they indicate continuation. In reduplication the final consonant of the first syllable gets deleted. When the stem ends with a consonant, it coalesces with the initial /*n/*.

- (147) *ki + nɔman* → *kikibɔman* 'was doing'
jim + nɔman → *jɪnjimɔman* 'was eating'
leber + nɔman → *leberɔnɔman* 'was sleeping'
- (148) *pancota be|oa bolia-te ebodom-oman.*
 five o'clock field-LOC play-IMPF
 'At 5 o'clock she was playing in the field.'
- (149) *aŋɖeka smesar ateta be|a am biri ki-kib-noman?*
 yesterday morning eight o'clock you what REDPL-DO-IMPF
 'What were you doing at 8 o'clock yesterday morning?'
- (150) *banɔ juaŋɖai-kia dando-te ko|eŋom-oman-kia.*
 two Juang.woman-DL road-LOC quarrel-IMPF-DL
 'Two women were quarrelling in the street.'
- (151) *gataj barɔ i-noman luko-re kala.*
 road cross be-IMPF man-ART deaf
 'The man crossing the road is deaf.'
- (152) *aŋɖeka ain muinto juangɖai-te kui-sero ju*
 Yesterday I one woman-ACC meet-PERF:PST who
panco-go|a gata gata-noman.
 five-CLSSFR language speak-IMPF
 'Yesterday I met a woman who can speak five languages.'

3.2.5 Mood

Various suffixes are added for indicating modal variation which occurs before number suffixes and after aspect suffixes. In subjunctive and conditional sentences /*tan/* is used for simple present; /*nomtan/* for imperfect/progressive and /*etan/* for perfect. Instead of /*nomtan/*, /*nomsetan/* is also used in instances indicating imperfect aspect. In habitual usage, /*nom-de tan/* and /*setan/* are interchangeably used depending on tense and aspect suffixes. In imperative /*mel*, /*pel* and /*rel* etc. are used depending on the person.

- (153) *Subjunctive*
 Examples of subjunctive forms are given below for the verb /*kib/* meaning 'do':

Simple subjunctive	<i>ara kib-tan</i>	<i>araki kib-tan-ki</i>
Imperfect subjunctive	<i>ara kipki-nomtan</i>	<i>araki kipki-nom-tan-ki</i>
Perfect subjunctive	<i>ara kib-setan</i>	<i>araki kib-setan-ki</i>
Simple subjunctive	<i>ain kib-tan</i>	<i>nin mi-kib-e</i>
Imperfect subjunctive	<i>ain kip-nom-setan</i>	<i>nin ni-kib-setan</i>
Perfect subjunctive	<i>ain kib-setan</i>	<i>nin ni-kib-e</i>
Simple subjunctive	<i>am mi-kib-tan</i>	<i>apperiki i-kib-tan</i>

Imperfect subjunctive	<i>ammi-kip-momtan-kilmi-kipki-setan</i>	<i>i-kipki-nom-setan</i>
Perfect subjunctive	<i>am kip-setan-ki</i>	<i>apperiki i-kip-setan</i>

(154) *Hortative*

The following examples illustrate the hortative forms of the same verb */-kib/* meaning ‘do’.

Simple hortative	<i>ara</i>	<i>kib-e-me</i>
Imperfect hortative	<i>ara</i>	<i>kikip-nom-deme</i>
Perfect hortative	<i>ara</i>	<i>kip-seke-me</i>
Simple hortative	<i>am</i>	<i>ki-be</i>
Imperfect hortative	<i>apa banog</i>	<i>kikip-moma-pe</i>
Perfect hortative	<i>aperiki</i>	<i>ni- kip-ke</i>

The affirmative–imperative markers are */-kul* and */-qel* for first person and third person. */qel* is used in cases when the speaker requests the hearer for something or persuades the hearer.

(155) *nijba-ku-ɔnɔ*

1DL-IMP-GO

‘Let both of us go.’

The second person imperative is formed by adding suffixes */-pa/* and */-pe/* along with the future marker */-nal/*.

(156) *doko-na*

sit-IMP.II

‘Sit down! (You alone)’

(157) *qak-e*

call-IMP.I

‘Call! (You alone)’

(158) *doko-na-pa*

sit-IMP.II-2DL

‘Sit down! (You two)’

(159) *doko-na-pe*

sit-IMP.II-2PL

‘Sit down! (You all)’

There are some verbs from class-II type which can take */-el/* and */-nal/*.

In negative imperatives */rol* and */-jenal* are used. */rol* is added to the past form of the verb, and */-jenal* is added to reduplicated stem or to the stem with */-kel/* as in present tense. A variant of */rol* is */-dol/*. */-dol/* is used with class-II verbs.

(160) *mi-jiŋ-o-ro*

2-ask-PST.I-PROHIB

‘Don’t ask’ (You alone)

(161) *me-toŋon-an-do*

2-stand-PST.II-PROHIB

‘Don’t stand’ (You alone)

(162) *a-jiŋ-o-ro*.
2DL-ask-PST.I-PROHIB
'Don't ask. (You two)'

(163) *i-jiŋ-o-ro*.
1PL-ask-PST.I-PROHIB
'Don't ask. (You all)'

Note that *-o* is the past suffix of class-I, *-an* of class-II.

More examples follow.

(164) *kanto me-ŋeŋe-o-ro*.
wall 2-jump-PST-PROHIB
'Please jump over the wall.'

(165) *aŋ-a paī me-lelej-an-do*.
I-GEN for 2-get.angry-PST-PROHIB
'Don't get angry with me!'

(166) *mino m-ono-ro*
alone FUT-go-NEG
'Don't go alone.'

3.2.6 Orientation/directionality

Not investigated for Juang.

3.2.7 Voice

Causatives are formed by affixations such as *l-b-l*, *lu-l*, *lo-l*, *la-l*, *lap-l*, *lam-l*, and *lab-l*.

(167)	<i>buŋur</i> 'get up'	<i>bu-b-ŋur</i> 'cause to get up'
	<i>tij-</i> 'carry on head'	<i>u-tij</i> 'cause to carry on head'
	<i>boi-</i> 'blow'	<i>o-boi</i> 'cause to blow'
	<i>bug-</i> 'play'	<i>a-bug</i> 'cause to play'
	<i>peŋe-</i> 'strike'	<i>a-peŋe</i> 'cause to strike'
	<i>maje-</i> 'rinse'	<i>am-maje</i> 'cause to rinse'
	<i>ŋim-</i> 'pinch'	<i>ab-ŋim</i> 'cause to pinch' (Matson 1964:26)

Juang has active as well as passive voice. Passive voice is made by the infixation of *l-ŋoman-l*, *l-riman-l*. Sometime passive is formed by just exchanging the position of the subjects and objects.

(168) *polici curɔ-re-ki-te sɔb-ɔ*
police thief-DEF-PL-ACC catch-PST
'Police caught thieves'

(169) *curɔ-re-ki sɔb-ŋoman-ki*
thief-DEF-PL catch-PSSV-PL
'The thieves got caught.'

(170) *juanŋdai-re-te muinŋo nua-ra ganŋua ŋi-ŋa-ki*
woman-DEF-ACC one new-DEF basket give-1-PST
'The woman was given a new basket in the village.'

- (171) *ajñ muinto juangadai-ra-te nua ganɔua dʒina*
 I one woman-DEF-ACC new basket gave
 ‘I gave basket to the woman.’

3.2.8 Finiteness

See section 4.2 on Syntax.

3.2.9 Negation

Negation is formed by several affixes depending on the tense and person/number features of the subject or the object with which the verb agrees.

The following table illustrates various negative markers for various persons and numbers.

(172)	Pronominal	Person negative marker	General negative markers
	1 SG	∅	<i>-jena</i>
	1 DL	<i>b-</i>	<i>ama-, mama-, am-, ab-</i>
	1 PL	<i>n-</i>	<i>-na</i>
	2 SG	<i>m-</i>	<i>nɔ-</i>
	2 DL	<i>a-</i>	<i>-rɔ</i>
	2 PL	<i>e-</i>	
	3 SG		
	3 DL	<i>-kia</i>	
	3 PL	<i>-ki</i>	

Some of the negative markers are prefixed and some are suffixed. There are instances of ambifixation and double negatives as well. The negated verb is sometimes reduplicated.

Reduplicated stem + *-jena*

- (173) *ajñ ɔn-ɔn-jena*.
 I REDPL-go-NEG
 ‘I don’t go.’

- (174) *ajñ kui-kui-jena*
 I REDPL-receive-NEG
 ‘I haven’t received.’

- (175) *ajñ jim-jim-jena*
 I REDPL-eat-NEG
 ‘I didn’t eat.’

As stems with more than one syllable don’t reduplicate, these verbs are negated simply by addition of *ljenal*.

- (176) *ajñ pera-i-jena*
 I return-AUX-NEG
 ‘I haven’t returned.’

- (177) *ajñ cake-jena*
 I taste-NEG
 ‘I have not tasted.’

There are certain instances where reduplication is expected to be possible but it doesn't occur.

(178) *ajɲ on-jena*
I hear-NEG
'I don't hear.'

(179) *arɔ i-jena*
he stop-NEG
'It has not stopped.'

The following are some of the possible combinations of double negatives.

(180) *Combination I*
a- + -jena 'NEG + NEG'
ama- + -jena 'NEG + NEG'

(181) *Combination II*
a+ -na 'NEG + NEG'
a+ ma- + -na 'NEG + NEG'

(182) *Combination III*
ma+jena 'NEG + NEG'
ari+ jena 'NEG + NEG'
aire+ jena 'NEG + NEG'
bama + jena 'NEG + NEG'

(183) *Examples*
a-den-de 'not driving out'
ama-kɔŋ-ke 'not knowing'
apa-ɖeke-jena 'not calling'
ama-ɔn-de-jena 'not coming'
a-ɖen-de-na 'not running away'

3.2.10 Derivation

As discussed in section 3.1.6, Juang has a verbal reciprocal which is /*ku*/. It doesn't have a verbal reflexive. The reciprocal prefix *ku-* is added directly to the root. Person and negative prefixes precede it. *ku-* has the allomorph *ko-* before roots with mid-high vowels (*e, o*).

(184) *ne-ku-lona-na*
IDL-RECIP-look-FUT
'We (all) will look at each other.'

(185) *a-ku-lona*
2DL-RECIP-look
'You (two) looked at each other.'

(186) *e-ku-lona*
2PL-RECIP-look
'You (all) looked at each other.'

(187) *arɔ-kia banɔʔ-bɔ diɔ a-ku-buji-ri-kia.*
3DL two-DIR well NEG-RECIP-love-PRS-DL
'They don't like each other.'

- (188) *arakara-kia dio ku-buji-ri-kia.*
 each other-DL well RECIP-love-PRS-DL
 ‘We stop talking (to each other).’

3.2.11 Nominal incorporation

Noun incorporation is not as productive in Juang as it is in some other South Munda languages. Instances of noun incorporation are noticed in case of body parts being used with verbs like ‘wash’.

- (189) *am am-a itim-dɛ mi-gui-di-agan*
 you YOU-GEN hand: 2-DEF 2-wash-hand-PST
 ‘You washed your hand.’

‘*iti*’ ‘hand’ becomes *tildi* in incorporated forms.

3.2.12 Auxiliary verb constructions

Not particularly investigated for Juang.

3.3 Expressives

Not investigated.

4 SYNTAX

4.1 Simple sentences

Juang is a verb-final language. It has most of the verb-final features. In sentences with a ditransitive verb, the unmarked order of the constituents in a sentence is subject-indirect object/(indirect object)-direct object/(direct object)-verb. Yes-no questions and tag-question markers occur post-verbally. Juang is a pro-drop language. There are no pleonastic or expletive elements in this language. It has a relatively free word order for pre-verbal elements.

Simple sentences can be in declarative, passive, active and causative. All these sentences can be negated and interrogated. As mentioned before, there can be dative subject and genitive subject constructions in Juang.

- (190) *ajɲ dʒɔ asi-ke*
 I well be-PRS
 ‘I am fine.’
- (191) *ajɲ am-te kɔŋ-sək-jena*
 I YOU-ACC know-AUX-NEG
 ‘I don’t know you.’
- (192) *ajɲ-te rasiŋ gonɔ-ke.*
 I-DAT flower smell-PRS
 ‘I can smell the flowers.’

- (193) *aj̄n-te emalō i-sere*
 I-DAT cold be-PRF
 'I have got a cold.'
- (194) *aj̄n-te ekoloŋ ɔnɔn-te mina*
 I-DAT now REDPL:GO-INF FUT
 'I must go now.'

Juang has constituent questions. Question words can take case markers. The question words are: *ađi* 'who', *ađite* 'whom', *ađiaba* 'whose', *ađialiŋa* 'on what', *ađiare* 'where'. Question words cannot really be reduplicated as productively as in the superstratum Oriya.

- (195) *ij̄na-re ađi bɔna-e-be*
 home-LOC who make-FUT-Q
 'Who will make the house?'
- (196) *ij̄na-re ađi a-bɔna-ere-ki*
 house-LOC who NEG-make-FUT-Q
 'Who doesn't make the house?'
- (197) *apa biri mi-kib-ɔ-ɖe*
 you what 2-DO-FUT-Q
 'What do you do?'
- (198) *apa biri ama-kib-ɖe*
 you what NEG-DO-Q
 'What don't you do?'
- (199) *apa a ɔ-na-tɔ*
 You EMPH GO-FUT-EMPH
 'Will you go?'
- (200) *ē, aj̄n m-ɔn-a*
 yes, I FUT-GO-FUT
 'Yes, I will go.'

Juang allows multiple question words in simple sentence. The question words can occur anywhere in the sentence because Juang is a free word order language.

- (201) (a) *bite ađi-te ađi ɖij̄n-ɔ*
 What who-ACC who give-PST
 'Who gave what to whom?'
- (b) *ađi-te bite ađi ɖij̄n-ɔ*
 (c) *ađi bite ađi-te ɖij̄n-ɔ*

Sentences in Juang can take yes–no echo question.

- (202) *aj̄n am-te ɖio buji-ke*
 I you-ACC well do-PRS
 'I love you?'
 (Do I love you?)
- (203) *am biri m-ur-e*
 you what 2-eat-FUT
 'Will you eat?'

Here the question word *biri* is used with the dubitative marker.

Anaphors find their antecedents inside the same domain. Anaphors in Juang may appear as *lapeinl*. The Oriya loan *lnijol* also occurs as an anaphor. The pronouns often find their antecedents in the same domain too. Detailed discussion on this topic can be found in Patnaik (1996, 2000).

- (204) *ain_i ain-te_i diɔ bodice*
I I-ACC like do-PRF
'I love myself.'
- (205) *sita nijɔ-te bɔi muintɔ sɔgan*
sita self-DAT book one bought
'Sita bought a book for herself.'
- (206) *sita nijɔ-te muintɔ babu jɔyɔ*
sita self-DAT one snake saw
'Sita saw a snake near herself.'
- (207) *gita nijɔ iɲa-re rɔere*
gita self house-LOC stay
'Gita stays at her own place.'
- (208) *ramɔ nijɔ-te nije ekɔ bɔi diɲ-ɔ*
ram self-DAT self one book give-PST
'Ram gave a book to himself.'
- (209) *sita arɔ-rɔ minɔg gata-nɔman.*
sita self-GEN with talk-PROG
'Sita was talking to herself.'
- (210) *sita ara suke arɔ lejdɔman*
sita she on she angry.was
'Sita got angry on herself.'
- (211) *ram gata-nɔman shyam aram-ke arɔ prɔsɔnsa kiban.*
ram say-PROG shyam he-ACC he praise do:PST
'Ram was saying Shyam praises himself.'
- (212) *ram shyamɔ-te ara iɲa baɲbaɲndɔ kete me-gata-yɔrɔ gamɔ*
ram shyam-ACC self house burn about NEG-say-INF say:PST
'Ram told Shyam not to tell anybody about self's burning house.'
- (213) *sɔbie nijɔ-te budi abamdeki*
every body self-ACC intelligent think:PRS
'Everybody thinks himself to be intelligent.'
- (214) *aɟi apeɲ kɔsɔt didiɲ-te cahete*
who self pain REDPL:give-INF want
'Who wants to hurt himself?'

Equational sentences. There can be sentences in Juang that do not carry a verb. They are called equational sentences. Such sentences are usually combinations of nouns with nouns, nouns with adjectives, nouns with adverbs, and nouns with question words.

- (215) *puri muiŋɔ sanɔ jaga*
 puri one small place
 ‘Puri is a small place.’
- (216) *puri muiŋɔ sanɔ jaga jena | kuba jaga*
 puri one small place NEG
 ‘Puri is not a small place.’
- (217) *arɔ dɔsɔmɔ sreŋilrɔ chaɬuapila*
 he tenth class:GEN student
 ‘He is a student of class ten.’
- (218) *ajɲ-a mamu-ɲɲ muinta mastero*
 I-GEN uncle-1 one teacher
 ‘My uncle is a teacher.’
- (219) *arɔ-ki dʒiɔ lokɔ-ki jena*
 he-PL good people-PL NEG
 ‘They are not good people.’

4.1.1 Typological features

Juang is a verb-final language.

- (220) *ajɲ tɔŋɔ ur-ke*
 I rice eat-PRS
 ‘I eat rice.’

The verb ‘be’ doesn’t surface when the verb is in the present tense. However, the verb ‘be’ surfaces when it is in the past tense.

- (221) *abu muiŋɔ dakotoro*
 Abu one doctor
 ‘Abu is a doctor.’
- (222) *sita men-ce-na catuapila asi-ana*
 sita last year student be-PST
 ‘Sita was a student last year.’

In sentences with a predicative adjective, the verb ‘be’ is not overtly present.

- (223) *ini polo-re dʒiɔ*
 this fruit-DEF good
 ‘This fruit is good.’

The forms of the verb ‘be’ and ‘become’ are different. The form for ‘be’ is *asi* and ‘become’ is *i*. The verb *i* ‘becomes’ carries the tense marker as shown in the following example:

- (224) *ɔlen dʒag i-an*
 Ice water become-PST
 ‘Ice became water.’

In the unmarked order adverbs occur pre-verbally. Adverbs do not normally carry postpositions or case markers though they can.

Yes–no question markers occur post-verbally.

- (225) *am ara-te buguca-te mi-ki-ro*
 You he-ACC garden-LOC 2-beat-Q
 ‘Did you beat him in the garden?’

Like all Munda languages, the non-singular imperative markers are */-pal* and */-pel* in the affirmative. In Juang */-rol* and */-rel* occur in negative imperative forms, and they occur post-verbally.

- (226) *apa somogarep ina-bo on-a-pe*
 you all home-LOC go-IMP-2PL
 ‘All of you go home.’

- (227) *tear me-den-ro*
 tomorrow 2-come-NEG.IMP
 ‘Don’t come tomorrow’.

Like typical Munda languages, the subject in Juang does not carry any nominative case marker as such. Definite markers occur with subject noun phrase. With non-subject noun phrases, definiteness is manifested either by a lexical accusative or a definite marker when the noun phrase is [+animate]. Thus, when the direct object is definite, either the definite marker occurs or the lexical accusative case marker occurs. When the direct object is [-animate], the accusative marker does not normally occur as is the case with many South Asian languages.

- (228) *Itu merəm-te kiʃg-a asara ɔnu*
 Itu goat-ACC tiger-GEN near keep:PST
 ‘Itu kept the goat near the tiger.’

- (229) *ar-a uaʃi-rehualʃi-qi-te aijn-ya alij-te pəkao*
 he-GEN child-DEF/child-DEF-ACC I-GEN on-LOC keep:PST
 ‘He kept the child on me.’

- (230) *itu ere boi-re ain-te dij-o*
 itu that book-DEF I-DAT give-PST
 ‘Itu gave the book to me.’

The modifier can precede as well as follow the noun in noun modification.

- (231) *kote sunduru*
 cloth beautiful
 ‘Beautiful cloth’

- (232) *sunduru kote*
 beautiful cloth
 ‘Beautiful cloth’

In qualifying adjectives, the qualifier precedes the adjectives and adverbs precede the verb.

- (233) *sugei qiɔ*
 very nice
 ‘Very nice’

- (234) *sima sugai tiri dajin-ke*
 Sima very slow walk-PRS
 ‘Sima walks very slow.’

Gapping is not very consistent in Juang. The bilingual Juangs prefer gapping whereas the monolingual Juangs do not prefer gapping. Backward gapping is preferred to forward gapping.

- (235) *ain karəŋ jim-ə aru təm mətəra jim-ə*
 I fish eat-PST and Tom peas eat-PST
 ‘I ate fish and Tom peas.’

The conditional clause always precedes the matrix clause.

- (236) *ara judi gəb-bə ən-de ajn-te pacola m-ə-r-ene*
 he if Keonjhar-DIR go-COND I-DAT shawl FUT-bring-FUT
 ‘If he goes to Keonjhar, he will bring a shawl for me.’

The inclusive, exclusive and emphatic particles follow the noun.

- (237) *tanu-jə ere-bə məən-a*
 Tanu-EMPH there-LOC 3:FUT:go-FUT
 ‘Even Tanu will go there.’

Juang employs reduplication as a strategy for morphological and syntactic processes, which is very widespread in most Munda languages. Reduplication is noticed for participles, infinitives and negatives. It performs an intensifying function too.

- (238) *ara-ka enteij enteij ualji-ki asi-ki-ki*
 he-GEN small small child-PL be-PRS-PL
 ‘They have little little babies.’

- (239) *ajin ini komo kikib-te buluŋ lanka den-se-re*
 I this work REDPL:DO-INF very far come-PRS-PRF
 ‘I have come so far to do this work.’

- (240) *ajin ini komo ki-kib-jena*
 I this work REDPL:DO-NEG
 ‘I don’t do this work.’

The dubitative marker *-ri* is post-verbal. It always occurs on both disjuncts.

- (241) *ara de-me-nari jena-ri ajin a-konko-jena*
 he come-FUT:DUBIT NEG-DUBIT I NEG-know NEG
 ‘I don’t know whether he will come or not.’

4.2 Complex sentences

This section presents in brief the way sentences in Juang can be coordinated, subordinated, and compounded using various devices.

Juxtaposition is one of the ways of sentence coordination. In this type, sentences get coordinated without any markers indicating coordination.

- (242) *ki|əg dakeə gau|ə dən-jə gucae-ə*
 tiger call-PST milkman come-NF save-PST
 ‘The tiger called and the milkman came and saved him.’

The dubitative marker /*ri/* which is also homophonous with the question marker /*ri/* occurs when two sentences get coordinated in the disjunctive ‘but’-construction.

- (243) *jən ɖen-gər ri-mima mari ɖen-gər arim-ɖe jena*
 John come-FUT but Mari come-FUT possible-DEF-NEG
 ‘John can come but Mary can’t.’

Juang borrows the Oriya coordinator /*laul*/ to mark coordination. Most often the coordinator /*laul*/ doesn’t occur and sentences get juxtaposed.

- (244) *aiŋ ere ʈukikən-ɖae-te muɪŋɔ kələmɔ ɖiŋ-ɔ au kənən-te*
 I that girl-FEM-ACC one pen give-PST and boy-ACC
muɪŋɔ penɔsilɔ
 one pencil
 ‘I gave a pen to that girl and a pencil to the boy.’

In the following example the coordinator ‘and’ doesn’t occur.

- (245) *arɔ ʃiʈu-te gaən-bɔ gɔg-ɔ gitu-te iŋa-bɔ*
 he jitu-ACC village-LOC take-PST gitu-ACC house-LOC
 ‘He took Jitu to the village and Gitu to home.’

There are instances of conjunction reduction in Juang in case of relative clauses: the correlative marker and the pronoun of the embedded clause get deleted in the second coordinate. In the place of a relative marker the correlative marker occurs.

- (246) *ju lukɔ-re-te aiŋ jɔy-ɔ-ju bulu gɔe-jɔ sugai*
 REL man-DEF-ACC I see-PST-REL Bulu hit-PRF be.PST
kɔrapɔ iana.
 bad was
 ‘The man who I had seen and Bulu had hit he was very bad.’

Gapping is a very productive phenomenon in coordination. Gapping is usually forward in nature in which case the second occurrence of the noun gets deleted. The conjunction marker doesn’t necessarily occur in these cases.

- (247) *aiŋ mɔdɔ urke-ju arɔ urur jena*
 I alcohol drink-REL he REDPL:drink NEG
 ‘I drink alcohol but he doesn’t drink alcohol.’

Like nouns even the second occurrence of the same verb can be put to deletion in the second coordinate structure.

- (248) *jən karɔŋ ʃim-ke-ju bil-ɔ manso*
 John fish eat-PRS-CONJ Bill meat
 ‘John eats fish and Bill meat.’

- (249) *ramɔ selɔg meɔjɔ bilei-te*
 ram dog killed cat-ACC
 ‘Ram killed the dog and killed the cat too.’

Equatives. Sentences get coordinated with the use of equatives too. Juang doesn’t have any equative marker of its own. The equative marker is borrowed from the superstratum Oriya.

- (250) *gita jitiki saparɔ sita jɔ sitiki saparɔ*
gita as much beautiful sita also that much beautiful
 ‘Gita is as much beautiful as Sita.’

Conditional clause. The marker *l-jɔl* is used to mark conditionality in both counterfactual/irrealis and realis conditions.

- (251) *apa ama den-ɔ-jɔ ajn ɔn-tan*
 You NEG come-PST-COND I go-SBJNCT
 ‘If you would have come I would have gone.’
- (252) *nijn nama ne-lara-serɔ-jɔ arɔ jek-setan*
 We NEG 1DL-laugh-PRF-COND she cry-SBJNCT
 ‘If we would not have laughed, she would not have wept.’
- (253) *arɔ a-ur-ke-rɔ-kɔ-ju a-ur-ɖeme*
 he NEG-eat-PRS-DEF-COND NEG-eat-CONT
 ‘If he does not want to eat let him not eat.’

Reason clause. Another type of subordination is done by the use of *letel* meaning ‘because’. The reason clause is marked by the *letel* which has been grammaticalized to become to a finite complementizer. The details of this category have been discussed in Patnaik (2000).

- (254) *apa a-ma-jim-ke ete ajn kikib-jena*
 You 2DL-NEG-eat-PRS that/because I REDPL:do-NEG
 ‘I didn’t do it because you don’t eat it.’

There are two other types of sentential subordination, namely finite and non-finite. Finite sentential subordination includes the relative clause formation and the complementizer clause formation. Non-finite sentential subordination includes gerunds, the infinitival construction and the participial construction.

The relative–correlative construction in Juang consist of two clauses that occur immediately preceding the head noun in unmarked order. However, there are no such restrictions on the order of the relative clause. The relative marker is ‘*ju*’ and the correlative markers in the main clause are the personal pronouns *aro*, or the demonstrative pronoun *ere* or *auri*.

- (255) *ju luko-re-te ajn rubaɣ-ɔ arɔ ajn-a jo[a-n-a kaka-ro*
 REL man-DEF-ACC I meet-PST he I-GEN friend-I-GEN brother-DEF
 ‘The man who I have met is my friend’s brother.’
- (256) *auri suduru bhangi kuru-cere ere ajn-a jo[a-n.*
 that red shirt wear-PRF that I-GEN friend-I
 ‘The boy who is wearing a red shirt is my friend.’

The complementizers in Juang are *ju*, *ete*, *gamojo* and *buli*. Of these, *ju* is a left peripheral complementizer, whereas *ete* and *gamojo* are right peripheral complementizers. *Buli*, which also occurs as a right peripheral complementizer, is a loan from the Oriya right peripheral complementizer *boli* which is a grammaticalized form of the verb *bol-* ‘say’. *ju* also resembles the left peripheral complementizer of Oriya, that is, *je*. However, it is not easy to say whether *ju* is borrowed from Oriya or originally

it was there in Juang. Apart from them, there are complement clauses which are not marked by any complementizer. Let us take a look at these constructions first.

- (257) *aro gan-gam-moman- [ju Ramo q̄emen-a].*
 he REDPL-say-PROG – COMP R. come-FUT.
 ‘He was telling that Ram would come.’
- (258) *ain babe-sero-ju aro nice gundugu[o kimib-e.*
 I think-PRF-COMP he surely trouble do/3:FUT/-FUT
 ‘I had thought he would definitely create troubles.’
- (259) *ain oŋ-s-ro-ju aro muinŋo muruko.*
 I hear-PRF-COMP he one fool
 ‘I had heard that he is a fool.’
- (260) *am [Ramo Sita-te kania m-un-e ete me-gam-o-jo]*
 you Rama Sita-ACC marriage FUT: 3-do-FUT COMP 2-say-PST-COMP
me-gam-o
 2-say-PST
 ‘You said Ram will get married to Sita.’
- (261) *arokia [sina muinŋo godo ete gam-o-kia-jo] gam-o-kia.*
 They:DL Sina one fool COMP say-PST-DL-COMP say-PST-DL
 ‘They (two) said Sina is a fool.’
- (262) *aro am-te godo ete gam-o-jo gam-ke.*
 he you-ACC fool COMP say-PST-COMP say-PRS
 ‘He says you are a fool.’

Infinitival clauses. Interestingly *ete*, which behaves like a finite right peripheral complementizer, also occurs with non-finite complement clauses.⁵ In this function, it always occurs attached to the verb, whereas as a finite complementizer it occurs separately. Consider the following examples.

- (263) *bulɔbulɔ q̄uq̄uŋ-te jage-re.*
 Bulbul REDPL:GO-INF want-DEF
 ‘Bulbul wants to go.’
- (264) *aro kuŋumo musi-eteere cakiri kib-o.*
 he family manage-INF job do-PST
 ‘He did the job to manage his family.’

Gerunds. Gerunds are constructed much like the infinitive but the */tel/* clause takes the definite marker.

- (265) *Ramo ere-bo on-on-q̄e-te ke q̄io a-buji-ki-ki.*
 Ram there-LOC REDPL-go-DEF-NF no one good NEG-understand-PRS-PL
 ‘No one likes Ram’s going there.’

Participle clauses. There are three types of participial clauses, namely perfect participial clauses, imperfect participial clauses, and conjunctive participial clauses.

(i) *Conjunctive participial clauses*

- (266) *bulbulinja-bɔ q̄enɔ-ɔ jiray-an-ɔ.*
 Bulbul house-LOC come-PRTCPL rest-came-PST
 ‘Having come home Bulbul took rest.’

When the conjunctive participle gets negated /*jol*/ becomes /*rol*/ and the negative prefix /*a-*/ is added to the root clause.

- (267) *arɔ kici a-unu-rɔ ba-arian*
 he anything NEG-eat/NMLZ/-NEG.PRTCPL leave-PST-PRF
 ‘Having not eaten anything he left.’

Sometimes the conjunctive participle gets negated simply by reduplicating the root verb.

- (268) *arɔ tɔɔŋŋɔ urur-rima-ri ruʔi jim-ɔ*
 he rice REDPL:eat-CAP-DUB chapati eat PST
 ‘Having not being able to eat rice he ate chapatti.’

(ii) *The imperfective participial clauses.* The root verb gets reduplicated and it carries /*-noman/* to construct imperfect participial constructions.

- (269) *gɔʔaŋ-te ɔngo-r-ongɔ-re-nɔman ere beʔa bulɔbulɔ*
 road-LOC REDPL:go-DEF-PROG that time Bulbul
lukɔrɔ-re-te jɔyɔ
 man-DEF-ACC see:PST
 ‘Bulbul saw the man while going on the road.’

(iii) *Non-finite conditional clause.* The non-finite conditional clause is formed by suffixation of /*-jɔ/* to the matrix verb.

- (270) *dɔrkarɔ ian-jɔ rusi saajɔ dimin-je gam*
 necessary happens-PRTCPL rusi help will:do said
 ‘Rusi said he will help, if necessary.’
- (271) *ɔlej ina-bɔ ɔngɔnma-bela gɔʔaŋ-te muinʔɔ kiʔɔ-te rubaŋ-ɔ*
 cow house-LOC while,going-at road-LOC one tiger-ACC meet-PST
 ‘The cow met a tiger while going home.’

5 SEMANTICS

Not specially investigated for Juang for this chapter.

6 LEXICON

6.1 Austroasiatic/Munda components

I provide below a set of selected words which are native to the tribe and language.

Domestic items

<i>ina</i> – home	<i>urunina</i> – kitchen
<i>utan</i> – winnowing fan	<i>ulir</i> – carrying rope
<i>kaka</i> – bow	<i>kɔsɔkɔm</i> – arrow

Selected body parts

<i>ɔmɔr</i> – eye	<i>ɔilaŋ</i> – tongue	<i>lutur</i> – ears
-------------------	-----------------------	---------------------

<i>itip</i> – belly	<i>iti</i> – hand	<i>ɔmɔak</i> – tears
<i>mɔteij</i> – nose	<i>ijij</i> – leg	<i>redijij</i> – heel
<i>sagiti</i> – left hand	<i>ɔiɔiti</i> – right hand	

6.2 Loan strata

Superstratum loans are quite prominent in Juang in a few cases. The entire number system other than the first four numbers and a few others has been borrowed from Oriya. As can be seen from the texts given here, the very term ‘Rusi’ whose descendants are the Juang is supposed to be is an Indo-Aryan word. There are many words in today’s Juang which are borrowed from Oriya. Some of the syntactic structures have been influenced by Indo-Aryan languages. The occurrence of retroflex sounds could be Dravidian influence via Indo-Aryan languages. The word structure in case of gender and adjective formation is influenced by Oriya. Kinship terms still follow the Munda pattern though.

The usage of */ka/* as a genitive seems to be from Hindi. Loss of final vowel in some words could be due to Hindi influence.

A few English words like ‘jail’, ‘court’, ‘school’ etc. seem to have been borrowed from Oriya as well.

7 BRIEF ANALYZED TEXTS

Juangs feel very proud to speak about their origin as one of the most ancient tribes of human civilization. The first story they speak when they are asked to tell a story is the story that talks about their origin. This story dates back to prehistoric days when there were two people, a man and a woman, who existed on this earth that was created by the God of Justice (Justice God). Those two people grew up to be saints, and they lived on a hill near Gonasika.⁶ All other places in the world were surrounded by water those days which is why not many humans were there. Humans basically had no place to live. Eventually the couple had a child. Justice God then decided to dry up a few places so that he could keep some humans. He had to do something very drastic to be able to do that. He had to kill the saint’s child and throw his blood everywhere. He asked the parents of the child whether he could carry out such an act. The father agreed, but the mother did not. The Justice God then sent a tiger to kill the child. The child was very brave, and he had mastered the art of protecting himself against wild animals with the help of bow and arrow. So when he saw the tiger he killed him at once using his sharp arrow. Then the Justice God had to devise another plan. Justice God told the father to prevent his child from taking the bow and arrow when he would go to the river to bathe. The obedient child did that. He did not carry protection along with him when he went to the river to bathe and finally a tiger killed him. His blood was thrown everywhere. That caused the water to dry up from lots of places. His limbs were thrown to several places which gave rise to mountains and hills. His hair was thrown out to several places too which gave rise to plants and trees. After that the couple had 12 sons and 12 daughters. The children of those children got to be known as Juangs.

This story was told to me by several people during my visit to the field. The first person who had written down this story for me was Kasturi Juanga, who worked as

an office assistant in a government college in Keonjhar district. She had apparently given the same story to Ms Sashmita that appears in a small sketch book prepared by her in 1991 for the Tribal Welfare Department of Government of Orissa. The story in that book doesn't have glosses though the story lines are the same as the ones given to me. This shows the exactness of the oral tradition that has been carrying this story on for ages. I heard this story lines from many people belonging to many age groups. All of them told the story exactly the same way, using the same language and intonation, as Kasturi did.

Needless to say, oral tradition is very strong in Juang which is shown from the songs and folk tales they carry. I have collected hundreds of stories and folktales which reinforces the strength of Juang oral tradition. As mentioned above, they still don't have a script system of their own and the education system is one of Orissa government in which the medium of schooling is Oriya.

The story as given below has few Oriya loans like 'subida' meaning 'comfort' as in 'subidha' of Oriya; 'sina' is an Oriya loan which means 'could' in a negative way; 'rusi' meaning 'saint' is the Indo-Aryan word for 'rusi'. How could so many Oriya loans could get into a Juang folk tale remains to be studied by folklorists.

7.1 Text 1

- (i) *juang-ka jɔnɔmɔ gata*
 Juang-GEN born story
 [The story of Juang's birth]
- (ii) *ɔnɔ-ga uruŋa gata*
 went-GEN old story
 'This is a story of very old days.'
- (iii) *dɔrɔm kɔilɔŋ pɔrtibi-te ɔmbaɬɔ muŋusa suruʃti kib-sere*
 Daram God world-LOC two people create do-PRF
 'A god called Daram God created two people in this world.'
- (iv) *arɔ-kia i-an-kia rusi au rusi-aŋi*
 he-DL be-PST-DL saint and saint-FEM
 'Both of them became saints.'
- (v) *arɔ-kia gɔnasika nikuʃti-a muinɬɔ uli-a rɔe-an-kia*
 he-DL gonasika near-GEN one hill-GEN stay-PST-DL
 'They lived on a hill near Gonasika.'
- (vi) *mulukɔ-a sabi-ete ɬa²k asi-an*
 area-GEN everywhere-LOC water be-PST
 'The entire area was full with water.'
- (vii) *erete muŋusɔ-ki rɔe-te subida ara-asi-an*
 thus people-PL stay-INF comfort NEG-be-PST
 'That's why it wasn't possible for humans to live in the world.'
- (viii) *misiŋ-a ara-kia muinɬɔ kɔnɔn jɔnɔmɔ-ian*
 one.day-GEN she/he-DL one child born-PST
 'They had a child.'

- (ix) *erebeḷa dɔrɔm kɔilɔŋ tankuŋiŋ ɖak-te kɔb-sɔrɔ-jɔ*
 that time Daram God all water-ACC dry-PRF-PST-PRTCPL
muŋusɔ-ki-te unɔn-te jage-an
 people-PL-ACC keep-INF want-PST
 ‘At this time Daram God wanted to dry up all the water and keep human beings in the world.’
- (x) *erete rusi kɔnɔn-te gɔe-kijinja-bɔ-ɖɔ subui-te jinṭi-te*
 so saint child-ACC kill-? house-LOC-DEF all-ACC blood-ACC
mɔḷɔŋ-a
 throw-PST
 ‘That’s why he (the Daram God) had to kill the saint’s child and spray his blood everywhere.’
- (xi) *ere-re dɔrɔmɔ kɔilɔŋ rusi-te jiy-ki*
 so Daram God saint-ACC ask-PRS
 ‘So the Daram God asked the saint.’
- (xii) *rusi sina mɔnge-ɔ, rusi-anji jena gam*
 saint could agree-PST saint-FEM NEG say
 ‘Saint agreed but his wife did not agree.’
- (xiii) *rusi kɔnɔn-te gege-te dɔrɔm kɔilɔŋ muŋṭɔ kiḷɔg kin-a*
 saint child-ACC REDPL:kill-INF Daram God one tiger send-PST
 ‘Daram God sent a tiger in order to kill the saint’s child.’
- (xiv) *rusi kɔnɔn kakag kɔkɔm-bɔ kiḷɔg-te ɔb-gɔj*
 Saint child bow arrow-INS tiger-ACC CAUS-die
 ‘The saint’s son killed the tiger with his bow and arrow.’
- (xv) *arɔ kakag kɔm parɔgɔm asi-an*
 he bow arrow perfect be-PST
 ‘He (the saint’s child) was perfect with bow and arrow.’
- (xvi) *dɔrɔm kɔilɔŋ rusi-te gam ɖaluŋ ɔn-jɔ kɔnɔn-mɔ-te*
 Daram God saint-ACC say bathe go-PRTCPL child-DEF-ACC
kakag gɔg-te mana mi-kib-e
 arrow take-INF prohibit FUT-do-FUT
 ‘Daram God told the saint that he should prohibit his child from taking an arrow when he goes to bathe.’
- (xvii) *ɔba-rɔ mane-jɔ kɔnɔn-ɖɔ kakag ma-gɔg-an*
 father-GEN obey-PRTCPL child-DEF arrow NEG-take-PST
 ‘Obeying his father’s words the child didn’t take arrow to the river.’
- (xviii) *kiḷɔg rusi kɔnɔn-te ɔb-gɔj*
 tiger saint child-ACC CAUS-die
 ‘The tiger killed the child.’
- (xix) *ar-a injab-ɖɔ subuaɖe kinṭi-ɔ-ki*
 he-GEN blood-DEF everywhere spray-PST-PL
 ‘His blood was spread all over.’

- (xx) *ere-tasun purtibi-ta dak kəsər-ə*
 that-after world-LOC water dry-PST
 ‘After that water dried up from the world.’
- (xxi) *ar-a ijij iti-r-bə-ta paaqə u-i-ian*
 he-GEN leg hand-DEF-LOC-ABL hills happen-PRF
 ‘From his legs and hands hills and mountains emerged.’
- (xxii) *bəkək-qa juntar-ta əlag sumusiŋ ian*
 child-DEF hair-ABL leaf flower happen:PST
 ‘From the child’s hair plants grew.’
- (xxiii) *eləmi-ta rusia-kia barəgə-ʔa-leka kənən kəncelan ian*
 this-ABL saint-DL twelve-DEF-each son daughter be:PST
 ‘The saint couple had 12 sons and 12 daughters after this.’
- (xxiv) *arə-ki juaŋ-ki-te jənəmə kib-ki*
 he-PL Juang-PL-ACC born do-PL
 ‘They gave birth to Juangs.’
- (xxv) *ere-tasun juaŋ jati sursji ian*
 this after Juang community create be:PST
 ‘This is how Juang community was created.’

Pinnow’s (1960) collection of texts has a different folk tale about the origin of Juangs. I could hear very few informants of older generation subscribing to this story. I read the story given in Pinnow out to them. The story I give below has been taken from Pinnow’s texts which is no different from what I heard from few villagers. However, I didn’t check the language of the story with anybody. According to this story, the God Shiva and his wife the Goddess Parvati gave birth to a saint couple who lived in the forest by eating fruits and roots of the forest. They were always in meditation. A few days after hair started growing from Rusi’s (male saint) tongue. After that whoever he cursed got cursed and died. This created fear among the villagers, and they evoked spirits from mountains, forests to help them. The spirits of mountains and forests appeared, and they fed the Rusi with all sorts of food to satisfy him. After this event, the hair started falling out of Rusi’s tongue. After the hair fell out, whenever he cursed it had no effect. No one got cursed as his evil magical power disappeared. After that, wherever he sat good things started to happen, such as water started flowing from there. Wherever he kept his feet water started to flow from there which is called ‘Rusi’s foot water’. Then eventually many children were born from the Rusi who got to be known as Juangs. Thus Juangs are the children of the saint or Rusi. This is the origin of Juangs according to another folk tale. Note the contrast that in the earlier story there was water everywhere which is why there was no human civilization and Justice God had to do something as nasty as killing a child to be able to dry up the water and create human civilization. By contrast, in this story, there was no water or presumably not enough for the human habitats; the problem was solved with the help of spirits, gods, and demons of mountains and forests who made the evil power of the Rusi disappear by which it became possible thereafter for the Rusi to sit in places where water started flowing. The earlier story talked about a phase in which human child sacrifice was probably in practice. The earlier story justifies that Justice God persuades the saint to sacrifice his child so

the child's blood and severed limbs can create nature and civilization. The story is unforgivably suggestive about child sacrifice which was a part of Hindu Shakti cult in the middle ages. The second story disqualifies one's ability to curse by terming it to be a bad thing and promotes the message that the disappearance of evil habits like cursing someone to die can create by contrast, instead of death, the magical power to create nature and water and settlements. These two stories signify the influence of the Hindu belief system of two phases in different ways. The second story too has some Oriya loans. Both stories reinforce that Juangs are children of 'Rusi' meaning 'saint' in Sanskrit and other Indo-Aryan languages. The stories may have been created after Juangs have come into contact with the Aryan culture somewhere in prehistory.

I provide that text below; the glosses are my own because some of the glosses given by Pinnow don't match the grammatical categories found out for those segments.

7.2 Text 2

juang-rɔ kahaŋi rusi rusirai-kia A history of the Juang: Rusi and Rusiani

- (i) *isuɔr parboti jonomo-en-kia*
Shiva Parvati born-PST-3DL
'Shiva and Parvati were born.'
- (ii) *au-rɔ-ta rusi rusi-aŋi jonomoen-kia*
they-GEN-ABL Rusi Rusi-FEM born:PST-3DL
'Rusi and Rusiani were born from them.'
- (iii) *arɔ-kia banɔg-eta-ju ere-tasun boŋɔ-a roen-kia*
she/he-DL two this-ABL-EMPH that-ABL forest-LOC stay-3DL
'Both of them lived in the forest.'
- (iv) *phɔ]ɔ mu]ɔ jim-ɔ-kia*
fruit root eat-PST-3DL
'They ate fruits and roots (in the forest for survival).'
- (v) *e-tasun rusi-ra elay-ɖa-te juan]a-ro hari-seran*
this after Rusi-GEN tongue-DEF-LOC hair-DEF came-PRF:PST
'In few days hair started hanging out from Rusi's tongue.'
- (vi) *eta hari-seran-]ɔ jie-te lejɖag-an-]ɔ arɔ gɔ]ɔ*
then come-PRF:PST-PRTCLPL who-ACC curse-PST-PRTCLPL he die:PST
'After (the hair) came out whoever he cursed he died.'
- (vii) *eta-ju pirho-pa]ho-koilay baro-rokomo u-jim-o*
then EMPH village forest/spirit/demon/god twelve CAUS-eat-PST
'When this happened the spirits of mountains and forests fed him all sorts of food.'
- (viii) *u-jim-o-]o eta-sun elay-ɖa oioŋ-o*
CAUS-eat-PST-PRTCLPL thus-after tongue-DEF fall.down-PST
'Having fed him thus the hair of his tongue fell out.'

- (ix) *ono-jo jemiti jie-te lej-ḍa-gan-ḵo etasun kei-te*
 go-PRTCPL when which-ACC which-GEN-??-PRTCPL then who-ACC
ma-lageana
 NEG-happen:PST
 ‘When the hair fell out whoever he cursed it didn’t happen to him.’
- (x) *etasun rusi-ra soto-ro budi-ana*
 then Rusi-GEN truth-DEF end-PST
 ‘Then the truth about Rusi (Rusi’s magical power) ended.’
- (xi) *etasun rusi jouti doko-seran-ḵo ḍio era mane*
 then Rusi wherever sit-PRF:PST-PRTCPL well this all
boraboro ḍag asi-ke
 again.and.again water be-PRS
 ‘After that wherever the Rusi sat there always came water.’
- (xii) *rusi-ra pado-ra-te ḍag hari-an*
 rusi-GEN feet-GEN-ACC water came-PST
 ‘Water came out of Rusi’s foot.’
- (xiii) *rusi-a pado-ra ḍag ere-te gam-ke-ki*
 Rusi-GEN foot-GEN water this-ACC say-PRS-PL
 ‘This is known as “Rusi’s foot water”.’
- (xiv) *hari-ana eta jonomo rusi-ra boũso-ro, juay-ki*
 came-PST then born Rusi-GEN family-GEN Juang-PL
jonomo-siri-ki
 descendants-PRF-PL
 ‘Then Rusi’s descendants were born who were known as Juangs.’
- (xv) *ereta manta jono-man-ki jota tangorota*
 from there place:ABL be.born-PST-PL which:ABL hill:DEF:ABL
jonomosiri rusi
 be.born:PRF Rusi
 ‘They (Juangs) were born in the place where Rusi was born.’
- (xvi) *ere-tasun juay-ki asi-ke-ki rusi-ra putro*
 thereafter Juang-PL be-PRS-PL Rusi-GEN son
 ‘Therefore the Juangs are called the sons of Rusi.’

NOTES

- * A generous grant given to me to be at Max Plank Institute of Evolutionary Anthropology at Leipzig, Germany has helped me to complete this long pending work in a relatively short time. My sincere thanks to Professor Bernard Comrie for making it possible for me to come here to be able to complete the work. Thanks also to the editor.
- 1 The word ‘Juang’ in Juang language means ‘man’. The alternative name Patua means wearer of leaf-dress. Juangs abandoned their traditional leaf dress when Captain J. Johnstone forced them to wear clothes in the nineteenth century.
 - 2 Matson and Pinnow report vowel lengthening in case of first singular subject though I noticed lengthening in all persons and numbers.

- 3 /-ka/ is a genitive marker in Hindi. It may be possible that due to industrialization in the area Juangs are exposed to Hindi-speaking people as a result of which -ka may have entered the language, but the usage of /-ka/ seems to older than industrialization in the area.
- 4 I will give my own examples of object agreements in the large grammar book on Juang that I am writing for Europa Lincom.
- 5 Non-finite verb forms in Juang are marked by -te, -ta or they may just be the reduplicated forms of the roots.
- 6 Those two people grew up to be saints and they lived on a hill called Gonasika. Gonasika is still a hill in Keonjhar district which is very picturesque. Most of the speakers on Gonasika hill are monolingual Juangs. There are some other stories about Gonasika also which relate to portions in the great epic Ramayana.

BIBLIOGRAPHY

- Anderson, G.D.S. (2001) 'A New Classification of Munda: Evidence from Comparative Verb Morphology,' *Indian Linguistics*, 62: 27–42.
- Bhattacharya, S. (1975) *Studies in Comparative Munda Linguistics*. Simla: Indian Institute of Advanced Study.
- Dalton, E.T. (1872) *Descriptive Ethnology of Bengal*, Calcutta: Office of Government Superintendent Printing.
- Dasgupta, D. (1978) *Linguistic Studies in Juang, Kharia-Thar, Lodha, Mal-Paharia, Ghatolai and Paharia*, Calcutta: ASI.
- Elwin, V. (1948) 'Notes on Juang', *Man in India*, XXVIII (1–2): 1–146.
- Grierson, G.A. (1967) *Linguistic Survey of India, Vol. IV*, Delhi: Motilal Banarasidas.
- Lyall, A.J. (1868) *Report on the Etymological Committees on Papers Laid before them, and upon Examination of Specimens of the Aboriginal Tribes brought to the Jabbalpore Exhibition of 1866–67*, Nagpur.
- Mahapatra, B.P. (1962a) 'Revised Munda lexical list (Juang)', Unpublished manuscript.
- (1962b) 'A note on Juang phonology', Unpublished manuscript.
- (1962c) 'A note on Juang morphology', Unpublished manuscript.
- (1976) 'Comparative notes on Juang and Kharia finite verbs', in P.N. Jenner, L.C. Thompson and S. Starosta (eds) *Austroasiatic Studies. Part I & II* (Oceanic Linguistics Special Publication no. 13), Honolulu: University of Hawaii Press, 801–814.
- (1979) 'Munda languages', *Bulletin of the Department of Comparative Philology and Linguistics*, 4: 22–42.
- Matson, D.M. (1962) 'A brief description of Juang', Unpublished manuscript.
- (1964) *A Grammatical Sketch of Juang*, PhD Dissertation, Madison, University of Wisconsin.
- Misra, U. (1978) *Banakua*, (The Primary Magazine on Tribal Life) Bhubaneswar: Language and Culture of Orissa.
- Mohapatra, Sashmita (1991) *Juanga*, Bhubaneswar: Tribal Affairs Office.
- Patnaik, M. (1996) 'Internal DP structure of Juang', Paper presented in DLA conference, Chennai, India.
- (1998) 'Quotative agreement in Juang', Paper presented in International Conference on Agreement, Delhi University, Delhi.
- (2000) *Aspects of Juang Syntax*, Unpublished PhD dissertation, Delhi University, Delhi.
- (In Preparation) *Juang Grammar*, to be published by Lincom Europa.

- Pinnow, H.J. (1960) *Beiträge zur Kenntnis der Juang Sprache*, Unpublished Monograph.
- (1963) 'The position of the Munda language within the Austro-Asiatic family', in H.L. Shorto (ed.) *Linguistic Comparison in South-East Asia and the Pacific*, London. SOAS, 140–152.
- Ray, A.K. (1958) 'A note of Juang anthropometry', *Man in India*, XXXVIII: 36–39.
- Samuels, E.L. (1856) 'Note on forest race called Puttooas or Juanga inhabiting certain of the Tributary Mehols of Cutack', *Journal of the Asiatic Society of Bengal*, XXV, 295–303.
- Stampe, D.L. (1965) 'Recent work in Munda linguistics', *International Journal of American Linguistics* 31(4): 332–341; 32(1): 74–80, 32(2): 164–175, 32(4): 390–397.
- (1966) 'Phulbadi Juang phonology', Unpublished Paper.
- Zide, A.R.K. (1978) 'A Note on glottalization release in Munda', *Indian Linguistics*, 39: 70–75.
- Zide, N.H.(ed.) (1966) *Studies in Comparative Austro-Asiatic Linguistic*, The Hague: Mouton.
- Zide, N.H. (1969) 'Munda and non-Munda Austroasiatic languages', in Thomas Sebeok (ed.) *Current Trends in Linguistics*, The Hague: Mouton, 5: 411–430.
- Zide, N.H. and Stampe, D.L. (1968) 'The position of Kharia-Juang in the Munda family', in B. Krishnamurti (ed.) *Studies in Indian Linguistics* (Professor M.B. Emeneau's *Śaṣṭipurti* volume), Poona: Centres of Advanced Study in Linguistics, Deccan College, 370–377.

REMO (BONDA)

Gregory D.S. Anderson and K. David Harrison

1 INTRODUCTION

Remo, also known as Bonda, Bondo, and Bondo/Nanga Poroja (in Remo *remosam*), has several thousand speakers. Remo is known mostly from a few sources, Fernandez (1968), some of which is published as (1983) from the Hill Remo (Munḍlipaḍa) dialect, Ramachandra Rao's phonological materials, Swain's study and Bhattacharya's (1968) brief text and more extensive lexical materials, which are mostly of the Plains Remo variety. In the following discussion, Remo is used to mean either Hill Remo alone or the Hill and Plains dialects together. The data presented in this chapter are drawn from Fernandez (1968), Bhattacharya (1968) and from field work carried out by the authors in 2005 and 2007, which included digital video and audio of Remo speakers (attributed as [SDM] in the chapter, representing our primary consultant Sukra Dangada-Majhi).

Remo is unwritten. Its only close relative in South Munda is Gutob, with which it forms the Gutob–Remo branch. This latter may form a larger unit jointly with Gta? in a Gutob–Remo–Gta? node as has been suggested by Zide (1965) or may have several shared innovations that included Gta?, but not form an actual genetic unit as suggested by Anderson (2001). Remo and Gutob also share certain structural/typological (genetic?) features with Kharia as well; this enigmatic situation remains an open research question for the future.

The Remo-speaking Bonda occupy the Jeypore Hills in southern Orissa, west of the Machkund river in Malkangiri district, centred around Mundlipada (Parkin 1991:32). The Plains Remo are found primarily in 35 villages in Khairpat block of Malkangiri district. While the total number of Plains Remo grew from 2565 to 4764 from 1941 to 1971, the total number of speakers of the language did not increase accordingly. Gradual attrition to Desia has been occurring in some parts of the Remo (Bonda) area, and Remo is likely endangered.

Hill Bonda (Remo) are one of the tribal groups in India that are best known for being different than 'civilized' Indian citizens, and have a reputation for ferocity of character (Elwin 1950). They are organized into exogamous patrilineal clans typically named for villages, and, at a superordinate level, into two moieties, tiger (*ontal*), and cobra (*killo* or *kukusa*) (Elwin 1950:28–29). Boys and girls reside in segregated village dormitories, marriage is by mutual consent, not arranged, and it was not uncommon for older women to marry young boys (von Furer-Haimendorff 1943). The women traditionally shaved their heads and wore only a cloth girdle and elaborate ornaments, enormous metal necklaces and earrings and bead strands that came down to the upper thigh. The men wore loincloths, carried bows and arrows and reportedly would shoot each other (and outsiders) without qualm. Material culture still practiced includes fibre extraction and weaving, construction and use of ploughs, hoes, hatchets, mortars, pestles, nets, fish traps, baskets, bird snares, looms,

stringed instruments, and drums. They erect and venerate stone megaliths with offerings of rice, mangoes, and sacrificed fowl (Elwin 1950). The Bondo remains a popular attraction in the Indian ethno-tourist industry, at a weekly market they attend when they come down from their hill villages.

2 PHONOLOGY

2.1 Vowel inventory

The vowel inventory of Remo is relatively straightforward within the South Munda context. It is a typical five-vowel system. It may have schwa phonetically, but this is weakly motivated as a phoneme by Fernandez (1968).

(1) Remo Vowels

i	u
e (ə)	o/ɔ
a	

(Fernandez 1968:7)

Diphthongs of various types are also attested: (Fernandez 1968:66) *lean* ‘tongue’, *bois* ‘age’, *guiḍaq* ‘to wash’, *mbainu* ‘2 people’, *otoi* ‘not to be’, *kencua* ‘earthworm’, *dau* ‘small’.

In Plains Remo, various laxing and weakening processes typical of different vowels are seen in a range of contexts (Fernandez 1968:40–42).

- | | |
|--------------------|--|
| (2) /i/ > [i] /__# | <i>miri</i> ‘why’ |
| /e/ > [ɛ] /__n/r/l | <i>kenda</i> ‘branch’, <i>sellari</i> ‘scolopendra’ |
| /o/ > [ɔ] /__N/T | <i>dɔb</i> ‘white’, <i>aʔdɔŋ</i> ‘fish’ |
| /a/ > [ə] /__Ca | <i>əmas</i> ‘new moon day’, <i>kəsa</i> ‘astringent’ |

Some Plains Remo forms showing vowel oppositions include the following: *biri* ‘field on hill’ vs. *bire* ‘stone’; *kesu* ‘bed sheet’ vs. *kasu* ‘coin’ vs. *kasa* ‘astringent’ vs. *kosa* ‘joint part of metal point and the wooden part of an arrow’; *bund* ‘tree trunk’ vs. *bond* ‘tank’.

2.2 Suprasegmental phenomena

Remo makes limited use of phonemic nasalized vowels. Thus, one finds oppositions like the following in Hill Remo (Fernandez 1968:14).

- | | |
|--|--------------------------------|
| (3) <i>nkwi</i> ‘father-in-law’ | <i>nkwi</i> ‘younger sister’ |
| <i>ḍonkwī</i> ‘husband’s brother’s wife’ | <i>ḍionkwi</i> ‘married woman’ |
| <i>busā</i> ‘grandfather’ | <i>usa</i> ‘skin’ |
| <i>ŋger-sū</i> ‘boy’s dormitory’ | <i>nsu</i> ‘knife’ |
| <i>batuʔgō</i> ‘twins’ | <i>tumuŋgo</i> ‘night’ |

Note that contrastive nasalization is also seen in diphthongs in Hill Remo (Fernandez 1968:14).

- | | |
|---------------------------|---------------------------|
| (4) <i>orōȳ</i> ‘sickle’ | <i>suyroy</i> ‘bracelet’ |
| <i>oʔōndāȳ</i> ‘baby’ | <i>nsonḍay</i> ‘squirrel’ |
| <i>ḍumāw</i> ‘little’ | |

There are at least six words of Hill Remo where a nasalized vowel alternates with an oral vowel followed by a nasal consonant (Fernandez 1968:13).

- | | | |
|-----|----------------------------------|-------------------|
| (5) | <i>āyun</i> ~ <i>ayun</i> | ‘common fly’ |
| | <i>kīyar</i> ~ <i>kīnar</i> | ‘wife’s sister’ |
| | <i>nōḍa?</i> ~ <i>nōḍa?</i> | ‘honey’ |
| | <i>sōk</i> ~ <i>sonk</i> | ‘throat’ |
| | <i>orōȳ</i> ~ <i>oroyn</i> | ‘sickle’ |
| | <i>ḡger-sū</i> ~ <i>ḡger-suy</i> | ‘boy’s dormitory’ |

Note that while diphthongs are not overly common, *ey* is generally noted in only two words in Hill Remo and its nasalized counterpart *ēȳ* in one: *tēȳno* ‘anthill’, *geyn-o?* ‘chewed it’, *leympuy* ‘waist’ (Fernandez 1968:13)

The sequence *āḍ* is found in a small number of stems, for example, *bāḍsi* ‘flute’, *seserāḍ* ‘newborn baby’, but **ao* only occurs across morpheme boundaries, for example *a-om* ‘do not keep in hand’. Also, **oa* is always broken up by a *-w-* phonetically, for example, *ow-al-o?-pe* ‘you caused to thatch’ *ow-am-o?-niḡ* ‘I caused to pick’.

Some Plains Remo pairs with nasal and oral vowels include the following:

- | | | |
|-----|-------------------------------------|---------------------------------------|
| (6) | <i>ḡkuy</i> ‘younger sister’ | <i>ḡkūy</i> ‘wife’s father’ |
| | <i>er</i> ‘winnow’ | <i>ēr</i> ‘endosperm of a mango seed’ |
| | <i>ḍiyōḍaya</i> ‘house, etc.’ | <i>ḍiyōḍayā</i> ‘serpentine’ |
| | <i>usa</i> ‘skin’ | <i>busā</i> ‘paternal grandfather’ |
| | <i>kuy</i> ‘poke with folded hands’ | <i>kūy</i> ‘scold’ |

2.3 Consonant inventory

The consonant system of Hill Remo is typical South Munda, it has a five-way place-contrast at nasal and voiced stops. Glottal stop is found frequently.

- (7) Hill Remo
- | | | | | |
|----------|----------|----------|----------|----------|
| <i>p</i> | <i>t</i> | <i>ṭ</i> | <i>k</i> | <i>ʔ</i> |
| <i>b</i> | <i>d</i> | <i>ḍ</i> | <i>g</i> | |
| <i>m</i> | <i>n</i> | <i>ṇ</i> | <i>ɲ</i> | |
| | <i>s</i> | | | |
| | <i>z</i> | | | |
| | <i>l</i> | | | |
| | <i>r</i> | | | |
| <i>w</i> | | <i>y</i> | | |
- (Fernandez 1968:7)

Plains Remo presents a slightly different picture. Here, a voiceless palatal is found, but no retroflex *ṇ* or *z*; Hill Remo *w* is realized as Plains Remo *v*.

- (8) Plains Remo
- | | | | | |
|----------|----------|----------|----------|----------|
| <i>p</i> | <i>t</i> | <i>ṭ</i> | <i>k</i> | <i>ʔ</i> |
| <i>b</i> | <i>d</i> | <i>ḍ</i> | <i>g</i> | |
| | <i>s</i> | | <i>c</i> | |
| | | | <i>j</i> | |
| <i>m</i> | <i>n</i> | <i>ɲ</i> | <i>ŋ</i> | |
| | <i>l</i> | | | |
| | <i>r</i> | | | |
| <i>v</i> | | <i>y</i> | | |
- (Ramachandra Rao 1981:10)

In a small number of stems, Hill Remo (HR) has lost a final consonant that Plains Remo (PR) preserves:

(9)	Hill Remo	Plains Remo	
	<i>suʔu</i>	<i>suʔug</i>	'blossom'
	<i>wa</i>	<i>wal</i>	'swim'

At one point in the history of Remo there were, and for speakers of certain varieties there remain, instances of word-initial syllabic nasals. This is realized in the form of a lexicalized prefix or 'pre-base' in a number of Remo words. Remo speakers vary between treating these as syllabic or as pre-nasalization on the following consonant. Thus, a word like Hill Remo *nsu* 'knife' or Plains Remo *nsaj* 'refuse' may be treated as either one or two syllables for different speakers.

In Plains Remo, aspirates in loans are often deaspirated, for example, *g^hoɖa* ~ *goɖa* 'horse' *t^hal* ~ *tal* 'leafy fan' (Ramachandra Rao 1981:28–29). Both voiced and voiceless stops can occur word-finally *bop* 'head' *kub* 'many' *kot* 'manure' *bond* 'tank' *soŋɖ* 'castrated bull'.

Note that the palatal affricate is realized as alveolar before round vowels: *tsoŋɖi* 'beak' *tsolɕa* 'wick' *cali* 'skin' *kanc* 'bottle' *cinta* 'worry' *dzudzu* 'sigh' *dzoɖi* 'field canal' *dzokto* 'under' *jamali* 'check' *pajraɖa* 'spring season' (Ramachandra Rao 1981:35).

Some examples of contrastive words demonstrating the phonemic inventory of Plains Remo are offered below:

(10)	<i>pali</i> 'river bank'	<i>bali</i> 'sand'	
	<i>tati</i> 'leafy plate'	<i>dadi</i> 'maternal grandfather'	
	<i>taku</i> 'mango seed'	<i>ɖaku</i> 'girl'	
	<i>parak</i> 'river'	<i>parag</i> 'spring pond'	
	<i>kubu</i> 'node'	<i>gubu</i> 'pig'	
	<i>ruŋɔ̃</i> 'chill'	<i>suŋɔ̃</i> 'fire'	
	<i>ranɖi</i> 'widower'	<i>lanɖi</i> 'a month [May–June]'	
	<i>may</i> 'pull'	<i>nay</i> 'cobra'	
	<i>nom</i> 'eggplant'	<i>ŋ[g]om</i> 'village'	
	<i>rot</i> 'charriot'	<i>roɭ</i> 'fat'	
	<i>maʔ</i> 'curry'	<i>mag</i> 'a month'	<i>vaʔ</i> 'bow'
	<i>miri</i> 'how'	<i>piri</i> 'bird'	<i>biri</i> 'field on hill'

Note that word-finally the sequence *-oŋ#* may be realized as a nasalized vowel + glottal combination, *ɔʔ#* in Remo.

- (11) *gutumɔ̃ʔ* 'forehead' = *gutumɔŋ* [SDM]

2.4 Syllable structure and phonotactics

The following comments can be made about the phonotactics of Plains Remo: Most consonants can be found word-initial except *ʔ-, while final *-l* and *-ɭ* are found only in loans, for example, *jal* 'fish net'. In addition *-ŋ*, *-v*, and *-c* are all lacking word-finally (Ramachandra Rao 1981:19). In terms of the phonotactics of vowels, these are generally less restricted except that nasalized *-ĩ* is only found finally, for example, *sĩ* 'sun', and nasalized *ẽ* is never found initially: *mẽ* 'younger brother' (Ramachandra Rao 1981:23). In contrast to Rao's analysis, the consonant inventory of Plains Remo as given in Bhattacharya (1968) reports *v* as a

phoneme, but gives no examples in his lexicon. He finds *w* only in a few Desiya loanwords (e.g. *terwa* 'eunuch', *parwa* 'pigeon') and likewise *c* and *h* only in loanwords.

The following medial heterosyllabic clusters are found in the Plains Remo given in Bhattacharya (1968).

(12)

	<i>p</i>	<i>b</i>	<i>m</i>	<i>t</i>	<i>l</i>	<i>d</i>	<i>ɖ</i>	<i>n</i>	<i>ɳ</i>	<i>ɲ</i>	<i>y</i>	<i>j</i>	<i>s</i>	<i>r</i>	<i>ʀ</i>	<i>l</i>	<i>k</i>	<i>g</i>	<i>ɣ</i>	<i>ʔ</i>	
<i>p</i>														(✓)			(✓)				
<i>b</i>							✓						✓		✓		(✓)			✓	
<i>m</i>	✓	✓	✓	✓	✓	✓								(✓)							
<i>t</i>				✓										(✓)				✓	✓		
<i>l</i>	✓	(✓)		✓														✓			
<i>d</i>						(✓)				(✓)				✓							
<i>ɖ</i>																					
<i>n</i>				✓			✓	✓		✓	(✓)						✓	(✓)			
<i>ɳ</i>											✓									✓*	
<i>ɲ</i>																					
<i>y</i>	✓*										✓									(✓)	
<i>j</i>														(✓)							
<i>s</i>				✓			✓							(✓)				✓			
<i>r</i>	✓	✓	✓	✓			✓							✓				✓	✓		
<i>ʀ</i>																					
<i>l</i>		✓																			
<i>k</i>				✓										✓	✓	(✓)	✓				
<i>g</i>		✓				✓	✓	✓						✓	✓	✓	✓				
<i>ɣ</i>	✓	✓		✓								✓	✓	✓	✓	✓	✓			✓	
<i>ʔ</i>	✓	✓	✓	✓			✓				✓		✓	✓	✓	✓	✓			✓	

Key: (✓) attested only in loanwords, ✓* attested only in proper nouns, ✓ geminate.

The following lexemes (simplex or complex) exemplify consonant clusters. (Note: In Bhattacharya's transcription, a final consonant followed by an apostrophe, for example, *k'*, *p'*, indicates a pre-glottalized, unreleased stop, for example, ²p', ³k'. We have retained this convention in the wordlist presented here.)

(13)	<i>p</i>	<i>supret</i>	superintendent of police
		<i>upkar</i>	aid (Desiya loanword)
	<i>b</i>	<i>bóbsi?</i>	headache
		<i>bóbɣa</i>	a toothless person
		<i>bóbgari</i>	front side
		<i>bóbker</i>	head (i.e. ear) of paddy
		<i>tiribɔak'</i>	cloud
		<i>dabla</i>	the white one (Desiya loan)
	<i>m</i>	<i>limbi</i>	catapult
		<i>sɔmdi</i>	son's wife's father
		<i>sɔmpu?</i>	muskrat
		<i>rimkur</i>	quarrelsome person
		<i>kammi</i>	deity, spirit
		<i>ɖumri</i>	fig (Desiya loanword)
		<i>ɖɔmɖi ɖiɔ</i>	inner part of house
		<i>umtari</i>	prostrate
	<i>d</i>	<i>buddi</i>	intelligence (Desiya loanword)
		<i>dadya</i>	bad (Desiya loanword)
		<i>badrai</i>	murderer
	<i>ŋ</i>	<i>sangɔ-</i>	to stop
		<i>sɔŋra</i>	a kind of rat
		<i>suŋbɔr-</i>	to be sold out
		<i>buŋte</i>	buffalo (male)
		<i>raŋlak'</i>	glutton
		<i>juŋrak'</i> -	to sink
		<i>juŋjuŋlu</i>	earring (upper) for male
		<i>baŋsa</i>	until, henceforth
		<i>ɖeruŋpe?</i>	funnel
	<i>g</i>	<i>laglak'</i>	cave
		<i>lɔglɔk'</i> -	to calve
		<i>snagsagrak'</i>	(Hill Remo) window
		<i>sinigbuŋ</i>	a worm which emits bad smell
		<i>mugsɪŋ</i>	smoke
		<i>bunúgɖin</i>	a drum
		<i>ɖágɣak'</i> -	to be wet
		<i>bitig-na</i>	salt-LOC
		<i>digdag</i>	flash of lightning
		<i>ɔɖugɖuk'</i> -	to send
	<i>t</i>	<i>sitgur</i>	restlessness
		<i>mɔtri</i>	minister (Desiya loanword)
		<i>titti ɖenta</i>	to bury a dead body
	<i>ʈ</i>	<i>maʈpɔ?</i>	a wrapper, a blanket

	<i>puḥbəl</i>	football (English loanword)
	<i>əkəʃkəʃte-</i>	to have something tightened
<i>l</i>	<i>siŋal-bai</i>	the head of a horned deer
<i>s</i>	<i>lɔsna</i>	also
	<i>ləktapʰ</i>	dew
	<i>siksəŋ</i>	bone, skeleton
	<i>mesreʃ</i>	magistrate (English loanword)
	<i>miste dʒi:n</i>	neighbour
	<i>uskəpʰ</i>	hiccup
<i>n</i>	<i>ontal</i>	cobra moiety
	<i>pətənlakʰ</i>	a kind of necklace
	<i>punni</i>	lotus
	<i>punɖay-</i>	to stop another
	<i>banji ~ banjala</i>	a childless couple
	<i>nɖe</i>	bring (second person imperative)
	<i>minɖa</i>	one alone
	<i>dinke</i>	daily, always (Desiya loanword)
	<i>dinsu</i>	article (Desiya loanword)
<i>ŋ</i>	<i>məŋgam-sam</i>	a festival held in the month of Pus
	<i>maŋjal</i>	very much
<i>y</i>	<i>buyʃiŋ-</i>	to play on a flute
	<i>pusaypa</i>	pitmāri tree*
	<i>bayg</i>	luck (Desiya loanword)
	<i>əɔɔjakʰ</i>	to make ready
<i>j</i>	<i>bejri</i>	tomato (Desiya loanword)
<i>r</i>	<i>sərlāi</i>	cucumber
	<i>sərkupʰ</i>	snail
	<i>səɾni</i>	a female thief
	<i>səɾmɔʔ-</i>	to awake from sleep
	<i>urkuri</i>	mad
	<i>orpa</i>	the last of the death rites
	<i>məɾtəkʰ</i>	star
	<i>məɾsiŋ</i>	chilli
	<i>burbuŋ</i>	a kind of wild fruit
<i>k</i>	<i>kakka</i>	father's younger brother
	<i>makɾa</i>	spider (loanword)
	<i>ɖokra</i>	man, male
	<i>əkser</i>	horizontal pole on which clothes are dried
<i>ʔ</i>	<i>reʔgi</i>	wooden necklace
	<i>saʔye</i>	ancestor spirit
	<i>sanɔʔsiŋ</i>	egg shell
	<i>liʔtapʰ</i>	a small basket
	<i>seʔma</i>	bug
	<i>meʔɖer</i>	a kind of bean
	<i>meʔɾer</i>	a kind of worm

<i>baʔle</i>	relatives on the father's side
<i>bɔʔre-sa</i>	swearing by a brother
<i>naʔpɔʔ</i>	cloth worn by Bonda women
<i>gusɔʔbur</i>	wild dog

In Plains Remo, the combination of /gg/ is realized as [ʔg]. *og-oʔ-t-ij* 'I have carried a child' vs. *oʔ-goy-t-ij* 'I can carry a child.'

- (14) *nsuk* 'knife' > *nsug-e* 'knives'
 > *nsug-beroy* 'with the knife'
 > *nsug-e-beroy* 'with the knives' (Ramachandra Rao 1981:83)

Regarding possible syllable types, the following are permitted (Ramachandra Rao 1981:24).

- (15) V
 VC
 CV
 CVC, CVVC
 CVCC (only in loans)
 CCV, CCCV

One important distributional fact about Remo consonantism is that initial CC clusters are virtually exclusively homo-organic (syllabic) nasal + obstruent clusters. There are a few loans that are an exception to this: *ḍruka* 'leopard/panther' and *gyapo* 'bastard' (Ramachandra Rao 1981:48–49, 63–64). As aforementioned, word-initially these are (often) not clusters or unit phonemes (i.e. pre-nasalized stops) but rather the initial nasal forms the nucleus of its own syllable.

- (16) *mp mpo* 'saree' *mpor* 'husband'
mb mbar '2'
ns nsap 'waist' *nsuga* 'banana'
nt ntopsij 'egg' *ntuʔ* 'near'
nd ndeŋɖa 'tree species'
ŋk ŋkuk 'rice'
ŋg ŋgo 'crab' *ŋgorgaŋ* 'black scorpion'
nj njur 'dawn'

As for possible three-term initial clusters #CCCV, only the following with initial, homo-organic nasals (Ramachandra Rao 1981:64) are permitted.

- (17) *ntra* 'morning liquid meal, gruel'
ntlo 'that day'
ŋḍrokaŋbiye 'big black ant'
ntni 'mushroom'

Three-term medial clusters VCCCV invariably involve a pre-nasalized cluster (Ramachandra Rao 1981:66). These vary between bi- and tri-syllabic pronunciations.

- (18) *gorŋɖi* 'rag' *suŋḍru* 'bamboo hat'
paŋɖka 'pigeon' *kondru* 'hump of bullock'

In final position, a similar observation can be made: two-term final clusters CC# are mostly homorganic nasal + obstruent clusters as well (Ramachandra Rao 1981:66).

- (19) *kond* ‘arm’ *sonɔ̃* ‘castrated bull’
bond ‘tank’

Possibly, Plains Remo has pre-nasalized stops appearing in coda position functioning as unitary elements phonologically. Were this suggestion taken to be true, then Plains Remo syllable structure might be reducible to the following:

- (20) V VC CV CVC CVVC CRVC

where C₁ is either a simplex consonant or a pre-nasalized stop or a pre-nasalized stop + liquid cluster. Syllabification phenomena support this hypothesis (Ramachandra Rao 1981:78).

- (21) *u.ŋɔ̃* ‘temple’ *su.ŋɔ̃ru* ‘bamboo hat’
n[.]d̩e.ŋɔ̃a ‘tree species’ *e.ŋke.te.ŋke* ‘crookedly’

Note that vowel + *n* + consonant (or pre-nasalized), that is, onC (or oⁿC) and ðC contrast word-finally in Hill Remo *ons* ‘duck, swan’ vs. *sðk* ‘throat’.

2.5 Intonation/stress

In Plains Remo stress is always on the second syllable. That is, it is in final position in disyllable words, otherwise second syllables get primary stress, and every subsequent even numbered syllable gets secondary stress. This suggests that the first syllable is marked as extra-metrical and stress is assigned by forming iambic (left-headed) binary feet, with left-headed prominent word structure:

- | | |
|--|---|
| <p>(22) <i>oíyoʔniŋ</i>
 <i>(o).í.yoʔ.niŋ</i>
 (Ø)S W S
 (*) (*
 *</p> | <p><i>bəbáɔ̃denkiŋ</i>
 <i>(bə).bə.ɔ̃den.ki.niŋ</i>
 (Ø)S W S W
 (*) (*)
 *</p> |
| <p><i>baɔ̃ʔsɔ̃suŋsɔ̃ɔ̃denkə̀ga</i>
 <i>(ba).ɔ̃ʔ.suŋ.sɔ̃ɔ̃den.kə̀.ga</i>
 (Ø)S W S W S W
 (*) (*) (*)
 *</p> | <p><i>gaysúŋsuŋɔ̃denkiŋ</i>
 <i>(gay).súŋ.suŋ.ɔ̃den.ki.niŋ</i>
 (Ø)S W S W S
 (*) (*) (*
 *</p> |

2.6 Morphophonology

Both dialects of Remo exhibit a range of morphologically triggered alternations. For example, in some stems ending in *-aj*, this monophthongizes and become *-e* before the non-past marker in Hill Remo (Fernandez 1968:16).

- (23) *bosaj-oʔ-niŋ* > *bose-t-iŋ* ‘I set aside (offer)’

With the stem *-iy/i* ‘return’, on the other hand, *iy* becomes *i* before consonant-initial affixes (Fernandez 1968:16).

- (24) *o-iy-oʔ-niŋ* ‘I caused to return’ *i-t-iŋ* ‘I return’ *i-g-niŋ* ‘I returned’

Assimilation to place with nasal + consonant combinations can occur across morpheme boundaries in Hill Remo (Fernandez 1968:17).

- (25) *taʔmi-baŋ-gə-ta* ‘he has sneezed’ < *taʔmi-bam*
dem ‘do’ > *ɖiŋ-ɖem* REDPL

With a small number of nasal final stems, in the first person singular past tense (conjugation-II), there is a coalescence between the final nasal, regardless of place and the velar consonant of the PAST.II suffix to yield *ŋ* (Fernandez 1968:17).

- (26) *bam* ‘get’ > *baŋ-niŋ* ‘I got’ < */bam-g-niŋ/*
toŋ-niŋ < *ton-g-niŋ* ‘I stood up’
sənsayŋ-niŋ < *ɲ-g-niŋ* ‘I was ahead’

In Hill Remo, stem-final *ɖ*, *s*, *l* > Ø before a consonant-initial suffix or word-boundary/hiatus. Note that in Plains Remo, *s* and *l* show the same behaviour while in this same environment, *ɖ* > ʔ.

- (27) Hill Remo
baɖ-oʔ-niŋ ‘I slapped’ *ba-t-iŋ* ‘I slap’ *baʔ* ‘slap!’
bal-oʔ-niŋ ‘I burned’ *ba-t-iŋ* ‘I burn’ *ba* ‘burn!’
sayks-oʔ-niŋ ‘I sifted’ *sayk-t-iŋ* ‘I sift’ *sayk* ‘sift!’ (Fernandez 1968:18)

- (28) Plains Remo
ruys-oŋ-niŋ ‘I milked (cow)’ *ruy-t-iŋ* ‘I milk’ *no-ruy* ‘milk!’
ju-l-oʔ-pa ‘you 2 saw’ *ju-t-pa* ‘you 2 see’ *no-ju* ‘see/look!’
beɖ-oʔ-naŋ ‘we 2 gave’ *beʔ-to-naŋ* ‘we 2 give’ *no-beʔ* ‘give!’
 (Ramachandra Rao 1981:101)

Note that the restriction on final clusters mentioned above in Hill Remo operates on the word-level (as seen in the form ‘sift!’ above), not on the stem-level, where clusters like *-ks* are permitted.

In Plains Remo, the combination of */gg/* is realized as [ʔg]. *og-oʔ-t-iŋ* ‘I have carried a child’ vs. *oʔ-goy-t-iŋ* ‘I can carry a child.’

In Plains Remo, stem-final voiceless consonants may be voiced when adding vowel- or voiced-consonant initial suffixes (Ramachandra Rao 1981:83).

- (29) *nsuk* ‘knife’ > *nsug-e* ‘knives’
 > *nsug-beroy* ‘with the knife’
 > *nsug-e-beroy* ‘with the knives’

Plains Remo shows a range of consonant assimilations:

- (30) */ŋ/* > [m] / ___b *ntəmbək* ~ *ntəŋbək* ‘brain’
/ll/ > [n] / V[+NASAL] ___ *nsɔ̃* > *nsɔ-le* ~ *nsɔ-ne* ‘spade’-PL
/ɖ/ > [r] / [r] ___ *buitur renta* ‘to spit’ (< *lɖental*)
 ʔ ~ g *nsuk* ‘big knife’ > PLURAL *nsuʔle* ~ *nsugle*
ɖ ~ ʔ ~ l *piriʔ* ‘bird’ > *piriʔre* ~ *piriʔde* ~ *piriʔle* ‘bird’-PL

According to data in Fernandez (1968:18), with a small number of stems ending in *-b*, there is a curious change in first person singular past forms. Both of the stems below belong to conjugation-II. Rather than **lob-g(i)-niŋ* and **lob-ga* as might be expected, one finds *lok-niŋ* ‘I cured’ *lo-ga* ‘he cured’ instead. Similarly, one finds

sab > *sak-niŋ* ‘I came’ and *sa-ga* ‘he came’. Note that non-past shows regular voicing assimilation, for example, *lop-t-iŋ* ‘I cure’.

The causative prefix *o-* is realized as *ow-* before stems beginning with *-a*. This is due to the restriction on surface sequences of **oa* in Remo: *ow-al-oʔ-pe* ‘you caused to thatch’ *ow-am-oʔ-niŋ* ‘I caused to pick’.

As discussed in the section below on verb morphology, reduplication plays a significant role in Remo; for example, the reduplicated verb stem allomorph must be used with the suffixes *-qen*, *-goŋ*, and *-qusuʔ*. Optimally, the reduplication takes the shape of CVC-. Certain vocalic changes are typically found in Remo reduplicated forms. With high and mid vowels, the reduplicant vowel is always high. With rounded syllable nuclei, the reduplicant has a rounded vowel [u], with unrounded base nuclei, the reduplicant vowel is unrounded [i]; otherwise, the vowel is ə. This can be summarized as follows:

- (31) *Ci* / *Ce* / *Cey* > *Ci-*
Coy / *Cwi* / *Cwo* > *Cu-*
Ca(y) > *Cə*

As alluded to above, not every stem copies the second consonant in the reduplicated form. Generally nasals are preserved, but *-b*, *-d*, *-l*, and *-s* are deleted, for example, *qab* ‘pound’ > *qə-qap-*. CVr stems also mainly lose the final *-r*, that is, they are realized as CV- reduplication except *tur-tur* ‘search for’ (Fernandez 1968:20).

A handful of stems show idiosyncratic behaviour in reduplicated forms. For example, stem-final consonants in targets as well as copies may be lost, unexpected devoicing of stem final consonants in the base and the reduplicant may occur, or the reduplicant appears with unexpected vocalism, etc. (Fernandez 1968:20–21).

- (32) *leq* > *lə-le* ‘squeeze’
lag / *tag* > *lak-lak* / *tak-tak*, ‘smoothe’ / ‘strip’
log / *tog* / *zog* > *lok-lok* / *tok-tok* / *zok-zok* ‘fall’ / ‘pick up’ / ‘trample’
bam > *bum-bam* ~ *bəm-bam*

As in Hill Remo, the reduplicated stem in Plains Remo is used with mono-syllabic verbs with the following suffixes: PROG *-qen* CAP *-goŋ* and DESID *-qusuʔ*. All result from fused auxiliary verb constructions (see section 3.2.12). As with Hill Remo, there are a number of lexically restricted idiosyncrasies in the reduplicated forms themselves. However, the systems also differ not insignificantly.

In Plains Remo, the only time one finds full CVC- reduplication is if in the stem, $C_2 = -g/-n$. Otherwise, there is CV- reduplication. Note that the reduplicated vowel is a faithful copy of stem V_1 . The vowel in the reduplicant is not restricted to high vowels and schwa as in Hill Remo.

- (33) *ki-kib* ‘pour’ *lu-luq* ‘borrow’
ga-gar ‘rip’ *jo-jor* ‘descend’
qe-qem ‘do’ *ru-ruis* ‘milk’
sug-sug ‘sweep’ *qon-qon* ‘carry on shoulders’

Note that with stems of the shape VC, reduplication copies only the vowel, not the coda, thus yielding a long vowel. This is true even if the coda consonant is *-n* or *-g*, for example, *u-un* ‘transplant’, *o-og* ‘carry child in sling’.

The PAST.II morpheme can have a zero-allomorph when used with a first or second person subject, and is realized as *-ga* when in word-final position with a third person subject.

The plural suffix *-le* may be realized as *-ne* particularly following stems that have a nasal consonant in the final syllable.

- (34) (a) *remo-ne gulay-lə remo-ne banṭi-tunṭi*
 person-PL! all-PL person-PL crookedly-ECHO
uriṅ[w]-o[ʔ]-ta
 walk-CV-[AUX]-NPST.II
 ‘All people are walking crookedly.’ [SDM]
- (b) *remo-le banṭi-tunṭi uriṅ-o-ta*
 person-PL crookedly-ECHO walk-CV-[AUX]-NPST.II
 ‘The people are walking crookedly.’ [SDM]

3 MORPHOLOGY

3.1 Nominal morphology

The system of nominal morphology in Remo is relatively straightforward. Both inflectional prefixes and suffixes are found, but only a relatively small total number. The derivational system appears to have been much richer at one point, with a wide range of lexicalized derivational elements attested.

3.1.1 Number

The plural morpheme is *-(l)e*, with numerous allomorphs. Its use is not obligatory but it appears with both animates and inanimates: *remo-le* ‘men’ (< *remo*), *kutom-e* ‘mallets’, *bob-e* ‘heads’, *tamaʔkuli-le* ‘fireflies’.

- (35) (a) *sayb-le joman-jimin ḍay-ga-aluṅ*
 sahib-PL policeman-ECHO climb-PST.II-SUB
 ‘The sahib policemen climbed beneath.’ (Fernandez 1968:98)
- (b) *gitin biri-bɔ saʔme su-sum-sa*
 that.CLOSE forest-LOC mandeya.CORN REDPL-eat-PURP
kubete piriʔ kukum gisaʔk sa-sap tɔ
 many bird peacock monkey REDPL.COME NARR.PRTCL
 ‘Many birds, peacocks, and monkeys used to come to that forest to eat the corn.’ (Bhattacharya 1968:147)
- (c) *gisaʔg-e semug bagbɔ ḍai-ga*
 monkey-PL tree on.top.of climb-PST.II.3
 ‘The monkeys climbed on the top of a tree.’ (Bhattacharya 1968:147)

Note that stem-final *-ʔ* > *g* and *ʔk* > *ʔg* before the plural morpheme, for example, *moʔ* ‘eye’ > *mog-e* ‘eyes’, *gisak* [gisaʔk] ‘monkey’ > *gisaʔg-e* ‘monkeys’.

- (36) *mɔmɔrtɔʔg-e tina-ga*
 star-PL be.seen-PST.II
 ‘Stars were visible.’ (Bhattacharya 1968:85)

The plural allomorph *-le* appears after final vowels (all data in this section from Bhattacharya 1968):

- (37) *ηeri-le* ‘bodies’ *jɔri-le* ‘streams’
gisi-le ‘lice’ *gine-le* ‘teeth’
gu-le ‘boys’ *ηgɔ-le* ‘crabs’

Some vowel final forms take the *-e* suffix, inserting a glide.

- (38) *tati-ye* ‘leaf-plates’

Some forms ending in consonants show optional or regionally varying deletion of *-l*.

- (39) *qeruŋ-e ~ -le* ‘horns’
gusuŋ-e ~ -le ‘men of the Dora community’
gisiŋ-e ~ -le ‘cocks’
girem-e ~ -le ‘cats’
kɔʔn-e ~ -le ‘these ones’
kɔʔ-e ~ -le ‘cots’
meʔqer-e ~ -le ‘kind of beans’
mbur-e ~ -le ‘crowbars’
bailɔʔg-e ~ -le ‘friends’ (< *bailɔʔk*)

Other consonant final forms are not reported as allowing *-l* deletion,

- (40) *barik-le* ‘village guard of the Dom caste’
dalait-le ‘office peon, chaprasi’
jɔil-le ‘jail’
kes-le ‘cock’s combs’
kūi-le ‘wells’
ig-le ‘excrements’ (< *iʔk*)

while a few are only listed as taking *-e* alone.

- (41) *ayer-e* ‘unripe mangoes’
gusɔʔbur-e ‘wild dogs’

Plains Remo also shows allomorphy *-le ~ -ʔe / ʔ*

- (42) *qeruŋpeʔ-ʔe ~ -le* ‘funnels’
guluʔ-ʔe ~ -le ‘hares’
gubuʔ-ʔe ~ -le ‘pigs’
bubuʔ-ʔe ~ -le ‘snakes’
nsaʔmɔʔ-le ~ -ʔe ‘eyelashes’
pɔʔmiʔ-le ~ -ʔe ‘leaf cups’
gieʔ-ʔe ~ -le ‘ropes’

Some apparent lexical exceptions are not reported to take *-ʔe*, despite a final glottal stop,

- (43) *tumɔʔ-le* ‘mouths’
nseʔmiʔ-le ‘noses’
suʔ-le ‘oils’

while a few forms appear to take only *-ʔe*.

- (44) *kurmeʔ-ɽe* ‘goat hooves’
bileʔ-ɽe ‘feathers’

Finally, some expanded allomorphy *-le ~ -ɽe ~ -de* may be found

- (45) *bieʔ-le ~ -ɽe ~ -de* ‘ants’
gusɔʔ-le ~ -ɽe ~ -de ‘dogs’

and a few irregular forms:

- (46) *kunnui-le ~ -se* ‘wives’
jaɲarɓɔi-se ‘Hill Bonda women’
tuna-le ~ -ne ‘younger sister’

Irregular plural formants or *pluralia tantum* may be found to a limited degree as well: *oʔon-ɖay* ‘son, daughter’ > *oʔon-ɖay-ɽe* ‘daughters, sons (someone else’s)’, or *sela-n-o-ɲger-o* ‘girls and boys’ (Fernandez 1968:65).

Note that a plural marked noun is still morphosyntactically singular in terms of verb agreement (see also section 3.2.1).

- (47) *bondagada-na remo-le uli sum-to*
 Bondagada-GEN person-PL mango eat-NPST.I
 ‘The people of Bondagada eat mango[es].’ [SDM]

3.1.2 Case

The most characteristic and typologically unusual feature of the Remo nominal inflectional system is its use of the objective case *a-*. It is nearly obligatory with pronouns and in a number of contexts with nouns as well (Fernandez 1968:66ff.). Note that the semantic role of the element varies considerably.

One area where the use of the objective *a-* is found is when two overt non-subject NPs are present (even if one is indexed pronominally). The *a-* appears on the element *not* fulfilling the patient/theme role in these constructions. It, thus, appears to function as a kind of oblique object element. In addition, it most commonly occurs on definite animate noun phrases so its function is partially deictic.

- (48) (a) *niɲ a-remo kiyaɲ bi-be-ɖen-t-iɲ*
 I OBJ-man rice REDPL-give-PROG-NPST-I
 ‘I am giving the man rice.’
 (b) *selane aɖoɲ a-leʔsiɲ poɖik-gə-ta*
 girl fish OBJ-basket put.in-PST.II-NPST.II
 ‘The girl has put fish in the basket.’ (Fernandez 1968:66)

Compare the above with the following:

- (c) *niɲ nsuɽaʔ sum-t-iɲ* (d) *niɲ nsuɽaʔ a-sum-t-iɲ*
 I banana eat-NPST-I I banana NEG-eat-NPST-I
 ‘I eat bananas.’ [SDM] ‘I am not eating a banana.’ [SDM]
 (e) *a-ɖio wi-ya* (f) *a-niɲ kiyaɲ be*
 OBJ-house GO-IMP OBJ-I rice give
 ‘go home!’ ‘give me rice’ (Fernandez 1968:112)

- (g) *niŋ a-gusoʔ kiyəŋ su-sum be-t-iŋ*
 I OBJ-dog rice REDPL-eat AUX-NPST-I
 ‘I am feeding the dog the rice.’ (Fernandez 1968:120)
- (h) *gitin remo a-monaʔbay selane kiyəŋ beq-oʔ*
 that.CLOSE man OBJ-fat girl rice give-NPST.I
 ‘That man gave rice to the fat girl.’ (Fernandez 1968:119)

If there is no non-patient present, *a-* may occur on the patient. Note that this applies to reflexive pronouns as well. All of these verbs could be considered special transitive subtypes in Remo, which subcategorize for a subject and an oblique, not a patient, and that triggers the use of *a-*.

- (49) (a) *niŋ a-remo kiyəŋ ŋruŋ o-mak-t-iŋ*
 I OBJ-man rice cook CAUS-learn-NPST-I
 ‘I teach the man to cook rice.’
- (b) *niŋ a-niŋ o-jul-oʔ-niŋ*
 I OBJ-I CAUS-see-PST.I-I
 ‘I saw myself.’ (made me see me) (Fernandez 1968:66)
- (c) *remo a-gisiŋ ju-to* (d) *niŋ a-niŋ ju-ti-ŋ*
 man OBJ-chicken see-NPST.I I OBJ-I see-NPST-I
 ‘The man sees the chicken.’ ‘I see myself.’
 (Fernandez 1968:124) (Fernandez 1968:125)
- (e) *niŋ tuwela a-mənqə remo jul-oʔ-niŋ*
 I yesterday OBJ-one man see-PST.I-I
 ‘I saw that one man yesterday.’
- (f) *niŋ a-remo tuwela jel-oʔ-oʔ-niŋ*
 I OBJ-man yesterday see-PST.I-PST.I-I
 ‘I saw the man yesterday.’ [SDM]
- (g) *remo a-niŋ jul-oʔ-ta*
 man OBJ-I see-PST.I-NPST.II
 ‘The man sees me.’ (lit. has seen)
- (h) *may a-may-ʔe tuŋsaŋ-to*
 she/he OBJ-she/he-PL derogate-NPST.I
 ‘He derogates them.’ (Fernandez 1968:67)
- (i) *a-remo jul-oʔ-no-ki*
 OBJ-man see-PST.I-2-Q
 ‘Did you see the man?’ (Fernandez 1968:67)

Experiencers are also marked with the objective *a-*, and occur with third person verb agreement.

- (50) (a) *a-niŋ lu-lor-qusoʔ-gə-ta*
 OBJ-I REDPL-VOMIT-DESID-PST.II-NPST.II
 ‘I wish to vomit.’
- (b) *a-oʔon jeri si-ta*
 OBJ-baby body hurt-NPST.II
 ‘The baby hurts.’ (Fernandez 1968:67)
- (c) *a-niŋ mabisom ruŋoʔ-gəta* (d) *a-niŋ kuru-gəta*
 OBJ-I very cold-PRF OBJ-I hungry-PRF
 ‘I feel very cold.’ ‘I am hungry.’

- (e) *a-may taʔmibaŋ-gəta*
 OBJ-she/he sneeze-PRF
 'he is sneezing.' (Fernandez 1968:112)

Structurally speaking, the *a-* objective is not a prefix, but rather a phrasal proclitic which targets the leftmost edge of the relevant NP that it serves to mark. Thus, it may appear on a demonstrative, a possessive pronoun, or anything likely to precede a noun or come initially in a Remo noun phrase.

- (51) (a) *a-kon bire kur* (b) *a-kon soka oyja*
 OBJ-that stone roll OBJ-that shirt how.much
 'roll down that stone' *dabu qi-ta*
 money COP-NPST.II
 'How much does that shirt cost?'
 (Fernandez 1968:67)
- (c) *niŋ a-niŋ-ŋa qio uriŋ-t-iŋ*
 I OBJ-I-GEN house walk-NPST-I
 'I will walk to my house.'
- (d) *a-nay-ŋ qio qi-ki-ŋa sum-nay*
 OBJ-we-GEN house COP-PRF-COND eat-IPL
 'If we'd gone to our house, we would've eaten.' (Fernandez 1968:67)
- (e) *gitin remo a-monaʔbay selane kiyaj beq-oʔ*
 that.CLOSE man OBJ-fat girl rice give-PST.I
 'That man gave rice to the fat girl.' (Fernandez 1968:119)

The objective marker also appears on the comparandum in comparative and superlative formations in Hill Remo.

- (52) (a) *a-selane-ŋ upre may tiur-bay* (b) *may a-gulay-qo-ŋ upre*
 OBJ-girl-GEN over he tall-ADJ s/he OBJ-all-HUM-GEN over
 'He is taller than the girl.' *oli-anra*
 bad
 'He is the worst of all.'
 (Fernandez 1968:67–68)

But in Plains Remo, comparative constructions without the objective marker may be found.

- (53) *niŋ nanɖe nɔ munaʔ*
 I in.comparison.with you big
 You are older than me. (Bhattacharya 1968:79)

While the use of the objective *a-* is (largely) obligatory in the domains outlined above, there are a number of constructions where it may appear in certain instances. Possessive constructions, possibly derived from an existential locative copular formation historically, represent one such formation. These are frequently, but not invariably, found where the possessor is marked with the objective *a-* (and for human nouns, the possessive suffix) *ŋ[a]*, followed by a copular verb.

- (54) (a) *a-semuʔ suʔu-qay qi-ta*
 OBJ-tree raw-fruit COP-NPST.II
 'There is raw fruit on the tree.' (Fernandez 1968:67)

- (b) *samba zunuyɬu tuɣ-oʔ-ta*
 armband earring wear-PST.I-NPST.II
 ‘He has worn an armband and an earring.’ (Fernandez 1968:69)

The possessive or genitive formant is *-ŋ(a)* in Hill Remo *-n(a)* in Plains Remo. It appears after the plural suffix: *sik-saŋ-ŋa* ‘the bone’s’, *remo-le-ŋa* ‘the men’s’, *ɖatipaɖa-ŋa remo-le* ‘the men of Dattipada’ (Fernandez 1968:65). It occurs in the shortened or altered form *-n* in a range of adnominal attributive and non-finite verb forms as well (see sections 3.1.10, 3.1.11, and 3.2.8 below).

- (59) (a) *remo ɖio oroy-ɖen-ta* (b) *nij remo-ŋa tumoʔ*
 man house build-PROG-NPST.II I man-GEN mouth
 ‘The man is building a house.’ *on-t-ij*
 hear-NPST-I
 ‘I hear the man’s voice.’
 (Fernandez 1968:126)
- (c) *kon monaʔbay selane-ŋa ɖio*
 this fat girl-GEN house
 ‘This fat girl’s house.’ (Fernandez 1968:127)

The genitive case (which is identical to what Fernandez calls the possessive) is usually *-na* in Plains Remo. It occurs on a possessor noun that precedes the possessum. Some examples of its use are offered below.

- (60) (a) *nij nej-na ŋgom malkangiri distrikt-bo ɖi-ta*
 I we-GEN village Malkangiri District-LOC COP-NPST.II
 ‘My..our village is (a place) in Malkangiri District.’ [SDM]
- (b) *ɔmɖi-le-na gusunŋger-e-na ninden gusuʔge kɔʔne*
 Omdi-PL-GEN Gusunŋger-PL-GEN primary.clan Gusuʔge DEIC-EMPH
kirsani remo-le
 Kirsani person-PL
 ‘The original clan for the Omdi and Gusunŋger is Gusuʔge; these are Kirsani people.’ (Bhattacharya 1968:79)
- (c) *gojtaŋ-na nzeʔmiʔ gojtaŋ-na nseʔmiʔ* (d) *sajkal-na tiksuy*
 bull-GEN nose bull-GEN nose bicycle-GEN leg/foot
 ‘the bull’s nose’ [SDM] ‘bicycle wheel’ [SDM]
- (e) *bondagada-na remo-le uli sum-to*
 Bondagada-GEN person-PL mango eat-NPST.I
 ‘The people of Bondagada eat mango[es].’ [SDM]

It can be used in lexical[ized] ‘compounds’ as well.

- (61) (a) *titi-n kirime* (b) *tiksuy-na kirime*
 hand-GEN nail foot-GEN nail
 ‘fingernail’ ‘toenail’ [SDM]

One of the uses of the genitive case is to distinguish indefinite generic possessive compounds ‘an elephant leg’ from definite, referential possessors, for example, ‘(the) elephant’s leg’. As the English glosses represent, such a patterning is far from unusual.

- (62) (a) *ati-na tiskuŋ certa* (b) *ati-na mō:ʔ mbaʔar*
 elephant-GEN leg four' elephant-GEN eye two
 'the elephant's four legs' 'the elephant's two eyes' [SDM]
 (c) *ati-na luntur mbaʔar* (d) *ati-na ɖaroʔ*
 elephant-GEN ear two elephant-GEN tusk
 'the elephant's two ears' 'the elephant's tusk' [SDM]
 (e) *ati-na luntur mbaʔar munaʔwe*
 elephant-GEN ear two big
 'the elephant's two big ears' [SDM]

Compare the above 'definite referential' possessor with the genitive case with the indefinite/generic compounds found below.

- (63) (a) *ati boʔb* (b) *ati mō[:ʔ]*
 elephant head elephant eye
 'elephant head' 'elephant eye'
 (c) *ati sunɖoʔ* (d) *ati tiksuŋ*
 elephant trunk elephant leg
 'elephant trunk' 'elephant leg' [SDM]

Finally, although the evidence is scant, the genitive case may also function as the subject (at least of a pronominal subject) of a dependent clause in Remo as well.

- (64) *ɖakra jul-ɔ jul-ɔ a-gəŋ-seʔta mayn kiyəŋ-ntra-maʔ*
 man see-PST.I see-PST.I NEG-CAP-SS he:GEN cooked.rice-gruel-curry
buŋɔ-ki-n-bɔ ui-ga
 put-PLUP-DEP-LOC/DIR go-PST.II
 'The man waited there for some time, but then unable to wait more he returned
 to the place where rice, gruel and vegetables were laid by him.' (Bhattacharya
 1968:149)

Other case forms are often called postpositional elements, though many appear in a bound form, so are well on their way to being actual morphological cases. Fernandez lists 19 of these, three (actually four, and probably more) of which can be used with non-finite verbs to mark dependent clauses of various sources (see also section 4.2 below). Some of these are clearly borrowings, such as the originally Indo-Aryan *pəlay*, one variant of the benefactive or purposive postposition/case clitic. With verbs, it takes a reduplicated stem. With pronominals, it takes the *a-* form in Hill Remo and the genitive in Plains Remo. A nearly functionally identical element is *-sa* with nearly identical formal patterning (e.g. it also occurs with purposive clauses with a reduplicated non-finite verb to mean 'in order to X' or 'for X-ing').

- (65) *susum-pəlay* vs. *a-niŋ-pəlay*
 REDPL:eat-FOR OBJ-I-FOR
 'for eating' 'for me' (Fernandez 1968:110)

Plains Remo

niŋ(-na) palay 'for me, for my sake' *ma palay* 'for what' (Bhattacharya 1968)

Another common case or adpositional clitic is the locative/illative in *-bo*, also found in other South Munda languages.

- (66) (a) *nij nej-na ŋgom malkangiri distrikt-bo qi-ta*
 I we-GEN village Malkangiri District -LOC COP-NPST.II
 ‘“My” “our” village is (a place) in Malkangiri District.’ [SDM]
- (b) *mari kiyaj tɔ ʃipni-bɔ kib-ɔʔ*
 then cooked.rice EMPH small.basket-LOC/DIR pour-PST.I
 ‘Then he poured the rice into a small basket.’ (Bhattacharya 1968:148)
- (c) *ntra patli-bɔ kib-ɔʔ*
 gruel small.pot-LOC/DIR pour-PST.I
 ‘He put the gruel in small pot.’ (Bhattacharya 1968:149)
- (d) *maʔ munʃi-bɔ sɔb-ga*
 vegetable.curry earthen.pot-LOC hold-PST.II
 ‘The veggy curry was held in an earthen pot.’ (Bhattacharya 1968:149)

Note that in its attributive function, the ‘genitive’ may attach to a locative/directional case-marked noun, which then serves to modify another noun, for example, *-bo-n[a]*. Attributive use of a genitive in a wide range of formations may be found in other Munda languages as well, for example, Gorum (Anderson and Rau, this volume).

- (67) *dɔkra bisar-ɔ ɔkɔʔna biri-bɔ-na taas qiʔ-ta*
 man think-PST.I so.much forest-LOC-DEP field.work be-NPST.I
a-dɔkri cucare bug-ɔ-beʔ-ɔʔ-nij
 OBJ-woman in.vain beat-CV-AUX-PST.I-I
 ‘The man then thought (in his mind), “There is so much work in the field; I beat her for nothing.”’ (Bhattacharya 1968:149)

3.1.3 Person

Person as an inflectional category of nouns in Remo is highly restricted and perhaps attested only in rapid speech. In our corpus, there is only one possible form of this type attested in the spontaneous utterance directed from one Remo speaker to another.

- (68) *pe-mba-na ma-y mi*
 2-father-GEN what=name
 ‘What is your father’s name?’

Such a formation is not overly uncommon with a restricted set of largely kin-terms across the South Munda languages (e.g. Gtaʔ). Indeed, the form *pemba* is also a possible form in Gtaʔ with the same meaning and the speakers so recorded may well have been influenced by this sister language spoken in the same region. Such forms are not attested, otherwise, in the admittedly small corpus of recorded Remo data. However, they may simply be of low frequency and, therefore, have not yet (other than this form) been identified in the corpus.

3.1.4 Definiteness

Definiteness *per se* is not a morphological category of Remo and the topic has not received a special investigation yet to date. However, note that in Remo, as in Sora, third person pronominal forms can be used attributively before a noun to serve as a marker of definiteness.

- (69) *māʔe remo-le i-ta*
 they Remo/person-PL go-NPST.II
 'the Remo people go.' [SDM]

3.1.5 Class/gender

Generally speaking, nouns are not distinguishable by gender. Some words, mostly loans, (including personal names) from Indo-Aryan have characteristic *-a* (masculine) or *-i* (feminine) and, thereby, distinguish gender.

- (70) *Buda* 'Masculine name' *Budi* 'Feminine name'
banza 'younger brother's son' *banji* 'younger brother's daughter'
 (Fernandez 1968:78).

A small number of adjective stems are similarly marked in this fashion.

- (71) *sut-i-bay sut-a-bay* 'lame (f, m)'
kunḡ-i-bay kunḡ-a-bay 'mute (f, m)' (also *konḡ*) (Fernandez 1968:88)

Also (and only) *bojr-* 'deaf' *kan-* 'blind' and *paḡḡr-* 'fair complected' (Fernandez 1968:88). The most common way of indicating gender in Remo nouns is with the prefixes (or first compound elements) *layʔ-* 'male' vs. *joḡ-* 'female', for example, *layʔ-bu* 'boar' vs. *joḡ-bu* 'sow' (Fernandez 1968:78). There are suffixal or N₂ compound elements (see section 3.1.10) that encode gender and are used to distinguish human male and female referents, for example, *-boy* 'woman' and *-rem* 'man', for example, *gor-boy* 'Dom woman' *gor-rem* 'Dom man'.

A human/non-human distinction also appears to be relevant in the numeral system of Remo, see below.

3.1.6 Pronouns

The system of Remo personal pronouns includes three persons (first, second, third) and three numbers (singular, dual, plural) to form a nine-way opposition. The forms of the pronouns are as follows. Note that, while dual is recognized in the pronominal system, and for first and second persons, at least is encoded morphologically in the verb, it is not an activated inflectional category for nouns in Remo.

- | | | | |
|------|------------|---------------|-----------------------------------|
| (72) | SG | DL | PL |
| 1 | <i>nij</i> | <i>naḡ</i> | <i>naḡ</i> |
| 2 | <i>no</i> | <i>pa</i> | <i>pe</i> |
| 3 | <i>may</i> | <i>may-pa</i> | <i>may-ʔe</i> (Fernandez 1968:81) |

Subject forms of both nouns and pronouns are unmarked in Remo, but inflected 'objective' forms as well as genitive/possessive forms of pronouns are found. Some examples of illustrating their use are offered below.

- (73) (a) *nij nijja-na ḡgom bondauḡa*
 I self-GEN village Bondauḡa (Bondagaḡa)
 'My own village is Bondagaḡa/Bondauḡa.' [SDM]

- (b) *niŋ-na ugbo?* (c) *no-na uybo?* (d) *may-na ugbo?*
 I-GEN hair you-GEN hair she-GEN hair
 ‘my hair’ [SDM] ‘your hair’ [SDM] ‘her hair’ [SDM]
- (e) *niŋ lay-t-iŋ* (f) *naŋ [m]bayo ton-te-no*
 I sit-NPST-1 we.2DL 2.people stand-NPST.II-1DL
 ‘I am sitting’ ‘We (2) are standing.’ [SDM]
- (g) *nāj glero toŋə-te-naj*
 we.PL all stand-NPST-1PL
 ‘We all are standing.’ [SDM]
- (h) *niŋ tugola toŋ-kə-niŋ* (i) *niŋ tuw/gela lay-ki-niŋ*
 I yesterday stand-T/A-1 I yesterday sit-T/A-1
 ‘Yesterday I was standing.’ ‘I was sitting yesterday.’ [SDM]
- (j) *a-niŋ kiyaj be*
 OBJ-I rice give
 ‘give me rice’ (Fernandez 1968:112)
- (k) *mā?e nju ru mura?-ta?a-maj biri-bo*
 they morning rise-NPST.II OBJ-they-[GEN] highland.field-LOC
biri biri wi-ta
 highland.field go-NPST
 ‘They get up in the morning and go to (work) in their highland fields.’
 [SDM]

Note that, the human numeral/quantifier forms *mbayyo* ‘2/pair of’ and *gulay-ɔo* ‘all of’ may be used with first and second person dual and plural, but the former element never and the latter only rarely appears with third person dual and plural.

- (74) (a) *naŋ (mbayyo) a-Muŋdli paɔa wi-tə-naŋ*
 we.2 pair.of OBJ-M go-NPST.II-1DL
 ‘The pair of us go to M.’
- (b) *naŋ (gulay-ɔo) miŋdip kiyaj sum-tə-nay*
 we all.of evening rice eat-NPST.II-1PL
 ‘We eat rice in the evening.’ (Fernandez 1968:69–70)
- (c) *may-?e iskol- bo? sa-ga*
 she/he-PL school-LOC/DIR come-PST.II.3
 ‘They came to school.’
- (d) *may-pa (*mbayyo) iskol-bo? sa-ga*
 she/he-DL *pair.of school-LOC/DIR come-PST.II.3
 ‘The two of them came to school.’ (Fernandez 1968:70)

3.1.7 Demonstratives

The demonstrative system of Remo is decidedly complex in terms of deictic space. There is a neutral or proximal element *kon* and four distal elements distinguished by the degree of closeness to the speaker or deictic space. These are *gitin* (close but most commonly used distal), *gusu* (mid-1), *ro* (mid-2), and *geta* (~ *gəta*) (distant or far).

- (75) (a) *kon gisiŋ niŋ-ŋa* (b) *kon remo* (c) *kon mbaʔar bire*
 this chicken I-GEN this man this two stone
 ‘This chicken is mine.’ ‘this man’ ‘these 2 stones’
 (Fernandez 1968:82)
- (d) *gitin gusoʔ* (e) *gitin remo-le*
 that.CLOSE dog that.CLOSE man-PL
 ‘that nearby dog’ ‘those nearby men’ (Fernandez 1968:82)
- (f) *gusu gibesoʔ ʒonɖaʔ su-sum-ɖen-ta*
 that.MID.1 bear honey REDPL-eat-PROG-NPST.II
 ‘That somewhat close bear is eating honey.’ (Fernandez 1968:82)
- (g) *no-ŋa taŋja ro*
 you-GEN axe that.MID.2
 ‘That axe (some distance away) is yours.’ (Fernandez 1968:82)
- (h) *geta remo kon remo anʔa*
 that.DIST man this man NEG
 ‘that man, not this man’ (Fernandez 1968:82)
- (i) *ɔmɖi-le-na gusuŋgere-na ninden gusuʔge kɔʔ-ne*
 Omdi-PL-GEN Gusungger-PL-GEN primary.clan Gusuʔge DEIC-EMPH
kirsani remɔ-le
 Kirsani person-PL
 ‘the original clan for the Omdi and Gusungger is Gusuʔge; these are
 Kirsāni people.’

Certain ones of these demonstrative elements may be used pronominally without an accompanying noun and, then, be marked for number. These number-marked demonstrative pronouns include *kon-e* ‘these’ *gitin-a* ‘those [2] (close)’ or *kon-oŋ-pa* ‘these 2’ (Fernandez 1968:83).

The Hill Remo interrogative pronouns include the following: *arn* ‘what/which’ *ma* ‘what’ *ja* ‘who’ *ja-ŋ[a]* ‘whose’.

- (76) (a) *arn ŋgom remo sa-ga* (b) *ro-ŋ ma*
 what village man come-PST.II.3 that.MID.2-GEN what
 ‘What village does the man come from?’ ‘What is that?’
 (Fernandez 1968:83)
- (c) *gitin remo ja* (d) *kon ja*
 that.CLOSE man who this who
 ‘Who is that man?’ ‘Who is this?’ (Fernandez 1968:83)
- (e) *kon ja-ŋ taŋja* (f) *gitin ja-ŋ oʔon-ɖay*
 this who-GEN axe that.close who-GEN child-KIN
 ‘Whose axe is this?’ ‘Whose child is that?’
- (g) *no-ŋ imi ma*
 you-GEN name what
 ‘what’s your name?’ (Fernandez 1968:83)
- (h) *oho kɔʔn ma kakurti inɖɔ mʔɔʔ ɖigidiɖigi.ɔm*
 EXCL this what trouble alas cloth in.spite.of
 ‘Oh! What is this trouble, in spite of there being a cloth.’
 (Bhattacharya 1968:64)

For some speakers ‘why’ is combinatorial *ma-sa*, literally ‘on account of what, for what’.

When functioning as subjects and referring to people, various kinds of indefinite and quantifier pronouns formally marked by the suffix *-do* are found in Hill Remo. This set of elements consists of the following: *gulay-do* ‘all’, *tate-do* ‘all’ (loan), *soman-do* ‘everyone’, *rapte-do* ‘many’, and *una-do* ‘few’. This element also appears with numeral stems (see section 3.1.8).

- (77) (a) *gulay-do qay-ga* (b) *tate-do dio-bo?* *wi-ga*
 all climb-PST.II.3 all house-LOC/DIR go-PST.II.3
 ‘All climbed.’ ‘All went home.’ (Fernandez 1968:85)
- (c) *rapte-do i-qen-ta* (d) *una-do a-Munḍlipaḍa wi-ga*
 many return-PROG-NPST.II few OBJ-M go-PST.II.3
 ‘Many are returning.’ ‘Few went to Mundlipada.’
- (e) *soman-do a-dio wi-ga*
 everyone OBJ-house go-PST.II.3
 ‘Everyone went home.’ (Fernandez 1968:85)
- (f) *tɔŋre ~ tɔŋrɔ* there (g) *ɖɔkre* there (Bhattacharya 1968:75–79, 82)

3.1.8 Numerals

The numbers ‘one’ through ‘six’ showed Munda features, mixed with Indo-Aryan features in Fernandez’s materials from the 1960s. Above ‘six’ and the numeral system increasingly converges with local Indo-Aryan models. For younger speakers now, the restriction has extended largely down to the number ‘three’ with numbers over ‘four’ obviously loaned from Desia.

- (78) Bondagada Remo
mujō[ʔ] ~ mujū? ‘one’
mbaʔar ‘two’
ŋgīʔi ~ ŋiʔi ‘three’
catta ~ carta ‘four’
pā[n]tra ‘five’
coḍa ‘six’
sa:d[rə] ‘seven’
atḷ ‘eight’
no: ‘nine’
dos ‘ten’
e:garo ‘11’
kuḷie ‘20’
ekusi ‘21’
baisi ‘22’ [SDM]

Numbers ‘one’ to ‘five’ show idiosyncratic marked forms which explicitly refer to humans. Generally, these appear to be derived from corresponding non-human forms by the addition of some element. These elements are called the ‘human classifier suffixes’ by Fernandez (1968:88). They include the following: *-da* found in *miṅḍa* ‘one’, *-yo* in *mbayyo* ‘two’, *-o* in *witio* ‘few’ (and the adjective *somano* ‘straight’), *-do* which occurs with *moloydo* ‘five’, *ŋgeṅdo* ‘three’ *uḍuṅdo* ‘four’ and the indefinite pronouns. Finally, numbers from ‘six’ and higher take the element *-lok* of Indic origin.

(79)	Human	Non-human
	<i>miŋɖa remo</i> ‘one man’	<i>muy gisiŋ</i> ‘one chicken’
	<i>mbayyo selane</i> ‘two girls’	<i>mbaʔar bire</i> ‘two stones’
	<i>ŋgeŋɖo ŋgere</i> ‘three boys’	<i>ŋgiʔi gusoʔ</i> ‘three dogs’
	<i>tiʔiri-lok remo</i> ‘six men’	<i>tiʔiri goytaŋ</i> ‘six cows’
	<i>witi-o ŋgere</i> ‘few boys’	<i>witi gisiŋ</i> ‘few chicken’

Note that nouns following numerals usually remain in a singular form in Remo, but may optionally appear in the plural (e.g. *tiʔiri-lok remo* ‘six men’ vs. *ŋgeŋɖo ŋger-e* ‘three boys’).

While numerals (adjectives) often precede the noun they govern, they may also follow them as well in Remo. This is true whether the numeral is an original Munda root or a loan element.

(80)	(a)	<i>ati-na</i>	<i>tiskuŋ</i>	<i>certa</i>	(b)	<i>ati-na</i>	<i>mō:ʔ</i>	<i>mbaʔar</i>
		elephant-GEN	leg	four		elephant-GEN	eye	two
		‘the elephant’s four legs’				‘the elephant’s two eyes’ [SDM]		

As aforementioned, other cardinal numerals show an Indic-looking form, at least on the root level. The actual attested form consists of the Indic root followed by an augment, followed by the particle *-ʔa* for non-humans (Fernandez 1968:89). The augments are *-gə* with *so* ‘6’, *sat* ‘7’, *aʔ* ‘8’ *no* ‘9’ and *dos* ‘10’; *-ə* with *bar* ‘12’ *ter* ‘13’ *pondr* ‘15’ *sol* ‘16’ *sort* ‘17’ and *otr* ‘18’, \emptyset with *egar* ‘11’ and *onis* ‘19’, and *-u* with *sod* ‘14’. Note that the non-human element *-ʔa* may appear with a very restricted number of adjectival forms as well, for example, *bolʔa* ‘pretty, good’; *gigebɖa* ‘hot’ (Fernandez 1968:89). Examples include *satgəʔa* ‘7’, *aʔgəʔa* ‘8’, *nogəʔa* ‘9’, *dosgəʔa* ‘10’, *egarʔa* ‘11’, *barəʔa* ‘12’, *terəʔa* ‘13’, *soduʔa* ‘14’, *pondrəʔa* ‘15’, *soləʔa* ‘16’, *sotrəʔa* ‘17’, *otrəʔa* ‘18’, *onisʔa* ‘19’, *koreʔa* ‘20’ (Fernandez 1968:93)

Lower numerals in Remo also show variation with numerals borrowed from Indo-Aryan.

(81)	Human	Non-Human	Indic-Origin Alternate/ Higher Number Form
	<i>miŋɖa</i>	<i>muy</i>	‘1’ ~ <i>ek</i>
	<i>mbayyo</i>	<i>mbaʔar</i>	‘2’ ~ <i>dwi</i>
	<i>ŋgeŋɖo</i>	<i>ŋgiʔi</i>	‘3’ ~ <i>tin</i>
	<i>uʔundɖo</i>	<i>uʔũ</i>	‘4’ ~ <i>sar</i>
	<i>moloyɖo</i>	<i>moloy</i>	‘5’ ~ <i>pans</i>
	<i>tiʔirilok</i>	<i>tiʔiri</i>	‘6’ ~ <i>sogəʔa</i> (Fernandez 1968:94)

The word for ‘20’ in Remo is *kore*. This serves as the basis for the Remo vigesimal numeral system. Generally speaking only ‘21’, ‘31’, etc. show the Munda lower numeral stem, while all other higher numbers utilize the Indic forms: *kore muy* ‘21’ but *kore dwi* ‘22’, *kore tin* ‘23’, etc. Higher numbers are derived from 20, thus one finds *kore dos* ‘30’, *kore sol* ‘36’, *kore onis* ‘39’. Then from *dikori* ‘40’ comes *dikori dos* ‘50’, *dikori sol* ‘56’, etc. On occasion, certain speakers will use compound forms of higher numbers where the tens uses the Indic word but the ‘1’–‘5’ shows the Munda stem: *kore uʔũ* ‘24’ *dikori moloy* ‘45’ (Fernandez 1968:94). Sometimes a final *-o* is found in higher numerals, for example, *kore otəro* ‘38’ or *kore solo* ‘36’ (Fernandez 1968:92). No information is currently available on the formation of ordinal, distributive, collective, or fractional numerals in Remo.

- (82) (a) *naŋbay ton-te-no*
 we.2DL=2.people stand-NPST.II-1DL
 ‘We two are standing.’
- (b) *naŋ bayo ton-te-no*
 we.2DL 2.people stand-NPST.II-1DL
 ‘We two are standing.’ [SDM]
- (c) *ati-na luntur mbaʔar* (d) *remo-na luntur mbaʔar[a] ɖove*
 elephant-GEN ear two man-GEN ear two small
 ‘the elephant’s two ears’ ‘the man’s two small ears’

3.1.9 Adpositions

Remo makes extensive use of bound/enclitic postpositional or case elements. These express both a range of local or directional semantic categories as well as benefactive, purpose, etc. As aforementioned, in many instances a noun may either appear in one of these postpositional or case forms or in the general objective form in *a-*. Such case/fused adpositional forms include the locative/illative *-bo[ʔ]* ‘in(to)’ the subessive *-aluŋ* ‘under’ and the adessive *-boroŋtere* ‘near’.

- (83) (a) *niŋ ɖio-boʔ gay-t-iŋ* (b) *niŋ korji-bagboʔ layk-t-iŋ*
 I house-LOC/DIR enter-NPST-I chair-SUPERESS/LAT sit-NPST-I
 ‘I enter the house.’ ‘I sit on the chair.’ (Fernandez 1968:68)
- (c) *oʔom gusoʔ-bitre ɖi-ta*
 arrow dog-INNESS COP-NPST.II
 ‘The arrow is in the dog.’
- (d) *umporãy bire-uŋder layk-gə-ta*
 husband:KIN stone-AGAINST sit-PST.II-NPST.II
 ‘The husband sat against the stone.’ (Fernandez 1968:68)
- (e) *gitin remo semuʔ-aluŋ layk-gə-ta*
 that.CLOSE man tree-SUB sit-PST.II-NPST.II
 ‘That man is sitting/has sat under the tree.’ (Fernandez 1968:98)
- (f) *may gusoʔ-boroŋtere layk-gə-ta*
 s/he dog-NEAR sit-PST.II-NPST.II
 ‘He has sat/is sitting near the dog.’
- (g) *tənaŋtur-nande kukusa gay-ga*
 fence-THROUGH tiger enter-PST.II.3
 ‘The tiger entered through the fence.’ (Fernandez 1968:98)

In a small number of instances, a noun may appear with both the objective *a-* and a case or bound postpositional element.

- (84) *selane a-remo-bagboʔ layk-gə-ta*
 girl OBJ-man-SUPER sit-PST.II-NPST.II3
 ‘The girl has sat upon the man.’ (Fernandez 1968:98)

According to Fernandez (1968:97), there are 19 bound postpositional or case elements in Remo that appear with nouns and pronouns. Three (actually four if one includes *-sa*) of these may be found with verbs as well, namely *-aluŋ* ‘under, beneath’ *-bagboʔ* ‘on,

upon' and *-pəlay* 'for'. In the case of the first of these, it appears to function like an incorporated adverb. In the last instance, this forms a purposive subordinate clause (see section 4 below).

Note that one basic way of conjoining nouns belonging to this sub-class of nominal elements in Remo is with the comitative or connective *bfojroŋ* 'with, and'.

- (85) *selane broŋ ŋgere*
 maiden:PL with young.man-PL
 'The young men and women together.' [SDM]

3.1.10 Derivation

Derivation in Remo as a productive process is relatively weakly developed, although there is evidence that a more elaborate machinery for deriving nouns was available in a previous stage of the language. Specifically, a large number of unproductive elements appear to have been used to derive the free forms of Remo nouns from their corresponding roots. Such a system is commonly found in South Munda languages.

Nouns occur as mono-syllabic or bi-syllabic and mono-morphemic, for example, *mo?* 'eye' *biŋe* 'stone' *ola?* 'leaf, paper' *reŋi* 'spine' *narom* 'vein, nerve' *suti?* 'clitoris' (Fernandez 1968:71). Nouns may also be bi- or poly-morphemic. This includes composite and compound nouns consisting of N₁ and N₂ (and N₃) components like *tiksuy* 'leg' or *ŋijja-buruŋ* 'house fence' (Fernandez 1968:72) or derived. The processes of derivation include prefixing, infixing, and suffixing, reduplication, etc.

Sometimes the meaning of one element or another is opaque in Remo, for example, *leym-puy* 'waist' *siy-e* 'coloured thread' *su-za* 'needle' (Fernandez 1968:72). Another common (historical) derivational process seen in Remo nouns is reduplication *lə-lap* 'butterfly' *si-ser* 'song' *titi* 'hand, arm' (Fernandez 1968:72).

Remo makes extensive use of the pan-Munda (and pan-Austroasiatic) nominal infix *-Vn-* to derive nouns from verb stems. Many of these have instrumental meanings.

- (86) *sug* 'sweep' > *slunlug* 'broom'
ŋənay 'ghat' < *ŋayks* 'climb'
sinia 'fish hook' < *siaŋl* 'to fish'
sini 'bird trap' < *siŋl* 'snare birds'
pine 'flute' < *peŋl* 'play musical instrument'
gine 'tooth' < *gey* 'chew'
gumura? 'spindle' < *gurag* 'spin'
ŋənaray 'metal lid' < *ŋəray* 'cover a pot'
tənaβ 'thatched roof' < *tab* 'remove' (Fernandez 1968:79)

Such nouns can be derived from verbs that have incorporated nouns in them, for example, *sugsugbo?* 'to comb' > *slunlugbo?* 'hair comb', with the incorporated noun *-bo?* 'head'.

One unusual and restricted use of an *-n* infix is to form duals with certain nominals. Rather than the *-n* quasi-instrumental nominalizing infix this element is, perhaps, more likely to reflect a semantic extension of the reciprocal prefix *-n* (see 3.2.7).

(87) *bīyaŋ* (younger) brother > *bīniyaŋ* ‘two brothers’ (Bhattacharya 1968:94)

A curious feature of Remo is that a small number of nouns historically are inflected verb forms, like *rim-oʔ-ta* ‘quarrel’ which appears to be a present perfect third singular verb form of ‘fight’ (Fernandez 1968:72).

Fernandez divides the structure of Hill Remo nouns into prebases (i.e. prefixes) roots and postbases (= suffixes). One type of noun consists of a Root and an opaque suffixal element. [Root+Suffix] *oʔon-day* ‘someone else’s daughter/son’ *kunɖaʔ-i* ‘buttock, rectum’ (Fernandez 1968:72). Some roots are always bound to a suffix or N₂ compounding element, for example, *sik-say* ‘bone, skeleton’. Prefix + Root is the most common combination.

(88) *gi-siŋ* ‘chicken’ *gu-soʔ* ‘dog’
ŋ-gom ‘village’ *ŋ-kuswī* ‘jackfruit’
so-ma ‘curry’ *ka-gilas* ‘glass’ (loan) (Fernandez 1968:73)

Complex forms are also attested to a limited degree. The patterns attested include Root+Gen+Root *tum-na-soʔ* ‘dog’s snout’ (Fernandez 1968:73).

(89) *titi-n* *kirime* *tiksuy-na* *kirime*
 hand-GEN nail foot-GEN nail
 ‘fingernail’ ‘toenail’ [SDM]

(Prefix+Root)+Root is the most common complex pattern *gisimɔʔ* ‘eyeglasses’ (chicken+eye) *n-seʔ-mi* ‘nose’ *gi-be-soʔ* ‘bear’ *n-sak-pi* ‘bird’s nest’ *ŋ-ger-sū* ‘boy’s dormitory’ (Fernandez 1968:73). In one anomalous form, the combination Root+(Prefix+Root) is found. This is *bayʔ-ŋ-gre* ‘cradle’ (Fernandez 1968:73).

Compound nouns are also found in Remo. Formally, these show a range of types. The most common type is Root+Root, for example, *suku-baʔi* ‘heart’ *taŋgi-mali* ‘brass necklace’ (Fernandez 1968:73); Prebase+Root+Root, for example, *ŋ-ger-sela* ‘unmarried female’ *n-toʔ-siŋ* ‘egg’ (Fernandez 1968:73). On rare occasion three roots may be combined: *e-taŋ-goy* ‘cow’s rope halter’ (Fernandez 1968:73) A Root may also appear with a participle in Remo compounds, *susum-ti* ‘right hand’ (cf. *sum* ‘eat’). Note that the same element may appear as the first or second part of a compound, for example, *tiksuy* ‘leg’ vs. *suyom* ‘arrow shaft’ (Fernandez 1968:72).

Most of the prefixes occur infrequently, perhaps, in a very small number or even a single form. Others occur relatively commonly or in quasi-definable semantic groups of words. For example, the ‘animal’ classifier *gV-* is attested in a number of words with allomorphy conditioned as follows:

(90) *gV-*
 > *gi-/__Ci/e*
 > *gu-/__Cu/o*
 > *gə-/__Ca*

Examples of this prefix are to be found in such common words as *gisimɔʔ* ‘chicken’ *gise* ‘grasshopper’ *gəga* ‘crow’ *gəlayʔ* ‘castrated bull’ *gu-bu* ‘pig’ *gu-soʔ* ‘dog’ (Fernandez 1968:74).

Another common prefix found in Remo is the historical syllabic nasal *N- which assimilated to the place of the following consonant (*m,n,ŋ*). This is found in a wide range of words: *mbur* ‘iron bar’ *nsu* ‘knife’ *ŋ[ɖ]rem* ‘bird wing’ *ŋgom* ‘village’ (Fernandez 1968:74). A small number of words suggest that perhaps *ŋ* is being generalized or that there is a separate nominal formant that is *ŋ-* (Fernandez 1968:75).

Such a non-homo-organic η - is found in such words as *ηber* ‘slingshot’, *ηsoη* ‘spade’, and *ηraηgay?* ‘rib’.

Opaque or unique initial elements may be found in Remo *so-ma* ‘curry’ *ka-gilas* ‘cup, glass’ *su-tubu* ‘earth, soil’ *bay?-ηgre* ‘cradle’ *la?-gomar* ‘scorpion sp.’ *oli-anηa-rem* ‘bad man’ (Fernandez 1968:75).

Some noun roots appear commonly in combination with other roots. These include *-rem* ‘man’, *-ti* ‘hand’, and *-qa?* ‘water’. They serve almost as semantic classifiers in some forms. Examples with these elements include *e-rem* ‘evil magician’ (< *el-*) *buso-rem* ‘adult’ *sio-rem* ‘human testicles’ *sakar-rem* ‘rich man’ (Fernandez 1968:76). *seserrem* ‘songster’ = (Ramachandra Rao 1981:85) *buso-rem* ‘adult’ *sakar-rem* ‘rich man’ *bol-rem* ‘good man’ (Fernandez 1968:80); *kirim-ti* ‘fist’ *gata-ti* ‘palm of hand’ *basa?-ti* ‘left hand’ *slunluku-ti* ‘elbow’ *tlanlay-ti* ‘arrow nock’ (Fernandez 1968:77); *bo-qa?* ‘natural spring’ *non-qa?* ‘honey’ *gay-qa?* ‘large clay pot’ *lumu-qa?* ‘dew’, *jur-qa?* ‘cloud’ *olo?g-qa?* ‘waterfall’ *kiη-qa?* ‘river’ *liη-qa?* ‘rain’ (Fernandez 1968:77).

More rare second elements include *-siη* ‘chicken’, *-buη* ‘buffalo’, *-boy* ‘feminine’, *-qay/-qēy/-āy* ‘kinship reference’, and *-mo?* ‘eye, face, mouth’, etc. (Fernandez 1968:77). A sample of the forms that occur with these include *kurlak-siη* ‘chicken wattles’ *joη-siη* ‘hen’ *o?on-siη* ‘chick’ *gupa-siη* ‘ceremony type with chicken featured prominently’ *rlunluk-siη* ‘courtyard’; *n-to?-siη* ‘egg’; *buηte* ‘buffalo’, *se-buη* ‘marriage feast where buffalo is slaughtered’, *joηbuη* ‘buffalo cow’ *lay?-buη* ‘buffalo bull’; *salag-boy* ‘marriageable girl’ *gor-boy* ‘Dom female’ *lu-boy-qa* ‘woman’s bead necklace’ *n-boy-bu* ‘adolescent female pig’ *liy-boy* ‘husband’s younger sibling’; *o?on-qay* ‘daughter, son’ (when speaking of someone else’s child) *kuni-qēy* ‘wife’ (when husband speaks) *umpor-āy* ‘husband’ (wife speaking); *nsa?-mo?* ‘eyelash’ *qa?-mo?* ‘tears’ *sar-mo?* ‘face’ *tu-mo?* ‘mouth, beak’. Highly restricted suffixes include *-ne* which is found in two forms only *sela-ne* ‘girl, woman’ and *gulay-ne* ‘boy’.

Some examples of basic vocabulary showing different kinds of compounding and other derivational processes include the following:

- (91) *sumukuti* ‘elbow’ *gulay titi* ‘arm’ [all hand] *titi* ‘hand’
 ō?onti ‘finger’ *jomti* ‘finger’
 ō?ō[n]suη ‘toe’ *jomsuη* ‘toe’
 tanarom ‘shoulder’

The possessive element *-ηa* straddles the border between derivation and inflection. Functionally, it appears inflectional, appearing in genitive formations. However, it appears to also have derivational functions as well, for example, it may combine with a participial verb form to create a noun formally: *su-sup-ηa* ‘round pot handle’ *tu-ηa-so?* ‘dog’s snout’ *tu-ηa-goy* ‘cow’s snout’ (Fernandez 1968:77). Note also the following morphophonological changes in compounding (participle + noun) and reduplication, respectively, *oy-rig* ‘ripe Sua millet’ < *oys* ‘harvest’, *mi-me[?]* ‘dance’ < *meq* (Fernandez 1968:79). Note that variation in noun formation can be seen in individual instances, for example, *rim-o?-ta* ‘quarrel’ (Fernandez 1968:72) vs. *ηrim* ‘quarrel’ < *rim* ‘to quarrel’ (Fernandez 1968:79). As alluded to above, Remo finite verbs may sometimes be lexicalized as nouns (Fernandez 1968:79).

- (92) *re* ‘cut’ > *regata* ‘wound’
 sil ‘be sick’ > *sita* ‘pain, sickness’
 jur ‘be chilled’ > *jurgata* ‘chill’
 aη ‘be taboo’ > *aηto* ‘taboo’
 suηol ‘greet’ > *suηo-t-iη* ‘Hello’ (Fernandez 1968:79)

3.1.11 Adjectives

Adjectives as a word class is perhaps more justifiable in Remo than it is in many other Munda languages. Adjectives may precede (most commonly) or follow nouns that they modify. When used attributively they remain uninflected (except the clitic objective prefix which then appears on the leftmost word in the phrase which may be an adjective), but an adjective may undergo zero-derivation and form nouns and then, like demonstratives or numerals, may take, for example, number inflection. Adjective–Noun case and/or number concord is not known in Remo grammar.

- (93) (a) *nāj glero toŋgə-te-naj*
 we.PL all stand-NPST.I-1PL
 ‘We all are standing, have all stood up.’
 (b) *ati-na luntur mbaʔar munaʔwe*
 elephant-GEN ear two big
 ‘the elephant’s two big ears’ [SDM]
 (c) *remo-na luntur mbaʔar qove*
 man-GEN ear two small
 ‘the man’s two small ears’ [SDM]
 (d) *monaʔ-bay selane*
 fat-ADJ girl
 ‘the fat girl’

Adjectives may be morphologically marked by the modifier suffix *-bay*, a syllabic nasal prefix, N-, or both. Note that this includes numeral stems as well.

- (94) *m-baʔar* ‘two’ *ŋ-giʔi* ‘three’
buso-bay ‘adult’ *sakar-bay* ‘rich’ *bol-bay* ‘good’
saysay-bay ‘yellow’ *bubay* ‘multicoloured’
tali-rosuno ‘garlic’ > *tuluy-bay* ‘white’
monaʔ ‘grow big’ > *monaʔ-bay* ‘big, fat’ > *monaʔ-ŋgom* ‘town’
n-teraʔ-bay ‘green, blue’

Note that *-bay* may attach to reduplicated verbal forms as well.

- (95) *gigeb-bay* ‘hot’ < *geb* ‘be hot’ *zuzurbay* ‘weak’ < *jur* ‘be chilled’ (Fernandez 1968:80)

Note also the idiosyncratic and unique formant *-lu-* in *goy-lu-bay* ‘dead’.

Some adjective stems must occur with a suffix in Remo, although these are often loans and copy with them some of the gender/number morphology from their Indo-Aryan source, as well as often taking the adjectival suffix- *bay*. Such adjective stems include *kan-* ‘blind’, *ur-* ‘smart’, *benʔ-* ‘young’, *qeng-* ‘tall’, *seroʔ-* ‘dirty’, and *rentem-* ‘lean’.

- (96) *kan-a-bay* *remo*
 blind-MASC-ADJ man
 ‘the blind man’

As aforementioned, the objective marker also appears on the comparandum in comparative and superlative formations in Hill Remo. In Plains Remo, comparative constructions without the objective marker may be found. Comparatives in Hill Remo use *upre* ‘over’ and superlatives add *gulay[ʔo]* ‘all’ to this.

- (97) (a) *a-sela-ne-ŋ upre may tiur-bay*
 OBJ-girl-GEN over he tall-ADJ
 ‘He is taller than the girl.’
- (b) *may a-gulay-ŋo-ŋ upre oli-anra*
 s/he OBJ-all-HUM-GEN over bad
 ‘He is the worst of all.’ (Fernandez 1968:67–68)
- (c) *niŋ nanɖe nɔ muna?*
 I in.comparison.with you big
 ‘You are older than me.’ (Bhattacharya 1968:79)

3.1.12 Adverb(ial)s

Reduplication is frequently found in adverbials. This kind of reduplication may belong historically to a subsystem of expressive discourse, a common feature of Munda and Austroasiatic languages, generally. Examples include *baŋi-baŋi* ‘well’ *rumaru* ‘etc.’ *sorosoro* ‘loosely’ *toŋrotoŋ* ‘crookedly’.

- (98) (a) *gulayne oŋom baŋi-baŋi twin-to*
 boy arrow well shoot-NPST.I
 ‘The boy shoots the arrow well.’
- (b) *may sapuŋ uɖ-o? rumaru*
 s/he wine take-PST.I etc
 ‘He took wine and what-not.’
- (c) *niŋ a-taŋja sorosoro sog-o?-t-iŋ*
 I OBJ-axe loosely hold-PST.I-NPST-I
 ‘I have held the axe loosely.’
- (d) *remo toŋrotoŋ uriŋ-to*
 man crookedly walk-NPST.I
 ‘The man walks crookedly.’

Other adverbials may be uninflected nouns functioning adverbially.

- (99) (a) *niŋ miŋɖip ɖio-gəri a-wi-t-iŋ*
 I evening house-ABL NEG-go-NPST-I
 ‘I do not go from the house in the evening.’
- (b) *naŋzur a-biri wi-t-iŋ*
 morning OBJ-forest go-NPST-I
 ‘In the morning I go the forest.’
 NB *naŋzur* ‘dawn’ (Fernandez 1968:104)

Munda numeral stems may be found in frozen form in a range of temporal adverbs referring to days: *misin* ‘one day’ < *muy* ‘one’ *barsi* ‘two days’ *ersi* ‘three days’ *bartoga* ‘day before yesterday’ *baurtoga* ‘two days before yesterday’ *inlo* ‘fifth day before yesterday; fifth day after tomorrow’ (Fernandez 1968:104). Some examples of temporal adverbs in use include the following:

- (100) (a) *niŋ tugola toŋ-kə-niŋ* (b) *niŋ tuwela lay-ki-niŋ*
 I yesterday stand-T/A-1 I yesterday sit-T/A-1
 ‘Yesterday I was standing.’ ‘I was sitting yesterday.’ [SDM]

- (c) *nij oʔeʔga toŋg-edi-n*
 we.PL right.now:EMPH stand-AUX:T/A-1PL
 ‘We are standing now.’ [SDM]

Adverbs referring to years include the noun *boros* ‘year’ seen in *borsek* ‘one year’ *diboros* ‘two years’ and the combining forms [CF]=*mo* seen in *oymo* ‘this year’ and *ndimo* ‘last year’ (Fernandez 1968:105).

The locatives include *aka* ‘here’ *ate* ‘there’ *ijkalīngəri* ‘from here’ (Fernandez 1968:106). The interrogative adverbs of Hill Remo consist of the following: *ari* ‘where’, *masa* ‘why’, *miri* ‘how’, *oyja* ‘how much’, *oyjabele* ‘when’, and *tor* ‘to where’.

- (101) (a) *masa remo wi-ḍen-ta-ki* (b) *goytaŋ tor wi-ḍen-ta*
 why man go-PROG-NPST.II-Q COW to.where go-PROG-NPST.II
 ‘Why is the man going?’ ‘(to) Where is the cow going?’
 (Fernandez 1968:107)
- (c) *no miri wi-to-no*
 you how go-NPST.I-2
 ‘How will you go?’
- (d) *no-ŋa soka oyja dabu-ḍen-ta*
 you-GEN shirt how.much money-PROG-NPST.II
 ‘How much did your shirt cost?’ (Fernandez 1968:107–108)
- (e) *no oyjabele lem-oʔ wi-to-no* (f) *remo ari ḍi-ta*
 you when sleep-PST.I go-NPST.I-2 man where COP-NPST.II
 ‘When will you go to sleep?’ ‘Where is the man?’
 (Fernandez 1968:108)

Lastly, *aŋra* is used to mean ‘no’ and ‘not’; *ansa* also means ‘no’. *oʔon* is ‘yes’.
 (Fernandez 1968:107–108)

3.2 Verbal morphology

Like the other Munda languages, the verbal morphology of Remo is where the greatest complexity is expressed, although the Remo verb is in some senses among those that are least morphologically developed of the Munda languages. Categories such as the person/number of the subject, a variety of TAM formations, and negation are to be included in the make-up of the Remo verb form. In addition, like other South Munda languages, a limited degree of mainly lexicalized noun incorporation is evident in its structure. Lastly, Remo makes extensive use of a diverse functional and formal array of auxiliary verb constructions (many fused into large unverbated complexes).

Verb as a formal word category in such North Munda languages as Mundari has generated a small body of literature (see Evans and Osada 2005 for a recent discussion). Largely, nominal stems may simply be used verbally by making them predicates and adding finite verbal morphology to them. Whether one can justify the word class *per se* in Remo is yet to be fully investigated.

- (102) *maḍpru ns̄-ga*
 god spade/hoe-PST.II
 ‘God became a spade.’ (Bhattacharya 1968:83)

3.2.1 Subject

As in Gutob (and Kharia), Remo uses a series of subject enclitics (103). In the case of the first singular marker, there are two allomorphs, with the vowel-initial allomorph used after consonantal TAM forms like the non-past in *-t(V)-*. The past tense uses a third singular allomorph in *-ga* for third person subjects (in the ‘second’ conjugation). Elsewhere, third person subject (of any number) has no formal realization.

(103)	1st	2nd	3rd
SG	<i>-niŋ, -iŋ</i>	<i>-no</i>	<i>-Ø, -ga</i>
DL	<i>-naŋ</i>	<i>-pa</i>	<i>-Ø -ga</i>
PL	<i>-nay</i>	<i>-pe</i>	<i>-Ø, -ga</i>

Thus, there is a singular: dual: plural opposition in the first and second persons but not in the third. Examples of some sample inflected forms showing the use of the various subject enclitics (or suffixes) are seen in (104).

- (104) (a) *ju-tə-naŋ* see-NPST.II-1DL ‘we two see’
 (b) *wi-g-no* go-PST.II-2 ‘you went’
 (c) *ju-to* see-NPST.I ‘she/he sees’ (Fernandez 1968:25)
- (d) *sap-gə-tə-nay* come-PST.II-NPST.II-1PL ‘We all have come.’
 (e) *sum-oʔ-ke-pe* eat-PST.I-PRF-2PL ‘Y’all had eaten.’ (Fernandez 1968:26, 22)
- (f) *sap-kə-pa* come-PLUP-2PL ‘You 2 had come.’
 (g) *maypa wi-ga* they.2 come-PST.II ‘They 2 went.’ (Fernandez 1968:26, 25)
- (h) *niŋ laj-t-iŋ* I sit-NPST-I ‘I am sitting.’
 (i) *niŋ tugola toŋ-kə-niŋ* I yesterday stand-T/A-1 ‘Yesterday I was standing.’ [SDM]
- (j) *naŋ [m]bajo ton-te-no[ŋ]* we.2DL 2.people stand-NPST.II-1DL ‘We two are standing.’ [SDM]
- (k) *bondagada-na remo-le uli sum-to* Bondagada-GEN person-PL mango eat-NPST.I ‘The people of Bondagada eat mango[es].’ [SDM]
- (l) *māʔē njuru muraʔ-taʔ ..a-māj biri-bo* they morning rise-NPST.II OBJ-they-[GEN] highland.field-LOC
biri biri wi-ta highland.field go-NPST.II ‘They get up in the morning and go to (work) in their highland fields.’ [SDM]
- (m) *nāj glero toŋgə-te-naj* we.PL all stand-NPST-IPL ‘We all are standing.’ [SDM]

Note that the third person past allomorph is used in the pluperfect forms with the perfect in *-kə-* suggesting that this was a second conjugation verb when it still functioned as an auxiliary, before becoming the tense/aspect marker that it currently is.

- (105) (a) *sum-oʔ-kə-ga* eat-PST.I-PRF-PST.II.3 ‘They 2 had eaten.’
 (b) *sap-kə-ga* come-PLUP-PST.II.3 ‘They had come.’

3.2.2 *Object types*

Unlike South Munda languages such as Gorum, Juang, or Sora, Remo makes no use of morphologically encoded object categories within the verb.

3.2.3 *Tense*

Tense(/aspect) marking shows a moderately complex system in Remo, with verbs dividing into two rough conjugational classes, at least as far as the past tense is concerned. Originally, the distinction may have been one of transitive (Class/Conjugation I) vs. intransitive (Class/Conjugation II) or ‘active’ vs. ‘middle’. However, Plains Remo appears to be generalizing Conjugation-I and many intransitive verbs are now part of this class. The seventeen Class-II verbs in Plains Remo are all Class-II in Hill Remo except *rag-* ‘tear (cloth or paper)’ which is Class-I.

The two sets of inflections in Remo are offered in (106)

(106)	Plains Remo	Hill Remo	Plains Remo	Hill Remo
	PST-I	<i>ɔʔ</i>	PST-II	<i>-ga, -gi-, Ø</i>
	NPST-I	<i>-to</i>	NPST-II	<i>-ta</i>

The system in Gutob, the language most closely related to Remo, has a similar tense-marking system, where the PST-II is *-gu* (called middle past by Griffiths, this volume) and the PST-I *-oʔ* (called active past by Griffiths, this volume). There is a potentially cognate element in Gutob relating to the NPST.I, but in Gutob it appears to have a customary or habitual aspect meaning, at least in some contexts. It is possible that the *-o* vocalism in the NPST.I is a secondary development in Remo, spread from the *o-/ɔ* vocalism of the past form in the conjugation. The NPST.II form seems to be the older one from a comparative Munda perspective (Anderson 2001, 2004, 2007).

Some examples of the various tense-markers in Remo are offered in (107).

- (107) Class-I
- | | |
|--|---|
| <i>lar-oʔ-nay</i> ‘we two winnowed’ | <i>lar-t-ij</i> ‘I winnow’ |
| <i>lar-to-pa</i> ‘you two winnow’ | <i>lar-oʔ-t-pe</i> ‘you (pl) have winnowed’ |
| <i>lar-oʔ-dij-nij</i> ‘I had winnowed’ | <i>lar-oʔ-dik-pa</i> ‘you two had winnowed’ |
| <i>lar-oʔ-dig-a</i> ‘she/he’ | |
| <i>la-lar-djen-t-ij</i> ‘I am cutting’ | <i>la-lar-djen-t-a</i> ‘she/he is cutting’ |
| <i>la-lar-djen-no</i> ‘you were cutting’ | <i>la-lar-goy-t-nay</i> ‘we two can cut’ |
| <i>la-lar-goy-dij-nij</i> ‘I could cut’ | <i>la-lar-goy-dik-pa</i> ‘we two could cut’ |
- (Fernandez 1968:92)
- Class-II (Ø-past)
- | | |
|--|---|
| <i>nsay-no</i> ‘you refused’ | <i>nsay-ga</i> ‘she/he refused’ |
| <i>nsay-t-ij</i> ‘I refuse’ | <i>nsay-t-pa</i> ‘you two refuse’ |
| <i>nsay-ta</i> ‘she/he refuses’ | <i>nsay-gi-t-nay</i> ‘we have refused’ |
| <i>nsay-dij-nij</i> ‘I had refused’ | |
| <i>nsay-dik-pe</i> ‘you (all)’ | <i>nsay-dig-a</i> ‘she/he had refused’ |
| <i>nsay-djen-t-nay</i> ‘we are refusing’ | <i>nsay-djen-nij</i> ‘I was refusing’ |
| <i>nsay-goy-t-no</i> ‘you can refuse’ | <i>nsay-goy-dij-nay</i> ‘we two could refuse’ |
- (Fernandez 1968:92)

Note that in non-third person subject contexts, the past tense form of Class-II appears as *-gi*, and does so for all persons with certain verbs.

Some more examples demonstrating the use of various Remo tense morphemes are offered below.

- (108) (a) *remo soka oluʔə-ta* (b) *remo tuyboʔə-da*
 man shirt wear-NPST.II man tie.turban- NPST.II
 ‘The man is wearing the shirt.’ ‘The man is wearing a head-wrap.’
 [SDM] [SDM]
- (c) *gusoʔ tuwəʔg-bo dʔiʔyi-da* (d) *gusoʔ roʔ-bo laʔig[i]-da*
 dog ground-LOC lie- NPST.II dog road-LOC sit-NPST.II
 ‘The dog is lying on the ground.’ ‘The dog is sitting on the street.’
 [SDM]
- (e) *ɲgere toŋo-da*
 young.man stand-NPST.II
 ‘The young man/bachelor is standing.’ [SDM]
- (f) *niŋ laj-t-iŋ* (g) *niŋ to[ʔo]n-t-iŋ*
 I sit- NPST.II-1 I stand- NPST.II-1
 ‘I am sitting.’ ‘I stand up.’ [SDM]
- (h) *naŋbaj ton-te-no[ŋ]*
 we.2DL=2.people stand-NPST.II-1DL
 ‘We (two) are standing.’
- (i) *naŋ bajo ton-te-no[ŋ]*
 we.2DL 2.people stand-NPST.II-1DL
 ‘We (two) are standing.’ [SDM]
- (j) *nāj glero toŋə-te-naŋ* (k) *niŋ nsuʔaʔ sum-t-iŋ*
 we.PL all stand-NPST.II-1PL I banana eat-NPST.II-1
 ‘We all are standing.’ [SDM] ‘I eat bananas.’ [SDM]
- (l) *bondagada-na remo-le uli sum-to*
 Bondagada-GEN person-PL mango eat-NPST.I
 ‘The people of Bondagada eat mango[es].’ [SDM]
- (m) *bondagada-na remo-le... k[i]jaŋ*
 Bondagada-GEN person-PL cooked.rice
 ‘The people of Bondagada eat cooked rice.’
- (n) *sum-to nd[ə]rā uʔu-to*
 eat- NPST.I morning.gruel drink-NPST.I
 ‘(they) drink the rice gruel.’
- (o) *niŋ tugola toŋ-kə-niŋ* (p) *niŋ tuwela lay-ki-niŋ*
 I yesterday stand-T/A-1 I yesterday sit-T/A-1
 ‘yesterday I was standing.’ ‘I was sitting yesterday.’ [SDM]
- (q) *niŋ tuwela a-mənda remo jul-oʔ-niŋ*
 I yesterday OBJ-one man see-PST.I-1
 ‘I saw that one man yesterday.’
- (r) *niŋ nsuʔaʔ tugola... sum-oʔ-niŋ*
 I banana yesterday... eat-PST.I-1
 ‘yesterday I ate/was eating a banana.’

Note that the second conjugation past may, in certain instances, be used without the final vowel.

- (109) (a) *munaʔ-bai gisag sun-ɔʔ ki kɔʔn dɔkra sag*
 big-DET monkey say-PST.I QUOT DEIC man come:PST.II
 ‘The big monkey said: “This one is not the wife, the man has come.”’
 (Bhattacharya 1968:148)
- (b) *dɔkra a-dɔkri sun-ɔʔ ɔy saʔmele sɔrla:ye*
 man OBJ-woman say-PST.I DISC mandeya.corn cucumber
bɔda-le gaʔaŋ guly bulu-g diʔ-ta
 boda.grain-PL beans all/many ripen-PST.II AUX-NPST
 ‘The man said to his wife: “Dear, the mandeya, cucumbers, boda grains,
 and jurunga beans, all have matured (in our field).”’ (Bhattacharya
 1968:149)

Regarding the distribution of stems into the conjugations in Hill Remo, based on their English meanings, which is, of course, not necessarily indicative of anything *per se* about the semantics of the Remo forms themselves, there does not appear to be anything semantically coherent about them. The Plains Remo Class-II forms do not share that much, except that almost all of them are canonical one-argument verbs, which is not that surprising given that it is assumed here that Class-II was originally intransitive (or ‘inactive’, middle, stative, etc.).

In Hill Remo, there is no obvious connection among all the single argument forms in Class-I, but there are some clear groups within the relatively large set. For example, a large number of bodily functions are Class-I: ‘cough’, ‘hiccough’, ‘vomit’, ‘fart’, ‘piss’, ‘shit’, ‘belch’, ‘have wrinkles’; verbs of oral action: ‘say’, ‘chirp’, ‘whistle’, ‘chat’, ‘bark’, ‘blow on fire’, ‘suck blood’, ‘palm-read’. All of these might be subsumed under some kind of ‘active’ or agent-initiated or -involved action.

However, there are also some stative forms: ‘be sticky’, ‘be naked’, ‘be spicy’, ‘be taboo’; various active verbs of manipulating vegetable matter: ‘cut wood’, ‘spin fibre on thigh’, ‘spread cloth’, ‘place yarn on warp of loom’, ‘pound fibre’, ‘ball yarn’, ‘tie wood’, ‘wear on upper arm’, ‘bend wood’, ‘cut paper’, ‘graze’, ‘wind yarn on frame’; certain forms with incorporated nouns ‘cut off head’, ‘squeeze out water’, and miscellaneous others ‘rain’, ‘run’, ‘arrive’, ‘shave self’ (if anything a canonical middle form from a Munda or South Asian perspective!), ‘dam stream’, or ‘spoil’.

A number of Class-I stems of Plains Remo are Class-II in Hill Remo. This set includes both ‘expected’ intransitives and various unexpected semantic transitives. Frustratingly, there doesn’t appear to be any systematicity *per se* across the group of ‘unexpected’ Class-II forms, that is, those that are semantically transitive in Hill Remo. There are a number of verbs of aggressive or vigorous physical action: ‘fight’, ‘beat’, ‘rub’, ‘beat on breast’, ‘jump across’, ‘jump over’, ‘climb’, ‘give birth’, ‘scratch’, ‘wash’, ‘bathe’, ‘dig’, ‘wrap with cloth’; other verbs include: ‘get’, ‘know’, ‘love’, ‘fear’, ‘desire’, ‘win’, ‘heat’, ‘fish with trap’, ‘recognize’, ‘hide’, ‘wear on head’, ‘marry’, ‘wear’, and ‘threaten’. Some might be reconcilable with original notions of ‘middleness’ (as argued for Gutob and Kharia), but it is clear that no such categorization is operative, synchronically, in Remo. In addition, a small number of Hill Remo roots may be used in this labile manner with either conjugation (110). Exactly one, perhaps two appear to be used so in Plains Remo (111). For a list of verbs falling into these two inflectional classes in the two different attested Remo dialects, see 6.3 below.

(110) ‘Labile’ roots in Hill Remo

<i>jul</i> (I) ‘see, look’	(II) ‘hang up’
<i>lug</i> (I) ‘dig’	(II) ‘set (of sun)’
<i>ruŋ</i> (I) ‘bring’	(II) ‘take’
<i>tul</i> (I) ‘throw’; ‘fix in ground’	(II) ‘hang, tie’
<i>tur</i> (I) ‘search for’	(II) ‘sprout’

(111) Plains Remo

<i>wai</i> (I) ‘call’	(II) ‘marry’
NB: <i>raŋ</i> (I) ‘cook’	<i>raŋdal</i> ‘cook’ (II)

Without a doubt it is pure coincidence that all such roots in Hill Remo are monosyllabic and have the vowel *-u-*. Some stems may be used variably or differently by different speakers, for example, *nsaŋ* is listed as class-II by (Fernandez 1968:92) but as class-I (Fernandez 1968:134), where it has an unmarked past allomorph (or Ø-past) anyway.

Also, a number of stems show a kind of mixed inflection in Plains Remo *-o?* < I / *-ta* < II, which suggests that these are being generalized. Note that the multiple tense-marked (‘present perfect’) forms show *-o?ta* and *-gəta* not **-o?to*, so this may have something to do with the cause of this change.

Although the extent to which this active process has not been investigated, there appears to be in Remo the possibility to make a passive off of a first conjugation verb by inflecting it in the second conjugation. There is but one, albeit clear, example of this in our corpus.

- (112) *ma?* *mundi-bə* *səb-ga*
vegetable.curry earthen.pot-LOC hold-PST.II
‘The veggy curry was held in an earthen pot.’ (Bhattacharya 1968:148)

Morphophonologically speaking, the tense-markers show a limited degree of assimilation or ‘harmony’ processes. These can be archiphonemically abstracted:

- (113)
- | | | |
|------|-------|-------------------------|
| | I | II |
| NPST | -tO | -tA |
| PST | (-o?) | -gI |
| PRF | -kI | -kI (Fernandez 1968:28) |

Note that the perfect (see section 3.2.3) form is undifferentiated for the two classes, and probably represents a more recent grammaticalization of an auxiliary verb construction (Anderson 2006). The processes of assimilation or harmony operative are listed below.

- (114)
- | | | |
|------------|---|--|
| <i>O/A</i> | > | Ø/___V |
| | | <i>e/___Ce</i> |
| | | <i>o/___Co</i> |
| | | <i>a/___Ca</i> |
| <i>O</i> | > | <i>o/___#</i> |
| <i>A</i> | > | <i>a/___#</i> |
| <i>-gI</i> | > | Ø/___1st/2nd marker(+V??) |
| | | <i>a/___#</i> (= Ø-3rd ending (= -a probably)) |
| | | <i>i/___tO</i> [npst]-V |
| | | <i>e/___Ce</i> |

*ol*__-Co
*al*__-Ca
 -kI > identical to following vowel (*ile/ola*)
 -tV = NPST -gV PAST and -kI

Importantly, the vowel qualities can show a rightward as well as a leftward spread, as in *gaj-go-to-no*, where the -o- vocalism in the two tense-markers is determined by the -o in the subject marker.

In terms of relative positionality of the tense (T), aspect (A), quasi-fused modal auxiliaries (M) and subject markers (S), the following Verb Templates in Remo are found (in non-negative conjugations; on negatives, see section 3.2.9).

- (115) V-T-S V-A-S V-T-A-S V-M-T-S
 V-M-A V-M-T-A-S

Examples of these include those in (116).

(116) Conjugation-I

baɖ-oʔ-no ‘you slapped’
ba-te-pe ‘you all slap’
baɖ-oʔ-tə-pa ‘you two have slapped’
baɖ-oʔ-ki-nij ‘I had slapped’
bə-ba-goŋ-tə-nay ‘we can slap’
baɖ-oʔ ‘she/he, they slapped’
ba-to ‘they, she/he slap’
baɖ-oʔ-ta ‘she/he has, they have slapped’
baɖ-oʔ-kə-ga ‘she/he has, they had slapped’ (Fernandez 1968:52–53)

Conjugation-II

<i>gay-g-no</i> ‘you entered’	<i>gay-g-a</i> ‘she/he, they know’
<i>gay-tə-naj</i> ‘we 2 enter’	<i>gay-to-no</i> ‘you enter’
<i>gay-ta</i> ‘she/he, they enter’	
<i>gay-gi-t-ij</i> ‘I have entered’	<i>gay-go-to-no</i> ‘you have entered’
<i>gay-gə-ta</i> ‘she/he has, they’ve entered’	
<i>gay-ko-no</i> ‘you had entered’	<i>gay-ki-gə</i> ‘she/he has, they had entered’

(Fernandez 1968:52–53)

As for the inflectional pattern of the ‘completive’ and other auxiliary verb constructions, see section 3.2.12.

3.2.4 Aspect

Remo has two or three morphological, inflectionally encoded aspectual categories (or tense/aspect combined), the perfect and the pluperfect, as well as the progressive. The perfect or present perfect consists of the conjugationally appropriate past tense-marker followed by the non-past marker, followed by a person marker if relevant, that is *-oʔ/gI-tV*; note that the third past form precedes the non-past marker in the perfect but follows it in the pluperfect. Remo is typologically unusual in having a unitary, unanalyzable pluperfect (or past perfect) suffix but a transparent, componential (present) perfect element *-kI-* for second conjugation verbs. First conjugation verbs, on the other hand, also have the past tense-marker preceding the pluperfect marker, namely *-oʔ-kI* (to which is added the third person

past marker *-ga* with third person subjects). Compare the forms in the following examples:

- (117) (a) *gay-gi-t-ij* (b) *gay-go-to-no*
 enter-PST.II-NPST-1 enter-PST.II-NPST-2
 ‘I have entered.’ ‘You have entered.’
- (c) *gay-gə-ta*
 enter-PST.II-NPST.II
 ‘She/he has, they have entered.’ (Fernandez 1968:52–53)
- (d) *gay-ko-no* (e) *gay-ke-pe*
 enter-PLUP-2 enter-PLUP-2PL
 ‘You had entered.’ ‘You all had entered.’
- (f) *gay-ki-gə*
 enter-PLUP-PST.II.3
 ‘She/he has, they had entered.’ (Fernandez 1968:52–53)
- (g) *baɖ-oʔ-tə-naŋ* (h) *baɖ-oʔ-ta*
 slap-PST.I-NPST-1DL slap-PST.I-NPST.II
 ‘We 2 have slapped.’ ‘He has slapped.’ (Fernandez 1968:52–53)
- (i) *baɖ-oʔ-ki-nij* (j) *baɖ-oʔ-kə-ga*
 slap-PST.I-PLUP-1 slap-PST.I-PLUP-PST.II.3
 ‘I had slapped.’ ‘He had slapped.’ (Fernandez 1968:52–53)

Progressive forms in Remo are encoded by the suffix *-qen-*, originally an auxiliary verb fused into a morphological complex. The verb stem, if mono-syllabic, appears in the reduplicated form. Note that the progressive in Remo is inherently unmarked for tense and takes the second conjugation non-past marker *-ta* to make a present progressive and, importantly, the *-ki-* and *-kə-* pluperfect marker (for both conjugations) to mark past progressives or imperfects.

The degree of bondedness between the progressive auxiliary and the lexical verb, that is, whether this remains a synchronically bi-partite auxiliary verb construction or has been unverbated into a larger morphological complex, remains a subject to be resolved by future research.¹

- (118) (a) *bə-ba=qen-ki-nij* (b) *gə-gay=qen-ki-nij*
 RDPL-slap-PROG-PRF-1 RDPL-die- PROG-PRF-1
 ‘I was slapping.’ ‘I was dying.’ (Fernandez 1968:54)
- (c) *bə-ba=qen-kə-ga* (d) *gə-gay=qen-kə-ga*
 RDPL-slap-PROG-PRF-PST.II.3 RDPL-die- PROG-PRF-PST.II.3
 ‘She/he was slapping.’ ‘She/he was slapping.’ (Fernandez 1968:54)
- (e) *bə-ba=qen-gi-tij* (f) *gə-gay=qen-gi-tij*
 RDPL-slap-PROG-PST.II-NPST-1 RDPL-die-PROG-PST.II-NPST-1
 ‘I have been slapping.’ ‘I have been dying.’ (Fernandez 1968:54)
- (g) *nij weka nsuɾaʔ sũ-sum [d]jem-t-ij*
 I now banana REDPL:eat AUX-NPST-1
 ‘I am eating a banana now.’ [SDM]

It is possible for the progressive auxiliary to appear without a tense-marker but with a subject marker in non-past formations.

- (119) *qɔkri jul-ɔ-seʔta ma qeɔem qen-nɔ qɔkra qɔ seʔta sun-ɔʔ*
 woman see-PST.I-SS what REDPL:do AUX-2 man QUOT SS say-PST.I
 ‘Seeing (him) the wife (said), “What were you doing, husband?”’ (Bhattacharya 1968:148)

A reduplicated (or unmarked?) verb stem alone can also function as a finite verb form in Remo, marking a kind of past habitual tense/aspect form.

- (120) (a) *gitin dɔkra biri-bɔ piriʔ kukum gisakʔ gubuʔ*
 that.CLOSE old.man forest-LOC bird peacock monkey pig
ɪnɔ-ɪnɔ-sa isa jag-bar ui.
 COPY.chase-PURP daily watch-AUGM go
 ‘The man would go there daily to watch and drive away the birds, peacocks, monkeys, and pigs.’ (Bhattacharya 1968:148)
- (b) *gitin qɔkri mai qɔkra-palay kiaɲ isa*
 that.CLOSE old.woman 3sg man-FOR cooked.rice daily
ruɲ bebeʔ
 bring REDPL.give
 ‘The woman used to bring food for her husband daily.’ (Bhattacharya 1968:148)

There is one progressive formation in Plains Remo that bears mention here. This is the aspectual clitic *-ni* perhaps borrowed from Desia. Bizarrely, it seems to function as a post-inflectional clitic which attaches to a fully finite (tense- and subject-marked) *past* tense form of a verb (appropriate to the inflectional class of the Remo stem). Some examples of this bizarre mixed construction are offered in (121). Similar formations are found in Gorum (Anderson and Rau, this volume).

- (121) (a) *ʔi-ga-ni* (b) *sum-ɔʔ-niɲ-ni* (c) *tɔʔ-ga-ni*
 go-PST.II.3-PROG eat-PST.I-1-PROG become.LOOSE-PST.II.3-PROG
 ‘He is going.’ ‘I am eating.’ ‘It is becoming loose.’
 (Bhattacharya 1968:79)

3.2.5 Mood

There are three basic modal categories morphologically encodable within the Remo verb (in addition there is a range of quasi-bound modal auxiliary formations discussed in section 3.2.12). The first of these is the imperative formation. For first conjugation verbs, this is the bare stem to which is added person/number markers except for the second singular form (which is \emptyset in this conjugation).² Third person imperatives are marked by the suffix *-aj* in both Hill Remo and Plains Remo. Second conjugation forms have an *-[l]a-* imperative form added to most persons including second singular.

- (122) Hill Remo
- | | | | |
|---|--|---|--------------------------------|
| (a) <i>sum-niɲ</i>
eat-1
‘Let me eat.’ | (b) <i>sum-naɲ</i>
eat-1DL
‘Let us 2 eat.’ | (c) <i>sum-naj</i>
eat-1PL
‘Let us all eat.’ | (d) <i>sum</i>
eat
‘eat’ |
| (e) <i>sum-aj</i>
eat-3
‘let him/her eat’ | (f) <i>sum-pa</i>
eat-2DL
‘eat you 2’ | (g) <i>sum-pe</i>
eat-2PL
‘eat(PL)’ (Fernandez 1968:59) | |

(123) Hill Remo

- | | | |
|--|---|--|
| (a) <i>qayks-a-pa</i>
climb-IMP-2DL
'climb you 2' | (b) <i>qayks-a-pe</i>
climb.IMP-2PL
'climb(PL)' | (c) <i>qayks-a-nij</i>
climb-IMP-1
'let me climb.' |
| (d) <i>qayks-a-naj</i>
climb-IMP-1DL
'let us 2 climb.' | (e) <i>qayks-a-naj</i>
climb-1PL
'let's all climb.' | (f) <i>qayks-a</i>
climb-.IMP
'climb' |
| (g) <i>qayks-aj</i>
climb-3.IMP 'let him/her climb' (Fernandez 1968:59) | | |

Plains Remo, on the other hand, has an imperative suffix in *-le* that is used in both conjugations. Sometimes this appears as *-la* as in *Gta?* or as *-a* as in *Gutob*. Note that in Plains Remo, first person imperative forms may have a deontic modal nuance.

(124) Plains Remo

- | | |
|--|---|
| (a) <i>baṭa qem-naṅ</i>
divide AUX-1DL
'We 2 should divide.' | (b) <i>la-le</i>
go-IMP
'go!' |
| (c) <i>sum-le</i>
eat-IMP
'eat!' | (d) <i>sum-naṅ</i>
eat-1DL
'we should eat.' (Fernandez 1968:59) |

Another modal form in Remo is the subjunctive in *-(l)ai* which is probably cognate with both *Gutob -e* and *Gta? -le?*.

- | | | | | | | |
|-------|-------------|--------------|--------------------|------------------|--------------------|------------------|
| (125) | | <i>Gutob</i> | <i>Plains Remo</i> | <i>Hill Remo</i> | <i>Plains Gta?</i> | <i>Hill Gta?</i> |
| | Subjunctive | <i>-e</i> | <i>-(l)ai</i> | <i>-(l)ai</i> | <i>-le?</i> | <i>-le?</i> |

The conditional in Remo is marked by the clitic *-na* which generally allows for no person/number indexing. This attaches to the past form of first conjugation stems or to the bare stem forms instead for Conjugation-II stems (or the Ø-allomorph of the past); a similar pattern is seen with the completive auxiliary as well. This *-na* also serves to mark different subject structures within the system of switch reference active in narrative genres.

- | | | | |
|---|---|---|---|
| (126) (a) <i>ui-na</i>
go- COND
'if I go' | (b) <i>qem-ᵛ-na</i>
do-PST.I-COND
'if you do' | (c) <i>nᵛ raja qeṅ-na nij nᵛn mᵛntri qem-a-nij</i>
you king COP-COND I you:GEN minister COP-IMP-1
'When you will be a king let me be your minister.' (Bhattacharya 1968:63) | (d) <i>sa?kur -ren-ta kiyayṅ a-sa?kur-ᵛ-na</i>
stir.with.ladle -AUX-NPST cooked.rice NEG-stir.with.ladle-PST.I-COND
<i>geb-ᵛ? suṅ-tᵛ?</i>
burn-CV AUX-NPST?
'If you do not stir, the rice will be fully burnt.' (Bhattacharya 1968:126) |
|---|---|---|---|

3.2.6 Orientation/directionality

Outside of auxiliary structures where such categories are expressed to a minor degree, directionality or orientation plays a quite limited role in Remo. The one

exception is the stem *ruŋ* ‘carry’ which when inflected in the first conjugation (*ruŋ-oʔ* ‘she/he brought’) has a cislocative/ventive speaker-deictic meaning of ‘bring’ and in the second conjugation (*ruŋ-ga* ‘she/he took away’) has a deictic hearer orientation or translocative meaning. A similar system is found in Gutob.

The auxiliary *ui* ‘go’ has numerous functions in Remo. One such function is in combination with other verbs of motion to form a translocative/-itive formation. Note that the lexical verb in such formations occurs in the converb form that is the same as the past tense.

- (127) *a-dʒkra gisaʔg-e jul-ɔʔ-seʔta saʔay-bar-sa*
 OBJ-man monkey-PL see-PST.I-SS scatter-AUGM=PURP
ur-ɔʔ ui-ga tɔ
 run-CV go-PST.II NARR.PRTCL
 ‘The monkeys saw the man and ran off to scatter his food.’ (Bhattacharya 1968:148)

3.2.7 Voice

Like many Munda languages, Remo has two morphologically encodable voice categories involved in stem formation. The first is the archaic Munda causative prefix, realized in Remo as *o-* (or *ə-*). A very small number of verb stems appear in a reduplicated form with the causative prefix. Note that the causative curiously appears with modal auxiliaries rather than the lexical verb over which it has semantic scope in certain AVCs (see 3.2.12), suggesting it has some kind of phrasal clitic status morphophonologically. The causative is highly productive in Remo and can be added to practically any verb stem.

- (128) *o-gi-geb* ‘caus-REDPL:heat’ vs. *o-log* ‘pluck’ (Fernandez 1968:40)

The reciprocal, on the other hand, appears as an infix *-n* and appears in a small number of lexical items. It seems likely that the Remo reciprocal is cognate with the infix part of the prefix-cum-infix reciprocal found in Gtaʔ, but whether this is, in fact, cognate with the Kherwarian infix *-p-* reciprocal (via **-m-*) remains a subject for future research. Note that both the causative and the reciprocal can appear together in a single stem (cf. ‘cause to fight’).

- (129) (a) *əsu-n-ob* ‘cause to embrace each other’ < *sob* ‘catch’ (Fernandez 1968:40)
 (b) *o-bu-n-ug* < *bug* ‘cause to fight’ *bug* > *bunug* ‘fight’ (Fernandez 1968:40)
 (c) *naŋ ɔ-s/ɔn/ɔbnaŋ*
 we.DL CAUS-hold/RECIP/-1DL
 ‘We two held each other.’ (Bhattacharya 1968:25)

3.2.8 Finiteness

Verb phrases can be made into modifiers by adding the adjectival suffix *-bay* and preposing the verb phrase before the noun that it modifies, forming a relative-clause type structure. For more on this, see the Syntax section below. At least four and probably more types of non-finite or non-final forms are found in Remo. Some of these have other functions in the grammar of the language. Thus, reduplicated verb stems are found in number of contexts in Remo. One area where they constitute an

obligatory non-finite verb form is with mono-syllabic stems of lexical verb in certain auxiliary verb constructions, for example, the progressive formation.

- (130) (a) *niŋ [n]sura? sūsum [d]em-t-iŋ*
 I banana REDPL:eat AUX-NPST-I
 ‘I am eating a banana.’
 (b) *niŋ nsura? sūsum den-t-iŋ*
 I banana REDPL:eat AUX-NPST-I
 ‘I am eating a banana.’ [SDM]

Note that with loan stems in constructions licensing a reduplicated stem allomorph, instead of reduplication, a stem augment *-bar-* is added.

- (131) *gisa?g-e dɔgɔy-bar a-gɔŋ-ga*
 monkey-PL jump-AUGM NEG-CAP-PST.II
 ‘The monkeys were unable to jump out.’ (Bhattacharya 1968:148)

Another non-finite form is the past tense converb or participle form that is used in various auxiliary structures as well as quasi-serialized formations (see 3.2.12). Originally this was restricted to first conjugation verbs, but this has been extending into all verbs over the past forty years or so.

- (132) (a) *remo diredire uriŋ[g]-o[ʔ] wi-en-ta*
 man slowly walk-CV GO-AUX-NPST.II
 ‘The man is slowly walking.’ [SDM]
 (b) *remo suri?suri? uri-oʔ-jen-ta*
 man quickly walk-CV-AUX-NPST.II
 ‘The man is walking quickly.’
 (c) *remo suri?suri? uriŋ-o[ʔ] o-jen-ta*
 man quickly walk-CV AUX-AUX-NPST.I
 ‘The man is walking quickly.’ [SDM]
 (d) *saʔkur renta kiyay a-saʔkur-ɔna*
 stir.with.ladle-AUX-NPST cooked.rice NEG-stir.with.ladle-PST-COND
geb-ɔʔ suŋ-tɔʔ
 burn-CV AUX-NPST?
 ‘If you do not stir, the rice will be fully burnt.’ (Bhattacharya 1968:126)

There are two kinds of conjunctive elements in Remo as well, one of which also functions as a conditional marker but is found in mainly complex sentences which do not share subjects across the clauses, and the other of which just seems to string together predicates across a complex sentence, albeit with a strong tendency to patterning with same subject structures. Thus, these are largely addressed in section 4.2.3.

However, there are some uses of *-ŋa* which neither seem to be different subject coordinative forms nor conditional subordinate clauses, here functioning as a kind of purposive marker, but the details of its use in this construction remain to be worked out.

- (133) *sunu?bo? tug-oʔ-ta susugbo?-ŋa*
 comb:head tie-PST.I-NPST comb:head-DEP
 ‘He has tied a comb (to his waist) for hair-combing.’ (Fernandez 1968:61)

Finally, a bare stem form of a verb (in Conjugation-II) may serve as a non-finite form in certain complex predicate structures (see section 3.2.12 for examples).

3.2.9 Negation

Negative verb forms in Remo are mainly formed by adding the negative prefix *a[r]*- to the verb template. This is true of most past tense forms (some with Ø-past allomorphs), non-past, perfect, pluperfect, and progressive. Remo stands apart from its close sister language Gutob in this way, which has among the world's most unusual systems of negation (Anderson 2007). Note that the second conjugation past tense negative has a zero-allomorph of the past with first singular subjects with certain stems.

- (134) (a) *guso?* *roʃ-bo* *dʒi?* *ar-iyi-da*
 dog road-LOC lie NEG-AUX-NPST
 'The dog is not lying on the street.' [SDM]
- (b) *niʃ* *dʒa?* *a-u?**t-ij*
 I water NEG-drink-NPST-1
 'I am not drinking water.' [SDM]
- (c) *a-sum-o?* (d) *a-sa-ga* (e) *a-sak-niʃ*
 NEG-eat-PST.I NEG-come-PST.3 NEG-come-1
 'She/he didn't eat' 'She/he didn't come' 'I didn't come'
 (Fernandez 1968:56)
- (f) *a-sum-t-ij* (g) *a-sum-to* (h) *a-sap-ta*
 NEG-eat-NPST-1 NEG-eat-NPST.I NEG-come-NPST.II
 'I don't eat.' 'She/he doesn't eat.' 'He doesn't come.'
 (Fernandez 1968:56)
- (i) *a-sum-o?**t-ij* (j) *a-sab-o?**t-ij*
 NEG-eat-PST.I-NPST-1 NEG-come-PST.I-NPST.II
 'I have not eaten.' 'He has not come.' (Fernandez 1968:56)
- (k) *a-sum-o?**k-ga* (l) *a-sap-ki-niʃ*
 NEG-eat-PST.I-PLUP-PST.3 NEG-come-PLUP-1
 'She/he had not eaten.' 'I had not come.' (Fernandez 1968:57)
- (m) *a-sum-dʒen-t-ij* (n) *a-sap-dʒen-ta*
 NEG-eat-PROG-NPST-1 NEG-come-PROG-NPST
 'I am not eating.' 'I am not coming.' (Fernandez 1968:57)
- (o) *a-sum-dʒen-kə-ga* (p) *a-sap-dʒen-ki-niʃ*
 NEG-eat-PROG-PLUP-PST.3 NEG-come-PROG-PLUP-1
 'She/he was not eating.' 'I was not coming.' (Fernandez 1968:57)
- (q) *nə-na* *pʁsɔlbai* *bɔb-le* *gisi* *a-dʒem-ɔ-na*
 you-GEN clean:ADJ head-EMPH louse NEG-appear-PST.I-COND
 'Lice would not have appeared if your head was clean.' (Bhattacharya
 1968:63)
- (r) *niʃ* *nsuʒa?* *a-sum-t-ij*
 I banana NEG-eat-NPST-1
 'I am not eating a banana.' [SDM]
- (s) *niʃ* *tugola* *da?* *a-u-[o]?**t-ij*
 I yesterday water NEG-drink-PST.I-NPST-1
 'Yesterday I was not drinking water.' [SDM]

With modal auxiliary formations, the negative appears on the modal auxiliary over which it has semantic scope. The lexical verb appears in a reduplicated form, as is required by these particular auxiliary verb constructions (see section 3.2.12).

- (135) (a) *susum a-goŋ-ta* (b) *səsap a-goŋ-ta*
 REDPL:eat NEG-CAP-NPST.II REDPL:come NEG-CAP-NPST.II
 ‘She/he can’t eat.’ ‘I can’t come.’ (Fernandez 1968:57)
- (c) *susum a-goŋ-ki-niŋ* (d) *səsap a-goŋ-kə-ga*
 REDPL:eat NEG-CAP-NEG.PST-I REDPL:come NEG-CAP-NEG.PST-PST.3
 ‘I couldn’t eat.’ ‘He couldn’t come.’ (Fernandez 1968:57)

The most curious aspect of negative marking in the Remo verbal system is the formation of negative imperatives or prohibitives. This system is in part shared by Gutob where the system is infinitely more complex (but may have triggered the shift in the system attested in this language), and also in Gta?.

- | | | | | | | |
|-------|-----------|---------|-------------|-------------|-------------|-----------|
| (136) | | Gutob | Plains Remo | Hill Remo | Plains Gta? | Hill Gta? |
| | PST-I | ar-X-tə | a-X-tV-ɔ? | a-X-ɔ? | a-X-ke | a-X-ti |
| | PST-II | ar-X-tə | a-X-ga/gi-Ø | a-X-ga/gi/Ø | a-X-ke | a-X-ti |
| | PROHIB-I | ar-X-gu | a-X-ɔ? | a-X-ɔ? | a-X-ge | a-X-gi |
| | PROHIB-II | ar-X-ɔ? | a-X- ga/gi | a-X-ga/gi | a-X-ge | a-X-gi |

That is, prohibitives in Remo have the formal structure of negative past forms, with which they are, in fact, ambiguous. Note that both Conjugation-I and Conjugation-II verbs take the appropriate respective past tense forms. Compare the forms in (137) with those in (134) above. For more on this, see Anderson (2007).

- (137) (a) *a-sum-oʔ-niŋ* (b) *a-sum-oʔ*
 NEG-eat-PST.I-1 NEG-eat-PST.I
 ‘don’t let me eat’ ‘don’t eat’
- (c) *a-ɖayk-gi-niŋ* (d) *a-ɖayk-ga*
 NEG-climb-PST.II-1 NEG-climb-PST.II
 ‘don’t let me climb’ ‘don’t climb’ (Fernandez 1968:59)

3.2.10 Derivation

Derivation is not an extensively used process in the Remo verbal system. Causative and reciprocal formations, as well as noun incorporation are semi-productive or lexicalized means of deriving verb stems, each discussed in relevant sections above and below. There are also lexicalized instances of what appears to be deverbal nominal forms functioning as verb stems in Remo as well, for example, note *sinig* can be a verb ‘to fart’ < \sqrt{sig} (Fernandez 1968:40), where a reciprocal meaning seems unlikely. In addition, a small number of verb stems are, historically, complex, consisting of compounding or lexicalized combinations of serial verb forms (in some instances having become auxiliaries). Generally these have two stems in combination, historically, forming a complex verb stem synchronically, for example, *kukup-bam* ‘have a cough’ where *bam* means ‘get’. Sometimes the second element is opaque, *gugu-sur* ‘level ground (for planting)’ -?? (Fernandez 1968:35). Forms with lexicalized, serialized or auxiliary (or ‘light’) verbs include *bulo-suŋ* ‘boil over’, where the second element means ‘throw’ (Fernandez 1968:35) or *bana-wi[y]* ‘forget’ and *goy-iy* ‘die, be dead’ (Fernandez 1968:36) where the second element means ‘go’.

- (138) (a) *goytaŋ goy-i-da* (b) *goytaŋ kalevay goy-i-da*
 bull die-AUX-NPST.II bull black die-AUX-NPST.II
 ‘The bull died/is dead.’ ‘The black bull died/is dead.’ [SDM]

Note also *lubdaʔ-goy* ‘drown’ (Fernandez 1968:36), with an incorporated noun *daʔ* ‘water’ (see 3.2.11) and a second verb meaning ‘die’. There appears to be only one tripartite complex of this type in Remo, viz *biba-ɖem-jinl* ‘to perform marriage ceremony’, literally, ‘marry-do-win’ (Fernandez 1968:37). Of course, there may be other such formations as well.

Reduplication is also an important process in the Remo verbal system as it is in the majority of, South Munda languages. This may not be a productive stem deriving process *per se*, but elucidating this issue requires further analysis. In general, Remo ‘regular’ roots take either CV- or CVC- reduplication, the longer copy appearing when C₂ is *ŋ* in Remo, for example, *sum* > *susum* ‘eat’, *sap* > *səsap* ‘come’, *suy* > *suyɣsuy* ‘finish’/‘throw’. Many Remo verbs are transparently derived by reduplicating mono-syllabic roots (Bhattacharya 1968). The semantics associated with such verb roots are those frequently encountered in verb stem reduplication cross-linguistically, for example, iterative, frequentative, repeated action, augmented action, etc. For a number of such stems, however, the semantics behind the reduplication remain opaque.

(139)	<i>sugsugbɔʔ-</i>	to comb
	<i>kɔtkɔtɛŋ-</i>	to cackle (of a hen)
	<i>gagaʔtɔm-</i>	to open mouth, gape
	<i>susu ɖenta</i>	to massage oil on oneself
	<i>suyɣsuy ɖenta</i>	to sell
	<i>tuytuy ɖenta</i>	to wear (shoe, hat, bangle, finger-ring or ear-ring)
	<i>tɪŋtɪŋ ɖenta</i>	to pierce; to card cotton; to shoot with arrow
	<i>táktag ɖenta</i>	to peel off, crack, flay
	<i>ɖuŋɖuŋ ɖenta</i>	to flee
	<i>ɖɔɖɔ ɖenta</i>	(fire) to be kindled
	<i>sɔsɔb-</i>	to hold, catch, buy
	<i>ɖɔɖɔ ɖenta</i>	to laugh
	<i>tɔrtɔr-</i>	to tremble

There are, however, a large number of ‘irregular’ roots in Remo that either show a somewhat unexpected or unusual copy or change in the root itself, or add dummy morphs rather than some phonologically relatable copied sequence when appearing in formations that require the reduplicated stem allomorph. Examples include *oŋ* > *oʔon* ‘hear’, *raŋ* > *ŋraŋ* ‘cook; pull’, *un* > *unkeɾ* ‘transplant seedlings’ (< *keɾ* = *keɾon* ‘rice’), *yon* > *noyon* ‘chase away’, *leɖ* > *ləle* ‘squeeze’, *baɖ* > *bəba* ‘slap’.

3.2.11 Noun incorporation and combining forms

A small number of verb stems in Remo, historically, consist of a verb stem together with a so-called combining form of a noun to form a verb stem. Many South Munda languages make use of an opposition between bi-moraic or bi-syllabic free forms of nouns that contrast with monosyllabic combining forms of these same stems. Note that verb stems with incorporation can be either in Conjugation-I or Conjugation-II in Remo. Often the root of the noun is the same as the combining form, and the free form is derived through one of a range of non-productive lexical means (see section 3.1.10). Thus, the combining form of the noun is lexically determined, but as elsewhere in Munda, tends to be the basic root of the noun involved (canonically but not obligatorily of the shape CVC, less commonly CV). Sometimes (as in the second

example in (140) below), the combining form and the free form stand in a suppletive relation to one another (at least synchronically). Older incorporated forms, as in other South Munda languages, show the earlier syntactic structure of VN in the incorporative complexes (140).

- (140) *old VN compounds*
ale-ɖag 'squeeze water'
buk-taɖ 'pound fibre' (= CF of *suta* 'thread')
susug-bog 'comb' (sweep-head) (Fernandez 1968:39)

New compounds, on the other hand, reflect the more recent syntactic structure of the language, with NV order: *ɖarak-jul* 'palm-read' (Fernandez 1968:38), which belongs to Conjugation-I inflectionally. Note that verbs and nouns with the same incorporated element can be used in the same sentence in Remo.

- (141) *sunu?bo?* *tug-o?*-*ta* *susugbo?*-*ŋa*
 comb tie-PST.I-NPST.II comb:head-COND
 'He has tied a comb (to his waist) for hair-combing.' (Fernandez 1968:61)

3.2.12 Auxiliary verb constructions and other complex predicate types

Like many Munda languages, Remo makes extensive use of auxiliary verb constructions and other complex predicate structures (see Anderson 2006 for a general theoretical perspective on this). The range of functions expressed by such constructions are generally typical of areal and cross-linguistic norms, as are, for the most part, the semantic origins of the particular auxiliaries involved and the paths of grammaticalization and functional specialization that they have undergone. From an inflectional perspective, Remo shows a Munda-typical system of AUX-headed auxiliary verb constructions, with the lexical verb in various formations obligatorily appearing either in a non-finite reduplicated form, in a basic stem form (or Ø-marked dependent form), or in a past tense form (called the 'general converb' form by Griffiths (this volume) in his discussion of similar structures in the closely related Gutob).

In Remo, the completive aspect form is marked by an auxiliary verb construction using *suy-*. First conjugation verbs appear in a past tense ('converb') form, while intransitive forms appear in an unmarked form. As an etymologically transitive/Conjugation-I verb (meaning 'throw'), the verb itself takes this past form as well, when it is inflected in the past, as well as the subject markers, both of which are expected as this is embedded within an AUX-headed auxiliary verb construction (Anderson 2006). Therefore, transitive stems exhibit a pseudo-split-doubled pattern with past tense-marking seemingly on both – however, one tense suffix is licensed by the actual tense specification of the event, namely the one on the auxiliary verb itself, while the other is licensed by the larger constructional parameters, that is, it is necessitated by the use of a Conjugation-I verb with this particular auxiliary in this function, and with subject encoded on the auxiliary, while Conjugation II stems rather show the basic pattern and an unmarked stem form of the lexical verb, with all inflectional categories realized on the auxiliary verb. Compare the following forms in this regard.

- (142) (a) *baɖ-o?* *suy-o?*-*niŋ* (b) *gay suy-o?*-*niŋ*
 slap-CV-COMPL-PST.I-I-1 die COMPL-PST.I-I-1
 'I finished slapping.' 'I finished dying.' (Fernandez 1968:55)

- | | |
|--|--|
| (c) <i>baɖ-oʔ suŋ-oʔ</i>
slap-CV COMPL-PST.I
'She/he finished slapping.' | (d) <i>gay suŋ-oʔ</i>
die COMPL-PST.I
'She/he finished dying.' (Fernandez 1968:55) |
|--|--|

The progressive formation, as aforementioned, in Remo consists of CV(C) reduplication of the lexical stem, followed by the tense and person marked auxiliary verb *ɖen-*. This AVC shows the AUX-headed inflectional pattern in Remo in the positive, with tense and subject appearing on the lexical verb.

- | | |
|--|---|
| (143) (a) <i>baba=ɖen-ki-niŋ</i>
REDPL:slap=PROG-PRF-1
'I was slapping.' | (b) <i>gəgay=ɖen-ki-niŋ</i>
REDPL:die= PROG-PRF-1
'I was dying.' (Fernandez 1968:54) |
| (c) <i>baba=ɖen-kə-ga</i>
REDPL:slap=PROG-PRF-PST.II.3
'She/he was slapping.' | (d) <i>gəgay=ɖen-kə-ga</i>
REDPL:die=PROG-PRF-PST.II.3
'She/he was slapping.'
(Fernandez 1968:54) |
| (e) <i>baba=ɖen-gi-t-iŋ</i>
REDPL:slap=PROG-PST.II-NPST-1
'I have been slapping.' | (f) <i>gəgay=ɖen-gi-t-iŋ</i>
REDPL:die=PROG-PST.II-NPST-1
'I have been dying.'
(Fernandez 1968:54) |
| (g) <i>niŋ [n]suraʔ susum [ɖ]em-t-iŋ</i>
I banana REDPL:eat AUX-NPST-1
'I am eating a banana.' | |
| (h) <i>niŋ nsuraʔ susum ɖen-t-iŋ</i>
I banana REDPL:eat AUX-NPST-1
'I am eating a banana.' [SDM] | |
| (i) <i>remo suriʔsuriʔ uri-oʔ-jen-ta</i>
man quickly walk-CV-AUX-NPST.II
'The man is walking quickly.' [SDM] | |

In the negative, however, a split pattern is seen with tense and subject on the auxiliary and negative on the lexical verb. This is a relatively common split inflectional pattern in auxiliary verb constructions among the world's languages (Anderson 2006).

- | | |
|---|--|
| (144) (a) <i>a-sum-ɖen-t-iŋ</i>
NEG-eat-PROG-NPST-1
'I am not eating.' | (b) <i>a-sap-ɖen-ta</i>
NEG-COME-PROG-NPST.II
'He is not coming.'
(Fernandez 1968:57) |
| (c) <i>a-sum-ɖen-kə-ga</i>
NEG-eat-PROG-NEG.PST-PST.II.3
'She/he was not eating.' | (d) <i>a-sap-ɖen-ki-niŋ</i>
NEG-COME-PROG-NEG.PST-1
'I was not coming.'
(Fernandez 1968:57) |

The capabilitive formation in Remo is structurally similar to the progressive, that is, the lexical stem appears in a reduplicated form, and the construction exhibits the basic inflectional pattern.

- | | |
|--|---|
| (145) (a) <i>baba=goy-t-iŋ</i>
REDPL:slap=CAP-NPST-1
'I can slap.' | (b) <i>gəgay=goy-t-iŋ</i>
REDPL:die=CAP-NPST-1
'I can die.' (Fernandez 1968:53) |
|--|---|

- | | |
|--|---|
| <p>(c) <i>bāba=gōŋ-ta</i>
REDPL:slap=CAP-NPST.II
'She/he can slap.'</p> <p>(e) <i>bāba=gōŋ-ki-niŋ</i>
REDPL:slap=CAP-PRF-1
'I could slap/ have slapped.'</p> <p>(g) <i>bāba=gōŋ-kə-ga</i>
REDPL:slap=CAP-PRF-PST.II.3
'She/he could slap/ have slapped.'</p> | <p>(d) <i>gəgay=gōŋ-ta</i>
REDPL:die=CAP-NPST.II
'She/he can die.'
(Fernandez 1968:54)</p> <p>(f) <i>gəgay=gōŋ-ki-niŋ</i>
REDPL:die=CAP-PRF-1
'I could die/ have died.'
(Fernandez 1968:54)</p> <p>(h) <i>gəgay=gōŋ-kə-ga</i>
REDPL:die=CAP-PRF-PST.II.3
'She/he could die/ have died.'
(Fernandez 1968:54)</p> |
|--|---|

The desiderative formation in Remo is also, formally, similar to the capabilitive and the progressive, that is, the lexical stem is reduplicated and the construction exhibits the AUX-headed inflectional pattern.

- | | |
|--|--|
| <p>(146) (a) <i>bāba=quso?-gə-ta</i>
REDPL:slap=DESID-PST.II.3-NPST.II
'wants to slap'</p> <p>(c) <i>bāba=quso?-kə-ga</i>
REDPL:slap=DESID-PRF-PST.II.3
'wanted to slap'</p> | <p>(b) <i>gəgay=quso?-gə-ta</i>
REDPL:die=DESID-PST.II.3-NPST.II
'wants to die'
(Fernandez 1968:55)</p> <p>(d) <i>gəgay=quso?-kə-ga</i>
REDPL:die=DESID-PRF-PST.II.3
'wanted to die'
(Fernandez 1968:55)</p> |
|--|--|

Note that there is some considerable idiolectal, regional, or dialectal variation in the selection of certain functional operators/auxiliaries in Remo. Thus, the desiderative in *-quso?* is in variation with *-luḍa-ḍen* (< ? *luḍ* 'borrow, beg, want') and the progressive in *-ḍen* is in variation with *-nen* and *-e* (Fernandez 1968:41–42).

Causative forms of capabilitive marked verbs show an unusual pattern. Here, the causative attaches to the auxiliary verb, rather than lexical verb, despite the fact that the causative scope is over the lexical verb, that is, 'able to make X' not 'make able to X'.

- | | |
|---|---|
| <p>(147) (a) <i>su-sum=o-gōŋ-t-iŋ</i>
REDPL:eat=CAUS-CAP-NPST-1
'I can cause to eat.'</p> <p>(c) <i>susum=o-gōŋ-ta</i>
REDPL:eat=CAUS-CAP-NPST.II
'She/he can cause to eat.'</p> <p>(e) <i>susum=o-gōŋ-ki-niŋ</i>
REDPL:eat=CAUS-CAP-PRF-1
'I could cause to eat.'</p> <p>(g) <i>susum=o-gōŋ-kə-ga</i>
REDPL:eat=CAUS-CAP-PRF-PST.II.3
'She/he could cause to eat.'</p> | <p>(b) <i>sə-sap=o-gōŋ-t-iŋ</i>
REDPL:come=CAUS-CAP-NPST-1
'I can cause to come.'
(Fernandez 1968:57)</p> <p>(d) <i>səsap=o-gōŋ-ta</i>
REDPL:come=CAUS-CAP-NPST.II
'She/he can cause to come.'
(Fernandez 1968:57)</p> <p>(f) <i>səsap=o-gōŋ-ki-niŋ</i>
REDPL:come=CAUS-CAP-PRF-1
'I could cause to come.'
(Fernandez 1968:57)</p> <p>(h) <i>səsap=o-gōŋ-kə-ga</i>
REDPL:come=CAUS-CAP-PRF-PST.II.3
'She/he could cause to come.'
(Fernandez 1968:57)</p> |
|---|---|

Note that, as aforementioned, negative stems are not reduplicated in the progressive construction in Remo and the negative attaches to the lexical stem. Note also that the past tense is always the second conjugation form, as the auxiliary verb, historically, belonged to this class.

- (148) (a) *a-sum=d̄en-gə-ta* (b) *a-sap=d̄en-gə-ta*
 NEG:eat=PROG-PST.II-NPST.II NEG:come=PROG-PST.II-NPST.II
 ‘She/he has not been eating.’ ‘She/he has not been coming.’
 (Fernandez 1968:58)

Conversely, the lexical stems are reduplicated in the negative capability, but the negative prefix, like the causative, attaches to the auxiliary verb.

- (149) (a) *susum=a-goy-t-ij* (b) *səsap=a-goy-t-ij*
 REDPL:eat=NEG-CAP-NPST-I REDPL:come=NEG-CAP-NPST-I
 ‘I cannot eat.’ ‘I cannot come.’ (Fernandez 1968:57)
 (c) *susum=a-goy-ta* (d) *səsap=a-goy-ta*
 REDPL:eat=NEG-CAP-NPST.II REDPL:come=NEG-CAP-NPST.II
 ‘She/he cannot eat.’ ‘She/he cannot come.’
 (Fernandez 1968:57)

Further, complex auxiliary verb constructions may be created from the complete construction. For example, the auxiliary verb is reduplicated when used in a progressive formation, rather than the lexical stem. However, the form in which the lexical stem appears remains the same as in the past complete, that is, the past for Conjugation-I stems and the unmarked form for Conjugation-II stems. In other words, each auxiliary requires its complement to the left, whether it is a lexical verb or another auxiliary, to be in the appropriate form of the construction. This is true of recursively embedded auxiliary structures in English as well, for example, *I will have been seeing her*, where *will* triggers a stem form, *have* the past participle and be the *-ing* form in its progressive meaning.

- (150) (a) *baq-oʔ=suy-suy=d̄en-t-ij* (b) *gay=suy-suy=d̄en-t-ij*
 slap-CV=REDPL-COMPLT=PROG-NPST-I die=REDPL-COMPLT=PROG-NPST-I
 ‘I am finishing slapping.’ ‘I am finishing dying.’
 (Fernandez 1968:55)
 (c) *baq-oʔ=suy-suy=d̄en-ta* (d) *gay=suy-suy=d̄en-ta*
 slap-CV=REDPL-COMPLT=PROG-NPST.II die=REDPL-COMPLT=PROG-NPST.II
 ‘She/he is finishing slapping.’ ‘She/he is finishing dying.’
 (Fernandez 1968:55)
 (e) *baq-oʔ=suy-suy=d̄en-ki-nij* (f) *gay=suy-suy=d̄en-ki-nij*
 slap-CV=REDPL-COMPLT=PROG-PRF-I die=REDPL-COMPLT=PROG-PRF-I
 ‘I was finishing slapping.’ ‘I was finishing dying.’
 (Fernandez 1968:56)

The above phenomena may be summarized as follows:

- (151) AUX/Function CAP-*goy* DESID-*qusoʔ* COMPL-*suy/-suy* CONT-*d̄en*
 Verb Form +Rdpl +Rdpl *** -*suy* after past +Rdpl
 -*suy* after past before *d̄en*

Remo makes limited but definite use of structures that are commonly known as ‘light’ verbs in the literature. This consists of an uninflectable lexical stem, often

a borrowed or onomatopoeic/sound symbolic element in its base form followed by an inflectable stem, often ‘go’.

- (152) (a) *kəl piṭ iṭ-ga*
 mechanism break go-PST.II
 ‘The mechanism broke.’ (Bhattacharya 1968:149)
- (b) *ɖəkri sun-ɔṭ nɔ tɔ jul-ɔṭ-nɔ niṭ mɔri kubetek*
 woman say-PST.I you NARR.PART see-PST.I-2 I how very.much
ruṅ-niṭ nɔn mɔṭ wag iṭ-ga ki ɖəkra
 take-1 you:GEN eye crack AUX-PST.II.3 VOC man
 ‘The wife replied, “But you saw, how have I taken more? Have your eyes burst, O husband?”’ (Bhattacharya 1968:149)

As aforementioned, although not the developed system seen in its sister language Gtaṭ, something akin to a (semantically) serialized structure is found in Remo in which the first verb appears in the converb form, followed by the appropriately uninflected verb form of the second verb (including an uninflected imperative).

- (153) *suṭ guis-ɔ ɖik-lɔ ɖɔ seṭta sun-ɔṭ*
 fire kindle.fire.by.blowing-PST.I stay-EMPH QUOT SS say-PST.I
 ‘Stay here kindling fire’ (Bhattacharya 1968:149)

3.3 Expressives

Expressive forms with reduplication are characteristic of virtually all modern Munda languages, where such systems may reach advanced levels of development. Mono- and di-syllabic nouns frequently undergo suffixing echo reduplication with overwriting of V₁ (and V₂, if present) (Bhattacharya 1968).

- (154) *ṅger ṅgar* young man
saṭme siṭmi a type of corn (mandeya)
titi tata hand
kekepkakap’ (animal name)
kerɔṅ kuraṅ paddy
giṭeṭ gaṭaṭ frog
piriṭ paraṭ bird
sulup’ salap’ four-horned antelope
semug sumak’ tree (< *semuk’*)
gabuṭ gabaṭ pig
guluṭ galaṭ hare
gusɔṭ gasaṭ dog
gisiṅ gasaṅ cock

Echo reduplication can also occur with overwriting of V₂ only (Bhattacharya 1968).

- (155) *gɔytaṅ gɔytiṅ* cow
burtan burtiṅ animal (in Desiya ‘goil’)
senla senli mat
siram sirim sambar
musri musra lentil
rigḍar rigḍir sua corn

While such forms are relatively easy to find attested, the expressive and pragmatic meaning of this type of reduplication has not been investigated and unfortunately awaits further study.

4 SYNTAX

The syntax of Remo remains largely unstudied. Only the briefest of comments will be addressed here. This is especially true of complex sentence structures which have not been widely attested in extant literature on the language. Our consultant did not produce such formations in his spontaneous utterances, and the texts that exist also have only extremely limited instantiations of complex syntactic structures.

4.1 Syntax of the simple sentence

The simple sentence in Remo is largely similar to that of many other Munda languages. In terms of clausal constituent order, Remo is strongly SOV – almost no utterances or text examples in the corpus do not end in a verb, and speakers often outright reject such formations under elicitation conditions. Not only verbs but nouns and adjectives may occupy the clause-final slot licensed for predicates. In some instances, no copula is necessary.

- (156) (a) *niŋ remo* (b) *niŋ-na-yimi sukra dangara maji*
 I Remo I-GEN-name S. D. M.
 ‘I am a Remo.’ ‘My name is Sukra Dangada Maji.’ [SDM]
- (c) *remo baŋi (d̥i-ta)* (d) *d̥io sero? (d̥i-ta)*
 man good COP-NPST.II house dirty COP-NPST.II
 ‘The man is good.’ ‘The house is dirty.’ (Fernandez 1968:112)
- (e) *nuŋia? baŋibe[y]*
 coconut round
 ‘The coconut is round.’ or ‘the round coconut’ [SDM]
- (f) *goytaŋ kalevay* (g) *mayn jeri kalevay*
 bull black 3:GEN body black
 ‘The bull is black.’ or ‘the black bull’ ‘Its body’ colour is black.’
 [SDM]
- (h) *om̄d̄i-le-na gusun̄ger-e-na ninden gusu?ge*
 Omdi-PL-GEN Gusun̄ger-PL-GEN primary.clan Gusu’ge
k̄o?ne kirsani rem̄o-le
 DEIC-EMPH Kirsani person-PL
 ‘The original clan for the Omdi and Gusun̄ger is Gusu’ge; these are Kirsāni people.’ (Bhattacharya 1968:79)

Note that this lack of copula is also possible with possessive structures in the present. Thus, such formations are of the structure Noun-GEN Noun [Copula].

- (157) *m̄āy?ina ... goytaŋ gubu[?] gisiŋ gime?*
 they-GEN ... cattle pig chicken goat
 ‘They have cattle, pigs, chickens, goats.’ [SDM]

Negative possessive structures, on the other hand, in the present use the negative copula form. Possessors may be omitted contextually as well.

- (158) *nij ob-a-goŋ-t-iŋ gine anra*
 I bite-NEG-CAP-NPST-1 tooth NEG.COP
 ‘I can’t bite it, I have no teeth.’ (Fernandez 1968:114)

Locational copular structures often use the copula form *qi-* in Remo.

- (159) *nij ney-na ŋgom malkangiri distrikt-bo qi-ta*
 I we-GEN village M. D. -LOC COP-NPST.II
 ‘My/our village is (a place) in Malkangiri District.’ [SDM]

In more marked TAM forms, for example, conditional or imperative, the copula *qem* may be used in Remo.

- (160) *no raja deŋ-na nij non məntri qem-a-nij*
 you king COP-COND I you\GEN minister COP-IMP-1
 ‘When you will be a king let me be your minister.’ (Bhattacharya 1968:63)

As aforementioned, like the other Munda languages, the basic clausal constituent order of Remo is S[ubject] O[bject] V[erb]. Other permutations are permitted under certain discourse conditions, but the textually unmarked order is SOV, as it is in most, if not all, modern Munda languages.

- (161) (a) *remo nuŋia? sa?go-qa* (b) *remo a-gisiŋ ju-to*
 man coconut holding-PROG/PRES man OBJ-chicken see-NPST.I
 ‘The man is holding the coconut.’ ‘The man sees the chicken.’
 [SDM] (Fernandez 1968:124)
- (c) *nij a-nij ju-t-iŋ* (d) *ni nsuŋa? sum-t-iŋ*
 I OBJ-I see-NPST-1 I banana eat-NPST-1
 ‘I see myself.’ [SDM] ‘I eat bananas.’ (Fernandez
 1968:125)
- (e) *remo qio oroy-qen-ta* (f) *nij remo-ŋ tumo? on-t-iŋ*
 man house build-PROG-NPST.II I man-GEN mouth hear-NPST-1
 ‘The man is building a house.’ ‘I hear the man’s voice.’
 (Fernandez 1968:126)
- (g) *a-nij mabisom ruŋo?-gəta*
 OBJ-I very cold-PRF
 ‘I feel very cold.’
- (h) *a-nij kuru-gəta a-may ta?mibaŋ-gəta*
 OBJ-I hungry-PRF OBJ-she/he sneeze-PRF
 ‘I am hungry.’ ‘He is sneezing.’ (Fernandez 1968:112)
- (i) *no ero?ga kiyaj sum-to-no-ki*
 you tomorrow rice eat-NPST.I-2-Q
 ‘will you eat rice tomorrow?’
- (j) *arn ŋgom remo sa-ga*
 which village man come-PST.II.3
 ‘From which village does the man come?’ (Fernandez 1968:112)

The OV order is maintained in imperative formations as well in Remo.

- (162) (a) *a-dio wi-ya* (b) *a-niŋ kiyaŋ be*
 OBJ-house go-IMP OBJ-I rice give
 ‘go home!’ ‘give me rice’ (Fernandez 1968:112)

Within the noun phrase or verb phrase, there is some variability permitted but the order of modifiers and heads are generally set. Thus, deverbal ‘participial’ adjectives almost always precede the noun in NPs in Remo.

- (163) *surisuri uriŋbay remo*
 fast walk-ADJ man
 ‘the fast walking man’ (Fernandez 1968:117)

4.1.1 Typological features

The relative order of numeral and noun is variable in Remo with both Num N and N Num order attested. Note that suffixed or enclitic classifiers, similar to those found in various other South Asian languages can appear on the numeral, encoding, for example, human vs. non-human oppositions. Note that this syntactic variability of the relative position of Numeral and Noun is not dependent on categories of humanness or animacy, but may be subject to as-of-yet undiscovered discourse/pragmatic, rather than strict semantic (either functional or lexical) features of the elements involved.

- (164) Num N ~ N Num
- (a) *miŋda remo ŋkwusi-aluŋ di-ta*
 one.HUM man jackfruit-SUB COP-NPST.II
 ‘One man is beneath the jackfruit tree.’
- (b) *bire muy gari-bo? di-ta*
 stone one.NONHUM path-LOC/DIR COP-NPST.II
 ‘One stone is on the path.’ (Fernandez 1968:117)
- (c) *ŋgiđi gisiŋ aka di-ta*
 three.NONHUM chicken here COP-NPST.III
 ‘Three chickens are here.’
- (d) *selane u?uŋđo a-kiŋda? kuma-đen-ta*
 girl four.HUM OBJ-river bathe-PROG-NPST.I
 ‘Four girls are bathing in the river.’ (Fernandez 1968:117)

Other order-based restrictions within the Remo noun phrase include the predominance of the following structural types: Gen N, Dem N, and Adj N. That is, the noun phrase of Remo shows a typical head-final patterning that is common in Eurasian SOV languages, generally speaking, and in South Asian languages, in particular, (e.g. Gen N; Dem N; Adj N), with notable exceptions. More expanded phrases regularly follow this pattern as well, for example, Num Adj N.

- (165) *mbayyo mona?-bay selane-le er-đen-ta*
 two.HUM fat-ADJ girl-PL thresh-PROG-NPST.II
 ‘Two fat girls are threshing.’ (Fernandez 1968:117)

However, post-nominal adjectival modifiers are also possible (i.e. in an N Adj. structure), even those bearing the adjectival or modifier suffix *-bay* (but not genitives or demonstratives (?)) in Remo. Thus, like numerals, adjectives in Remo may appear post-nominally or in internally headed structures like the second example below, with Num N Adj order possible. Variation to this effect is common; what, if any, semantic difference exists between these variants requires further research to elucidate. Whether these post-nominal structures are archaic artefacts of an earlier SVO or VSO structure, an older order that is suggested by both internal-Munda and external-Austroasiatic comparative evidence, is unknown at present, but is at least plausible (unless these variants can be proven to be used only in particular discourse/pragmatic contexts and/or of very recent origin). Thus utterances of the following type are also grammatical.

- (166) (a) *remo mona?bay*
 man fat
 'fat man'
- (b) *muy gisiŋ kaylabay ~ muy kaylabay gisiŋ*
 one.NONHUMAN chicken black
 'one black chicken' (Fernandez 1968:127)
- (c) *kaŋabay remo diredire uriŋ-o?*
 blind.MASC:ADJ man slowly walk-PST.I
 'The blind man walked slowly.'
- (d) *mona?bay selane sum-o?*
 fat girl eat-PST.I
 'The fat girl ate.' (Fernandez 1968:116)
- (e) *goytaŋ kalevay goy-i-da*
 bull black die-AUX-NPST.II
 'The black bull died/is dead.' [SDM]
- (f) *ati-na luntur mba?ar muna?we*
 elephant-GEN ear two big
 'the elephant's two big ears' or 'the elephant has two big ears.' [SDM]
- (g) *remo-na luntur mba?ar[a] qove*
 man-GEN ear two small
 'the man's two small ears' or 'The man has two small ears.' [SDM]

A four-term noun phrase without a numeral expression also usually follows the head-final structure that seems characteristic of Remo, as in example (167) with Dem-Adj-Gen-N order.

- (167) *kon mona?bay selane-ŋ qio*
 this fat-ADJ girl-GEN house
 'This fat girl's house.' (Fernandez 1968:127)

Adverbs generally precede verbs in Remo, which as aforementioned, are usually the final element of a clause (whether this is a simplex lexical verb or complex predicate with a light or auxiliary verb).

- (168) *kaŋabay remo diredire uriŋ-o?*
 blind.MASC:ADJ man slowly walk-PST.I
 'The blind man walked slowly.' (Fernandez 1968:116)

However, the placement of adverbs is relatively free, and they can appear clause-initially, in second position, etc.

Further, within the verb phrase, the relative order of ‘indirect’ (i.e. semantic recipients, beneficiaries, goals, etc.) and ‘direct’ (i.e. semantic patients or themes), usually sees the indirect/human/animate object preceding the direct/inanimate object. Note that in terms of case-marking (which as mentioned in section 3.1.2 is a phrasal proclitic, appearing on the leftmost element of the noun phrase designated as ‘object’ (but not encoded in the verb, as Remo has lost morphological encoding of object properties in the verb that was characteristic of both Proto-South Munda and Proto-Munda (Anderson 2001, 2004, 2007)), it is the ‘indirect’ object that is case-marked (as this is the most likely argument to be animate/human/definite), and thus Remo shows the pattern something akin to that typically known as the ‘primary object’ pattern (Dryer 1986) in the literature, albeit it represents a system in which case is marked on the dependent noun not on the verb in Remo (as it is in Juang, Gorum, or Sora).

- (169) (a) *gitin remo a-mona?bay selane kiyay beq-o?*
 that.CLOSE man OBJ-fat:ADJ girl rice give-PST.I
 ‘That man gave rice to the fat girl.’ (Fernandez 1968:119)
- (b) *niy a-guso? kiyay su-sum be-t-iy*
 I OBJ-dog rice REDPL-eat AUX-NPST-I
 ‘I am feeding the dog the rice.’ (Fernandez 1968:120)

Interrogative sentences without ‘Wh-words’ in Remo are formed with the post-inflectional clitic *-ki*. This, like the post-inflectional progressive (see 3.2.4) is likely a loan from Indo-Aryan.

- (170) (a) *sum-o?-no-ki* (b) *sum-to-no-ki*
 eat-PST.I-2-Q eat-NPST.I-2-Q
 ‘Did you eat?’ ‘Do you eat?’ (Fernandez 1968:112)

There appears to be two elements used as quotatives in Remo, one native and the other borrowed. Generally they have different syntax. The native element is *dɔ* which functions as a kind of defective verb meaning ‘say’. It usually follows the quote and precedes the finite verb meaning ‘say’.

- (171) (a) *dɔkri jul-ɔ se?ta ma dɛdɛm dɛn-nɔ dɔkra dɔ*
 woman see-PST.I SS what REDPL:DO AUX-2 man QUOT
se?ta sun-ɔ?
 SS say-PST.I
 ‘The wife saw him and said “What are you doing, husband?”’
- (b) *suɔ guis-ɔ dɔk-lɔ dɔ se?ta sun-ɔ?*
 fire kindle.fire.by.blowing-PST.I stay-EMPH QUOT SS say-PST.I
 ‘Stay here kindling fire’
- (c) *dɔkra sun-ɔ? nɔ baʔa-rem kub sɔb*
 man say-PST.I you distribute-person very.much hold/catch
dɔt-nɔ dɔ se?ta dɔkra sun-ɔ?
 AUX:NPST-2 QUOT SS man say-PST.I
 ‘The man said, “You, who did the distribution, have taken more (meat).”’

In other uses, it appears to convey the meaning ‘say’ alone with no additional verb, or with it deleted.

- (172) (a) *gisaʔg-e jul-ɔʔ-seʔta dɔkra ui-ga-ni dɔ*
 monkey-PL see-PST.I=SS man go-PST.II.3-PROG QUOT/say:PST.I
 ‘The monkeys saw (her) and said “there goes the man.”’
- (b) *sum-ti-ŋ dɔ seʔta titi bɔrɔŋ sɔb-ɔ-seʔ jikl-ɔ*
 eat-NPST-I QUOT SS hand with hold-PST.I-SS pull-PST.I
 ‘Saying “I will eat” they held the beans with their hands and pulled.’
- (c) *cih mɔr dem-ti-ŋ be dɔ seʔta*
 fie how do-NPST-I DISC QUOT SS
gitin gisag-na siksəŋ mui dʒiʔ-ga
 that.CLOSE monkey-GEN bone one be-PST.II.3
 ‘“Fie upon me! What will I do now?” At that time a piece of bone of one of the monkeys was lying there.’

The other quotative is *ki* which likely derives from an Indo-Aryan loan source. This often precedes the quote that is sets off in the sentence.

- (173) (a) *dɔkri sun-ɔʔ ki anɔa dɔkra nɔ-na kiaŋ*
 woman say-PST.I QUOT no man, you-GEN cooked.rice
gisaʔg-e saɣay-ɔʔ sum-ɔʔ
 monkey-PL scatter-PST.I eat-PST.I
 ‘The woman said “no, husband, monkeys scattered and ate your food.”’ (Bhattacharya 1968:148)
- (b) *dɔkra sun-ɔʔ ki niŋ mui buddi dem-t-iŋ*
 Man say-PST.I QUOT I one trick do-NPST-I
 ‘The man said “let me do a trick”.’ (Bhattacharya 1968:148)
- (c) *munaʔ-bai gisag sun-ɔʔ ki kɔʔn dɔkra sag*
 big-ADJ monkey say-PST.I QUOT DEIC man come:PST
 ‘The big monkey said: “This one is not the wife, the man has come.”’

Note that the use of a quotative is not obligatory (although quite common) in Remo narrative discourse.

- (174) *dɔkra sun-ɔʔ ma palay bug-ɔ-beʔ-ɔʔ-niŋ dɔkri gɔy-ga*
 man say-PST.I what for beat-CV-AUX-PST.I-1 woman die-PST.II.3
 ‘The husband said, “why did I strike? The wife is dead.”’ (Bhattacharya 1968:149)

4.2 Complex sentence structure

As aforementioned, while the investigation of most features of Remo grammar need more and better data so we can advance our understanding of its structure, one area of Remo grammar that stands out as in particular need of further intensive investigation is the formation and structure of complex sentences. Indeed, virtually no data on complement clauses are available for the language. An object complement clause may appear in reduplicated form. Note that this is exactly the structure that

typifies many auxiliary verb constructions in Remo and it is from such complement structures that many such auxiliary verb constructions probably have arisen.

- (175) *ɲiŋ tutu mak-t-iŋ*
 I REDPL:throw/tie know-NPST-I
 'I know how to throw.' (Fernandez 1968:122)

4.2.1 Relative-type clauses

Like complement clauses, relative clauses are mainly lacking in our corpus. A verbal element can be turned into a participle with relative-type functions by adding *-bay* to the basic stem for bi-syllabic verbs and a reduplicated stem for mono-syllabic ones.

- (176) *surisuri uriŋ-bay remo ɖio-bo? wi-ga*
 fast walk-ADJ man home-LOC/DIR go-PST.II.3
 'The fast walking man goes home.' (i.e. the man who walks fast') (Fernandez 1968:117)

4.2.2 Other subordinate clauses (time, manner, cause, purpose, etc.)

Remo does appear to have a range of means of creating subordinate clauses of various functional subtypes. Unfortunately the extent of their use, and how seemingly synonymous constructions differ from each other either in meaning, contexts of appropriateness and connotation remains uninvestigated.

Like many other languages of the world, Munda languages included, clitic postpositional elements (or case forms) may be attached to non-finite verbs to form subordinate clauses of various types in Remo. These may attach to a reduplicated form if the stem is mono-syllabic. They may also attach to a copular form.

- (177) *badɔl and-sa paʔɾ-ɔ mɔmɔrtɔʔg-e tina-ga*
 ?cloudy COP.NEG-DEP sky.to.clear-PST.I star-PL be.seen-PST.II.3
 'Because there was no cloud, the sky has cleared up, stars are visible.'
 (Bhattacharya 1968:85)

Some postposing of subordinate structures is found in Remo. Nominalized or case-marked complements may be found in a post-verbal position, which, therefore, appears sentence medially.

- (178) (a) *sunuʔbo? tug-oʔ-ta su-sugboʔ-ŋa*
 comb:head tie-PST.I-NPST.II REDPL.combing:head-DEP
 'He has tied a comb (to his waist) for hair-combing.' (Fernandez 1968:61)
- (b) *gor-boy suta ruŋ-oʔ-ta kəja-pəlay bodobel-bo?*
 Dom.FEM thread bring-PST.I-NPST.II barter-FOR Bodobel-LOC/DIR
 'A Dom woman has brought thread for bartering at Bodoballe.'
 (Fernandez 1968:61)
- (c) *gurume-ŋ ŋgom remo bire ɖon-oʔ-ta*
 Andrahal-GEN village man stone carry-PST.I-NPST.II
gunom jukjuk-pəlay
 memorial building-FOR
 'The men of Andrahal village have carried a stone for building a memorial.' (Fernandez 1968:61)

Some elements may attach to what otherwise appears to be a finite verb form in the language.

- (179) *ruŋ-qen-kə-ga-sa?*
bring-PROG-PERF-PST.II.3-SS
'while they were taking him in' (Fernandez 1968:100)

Conditional forms, already mentioned in section 3.2.5 are also, properly speaking, a subtype of complex sentence and, therefore, should be included herein as well.

- (180) (a) *a-nay-ŋ qio qi-ki-ŋa ma losuna sum-niŋ*
OBJ-we-GEN house COP-PRF-COND what ADD eat-1
'If we had food in our house, I would also have eaten something.'
(Fernandez 1968:101)
- (b) *gulayne kiyay a-sum-o?-ŋa goy-g-wi-ta*
boy rice NEG-eat-PST.I-COND die-PST-AUX-NPST.II
'If the boy does not eat, he will die.' (Fernandez 1968:101)

An accompaniment converb or dependent form in Remo is also found in *-lo*. Cognate elements are found in Sora, for example, as well.

- (181) *niŋ wi-lo gisiŋ goy-wi-ga*
I go-WHILE chicken die-AUX-PST.II.3
'While I was gone, the chicken died.' (Fernandez 1968:101)

As aforementioned, it may be the case that the genitive case of a pronominal may also function as the subject of a dependent clause in Remo.

- (182) *qɔkra jul-ɔ jul-ɔ a-goŋ-se?ta mayn kiyay-ntra-ma?*
man see-PST.I see-PST.I NEG-CAP-SS -he:GEN cooked.rice-gruel-curry
buŋ-ɔ-ki-n-bɔ ui-ga
put-PST.I-PRF-DEP-PLACE go-PST.II.3
'The man waited there for some time, but then unable to wait more he returned to the place where he lay the rice, gruel and vegetables.' (Bhattacharya 1968:149)

In this sentence the subordinate clause verb has the dependent/attributive suffix *-n* to which the locative/directional case or adposition has attached. Here the meaning created reflects the origin of the element 'place' and makes a relative-type subordinate clause meaning 'the place where.'

Examples of other verb forms functioning as complements of postpositions, appearing in the dependent *-n* form include the following in Remo:

- (183) *qɔkra qɔkri i?-ki-n turgu biri-bɔ gisa?g-e*
man woman go-PLUP-DEP afterwards forest-LOC monkey-PL
sa?me su-sum-sa i?-ga
mandeya.corn REDPL-eat-PURP go-PST.II.3
'The husband and wife being gone, the monkeys went to the field to eat mandeya.' (Bhattacharya 1968:149)

This dependent verb form in *-n* can operate itself to mark a temporally subordinate clause in Remo in, at least, certain instances.

- (184) *gitin kata sun-ɔʔ-ki-n dapre dɔkra*
 that.CLOSE utterance say-PST.I-PRF-DEP suddenly man
maʔ dɔy-ɔ-ki-n ŋkuĩ brɔŋ a-dɔkri
 vegetable.curry cook-PST.I-PRF-DEP earthen.pot INS OBJ-woman
bug-ɔ-beɣ-ɔʔ
 beat-CV-AUX-PST.I
 ‘When this was uttered the man suddenly struck his wife with the pot in
 which the curry was cooked.’ (Bhattacharya 1968:149)

4.2.3 Coordination, co-subordination and switch reference

In the Remo texts in Bhattacharya (1968), there are certain non-finite elements that appear to function as a coordinating ‘conjunction’ but which have an predominant patterning in sentences where the two clauses share or do not share a subject argument. Such a pattern led Anderson and Boyle (2002) to suggest that at least for certain narrative genres, such a distribution seems to represent a switch reference system. The Remo switch reference markers are *seʔ(ta)* for same subject and *-na* for different subject.³ See examples in (185) for same subject and (186) for different subject.

(185) Same subject

- (a) *dɔkri dɔkran-bɔ ui-seʔta mayn lɔge tɔŋg-ɔ*
 Woman man-LOC go-SS she/he:GEN side stand-PST.I
 ‘The woman went up to and stood by her husband.’
- (b) *nɔ gɔsig-seʔta biri-bɔ i-ya*
 you wear.cloth(by.men)-SS forest-LOC go-IMP
 ‘dress like a man and go to the forest’
- (c) *kukusag gine gijɔd-lɔ=seʔta kirime ɔtur-ɔʔ atin*
 tiger teeth gnash-while-SS claw take.out-PST.I that.far.off
gu naŋlili? sugɔ dem-ɔʔ sa, kɔʔn gu-lɔna
 boy embracing like do-PST.I and.then this boy-also
kirime ɔtur-ɔʔ=seʔta gijɔd-lɔ=seʔta a-kukusag naŋlili?
 nail take.out-PST.I-SS gnash-while-SS⁴ OBJ-tiger emb.
sugɔ dem-ɔʔ
 like do-PST.I
 ‘The tiger then gritted his teeth and bringing out his claws moved as if
 to embrace the boy, at which the boy also exposed his fingernails gritted
 his teeth and moved as if to embrace the tiger.’

(186) Different subject

- (a) *a-nij dɔʔtɔr a-beʔ-tɔ-nɔ-ki dɔʔ-na nɔn*
 OBJ-I milk NEG-give-FUT.II-2-Q say-DS you:GEN
baʔagari daktɔr ruis-ɔ sum
 distributed.property milk milk-CV drink(eat)
 ‘(I say), “will you give me milk or not” and he says “milk your side of
 the partitioned property and drink (it)”.’ (Bhattacharya 1968:150)
- (b) *pɔrɔkʔ dɔkri sag-na kiaŋ saɣay-ɔ sum-nay*
 Next.time woman come-DS cooked.rice scatter-PST.I eat-1PL
 ‘Next time the woman comes, let’s scatter the food and eat it.’

Usually same subject coordinate constructions are not simply a juxtaposed finite clause sequence, but this structure does occur in Remo.

- (187) (a) *niŋ sa-g-niŋ a-remo jul-oʔ-niŋ*
 I come-PST.II-1 OBJ-man see-PST.I-1
 ‘I came and saw the man.’
- cf. (b) *niŋ sak-seta a-remo jul-oʔ-niŋ*
 I come-SS OBJ-man see-PST.I-1
 ‘I came and saw the man.’ (Fernandez 1968:113)
- (c) *gulayne sum-oʔ sa-ga*
 boy eat-PST.I come-PST.II.3
 ‘The boy ate and came.’ (Fernandez 1968:113)
- (d) *māʔē njuru muraʔ-taamaŋ biri-bo*
 they morning rise-NPST.II OBJ-they-[GEN] highland.field-LOC
biri biri wi-ta
 highland.field GO-NPST.II
 ‘They get up in the morning and go to (work) in their highland fields.’ [SDM]

Quasi-serialized structures with the first verb in a (past tense/participle or) converb form are also found to a certain extent in Remo clause structure. Since the first verb appears in semi-dependent form, these are not classic serial structures of the type found in Gtaʔ (Anderson this volume).

- (188) *ɖaktɔr ʔuis-ɔ sum*
 ‘milk milk-CV drink(eat)
 ‘milk (it) and drink its milk!’ (Bhattacharya 1968:150)

5 SEMANTICS/DISCOURSE

5.1 Semantics

This has not been extensively examined in Remo to any degree.

5.2 Discourse

Nor has discourse been examined in Remo and little will be offered here. There are, of course, discourse sensitive particles, for example, *-le*, the use of which remain obscure.

- (189) *nɔ-na pɔrsɔlbai bɔb-le gisi a-ɖem-ɔ-na*
 you-GEN clean:ADJ head-EMPH louse NEG-appear-PST-COND
 ‘Lice would not have appeared if your head was clean.’ (Bhattacharya 1968:63)

Another such element is the particle *lɔ* which is frequently found in singular imperative forms and adds a sense of heightened urgency to the command.

- (190) (a) *suʔ guis-ɔ ɖik-lɔ ɖɔ seʔta sun-ɔʔ*
 fire kindle.fire.by.blowing-PST.I stay-EMPH QUOT SS say-PST.I
 ‘Stay here kindling the fire’ (Bhattacharya 1968:149)

- (b) *ɔ dɔkri siriʔ-lo*
 DISC woman quickly-EMPH
 ‘O wife, come quickly, come quickly’ (Bhattacharya 1968:149)

Other discourse sensitive elements, possibly restricted to particular genres (e.g. narrative style) are also commonly found in Remo. One such element is the narrative particle *tɔ*.

- (191) (a) *gitin biri-bɔ saʔme kubete bulu-ga tɔ*
 that.CLOSE forest-LOC mandeya.corn a lot ripen-PST.II.3 NARR.PRTCL
 ‘A lot of mandeya corn ripened in the forest-field.’ (Bhattacharya 1968:148)
- (b) *gitin biri-bɔ saʔme su-sum-sa kubete*
 that.CLOSE forest-LOC mandeya.corn REDPL.eat-PURP many
piriʔ kukum gisakʔ sa-sap tɔ
 bird peacock monkey REDPL.COME NARR.PRTCL
 ‘Many birds, peacocks, and monkeys used to come to that forest to eat the corn.’ (Bhattacharya 1968:148)

Another is the rhetorical particle *be* which seems to be more typical of conversational genres or in conversation in narratives.

- (192) *cih mɔr dem-ti-ŋ be dɔ seʔta gitin gisag-na*
 fie how do-NPST-I DISC QUOT SS that.CLOSE monkey-GEN
siksaj mui diʔ-ga
 bone one be-PST.II.3
 ‘“Fie upon me! What will I do now?” At that time a piece of bone of one of the monkeys was lying there.’ (Bhattacharya 1968:149)

As in many South Asian languages, for example Burushaski or Gtaʔ, there is a structure that is used to string along sentences in narrative discourse. Commonly known as ‘head-to-tail linkage’, this structure consists of a copy of the finite verb in clause-final position from the preceding sentence as a non-finite verb form in clause-initial position in the following sentence. In Remo, this structure is not as common as it is in the aforementioned two languages (one of which, Gtaʔ, is Remo’s sister language). Formally speaking, the verb often takes what appears to be the purposive subordinator *-sa* or in the dependent form *-n* (193–194) while in others (195) it appears to be embedded within the switch reference system (see Anderson this volume for more on this latter structure in Gtaʔ).

- (193) (a) *gitin nsuʔg burɔŋ aten munaʔ-bai gisak*
 that.CLOSE knife with that.one big-ADJ monkey
pug-ɔ beʔ-ɔʔ
 gore-PST.I AUX-PST.I
 ‘The man gored the big monkey with that knife.’
- (b) *pug-ɔ beʔ-ɔʔ-sa dau-dau gisaʔg-e a-munaʔ-bai*
 gore-PST.I AUX-PST.I-DEP REDPL-young monkey-PL OBJ-big-ADJ
gisag jul-ɔ
 monkey see-PST.I
 ‘(he was stabbed and) the younger ones looked at the big monkey.’

- (194) (a) *tul-ɔʔ beʔ-ɔʔ seʔta dɔkra mar dɔkri dʒiɔ-gari iʔ-ga*
 tie-CV tie-PST.I SS man and woman house-PATH go-PST.II.3
 ‘After that the husband and wife went home.’
- (b) *dɔkra dɔkri iʔ-ki-n turgu biri-bɔ gisaʔg-e*
 man woman GO-PLUP-DEP afterwards forest-LOC monkey-PL
saʔme su-sum-sa iʔ-ga
 mandeya.CORN REDPL-eat-PURP go-PST.II.3
 ‘The husband and wife being gone the monkeys went to the field to eat mandeya.’
- (195) (a) *biri-bɔ-n gaʔaŋ puŋek sugɔ ruŋɔʔ-seʔta dʒnɔg-bɔ*
 field-LOC-DEP beans ??? like bring-PST.I-SS trap-LOC
tul-ɔʔ beʔ-ɔʔ
 tie-CV AUX-PST.I
 ‘They brought a bunch of jurunga beans (like *puŋek*) from the field and tied it inside the trap.’
- (b) *tul-ɔʔ beʔ-ɔʔ seʔta dɔkra mar dɔkri dʒiɔ-gari iʔ-ga*
 tie-CV tie-PST.I SS man and woman house-PATH go-PST.II.3
 ‘After that the husband and wife went home.’

As is clear from the glosses in these sentences, it is often best to leave the head-to-tail linkage part out of the translation into English, as this is not a part of English narrative structure and the result is odd sounding in translation, and in the case of different subject structures like (193) hard to fit into the translation at all.

6 LEXICON

6.1 Austroasiatic/Munda components

Remo has a large amount of its basic vocabulary inherited from its Proto-Munda and Proto-Austroasiatic ancestral languages. Some examples of basic vocabulary items of an indigenous origin include the following:

- (196) *likimoʔ tenkur[u]mōʔ nseʔmiʔ*
 seed-eye
 ‘eye pupil’ ‘eyebrow’ ‘nose’ [SDM]
- nsuʔrãʔ nsuʔgʔaʔ* ‘banana’
- | | | | |
|----------------------------|---------------------------|----------------------|---|
| <i>daʔ daʔ</i> ‘water’ | <i>gusoʔ</i> ‘dog’ | <i>tiksuy</i> ‘leg’ | ‘foot’ <i>luntur</i> ‘ear’ |
| <i>gine</i> ‘tooth’ | <i>mōʔ</i> ‘eye’ | <i>boʔ[b]</i> ‘head’ | <i>gutumōʔ</i> ‘forehead’
= <i>gutumɔŋ</i> |
| <i>buli</i> ‘thigh’ | <i>gire</i> ‘chest’ | <i>suloy</i> ‘belly’ | <i>guʔaŋ</i> ‘back’ |
| <i>tumoʔ</i> ‘mouth’ | <i>kurtɔm</i> ‘moustache’ | | <i>leʔaŋ</i> ‘tongue’ |
| <i>tənarɔm</i> ‘shoulder’ | <i>sumúkuti</i> ‘elbow’ | | <i>titi</i> ‘hand’ |
| <i>ōʔonti</i> ‘finger’ | <i>ōʔōʔn]suŋ</i> ‘toe’ | | <i>jomsuy</i> ‘toe’ |
| <i>kunɔay</i> ‘butt’ [SDM] | | | |

6.2 Loan strata

Naturally, a wide range of loan words has entered the Remo lexicon from a range of sources, mainly local Indo-Aryan ones, but also including English via Indo-Aryan.

See Bhattacharya (1968) for an excellent Remo lexicon and still the best source on Remo vocabulary.

(197) *gari* 'truck' *cini* 'sugar'

6.2.1 Hill Remo verb lexicon

In the following two sections, we briefly offer a list of verb stems and their inflectional classes in both the Hill Remo and Plains Remo varieties.

(198) Unexpected (?) Class-I	Sample maybe unexpected Class-II
<i>kub</i> 'cough'	<i>kirol</i> 'shout'
<i>koŋto-beq</i> 'hurt'	<i>kukwi</i> 'wrap, cover with cloth'
<i>kurob</i> 'belch'	<i>kumal</i> 'bathe'
<i>lor</i> 'vomit'	<i>lobl</i> 'love' (loan)
<i>mag</i> 'know'	<i>loklog</i> 'give birth'
<i>nana-dji</i> 'be naked'	<i>lugbar</i> 'dig'
<i>nasto-qem</i> 'spoil'	<i>obunug</i> 'fight'
<i>ŋdɨg</i> 'cut paper'	<i>olug</i> 'put on'
<i>og</i> 'carry child in sling'	<i>orig</i> 'wear, wrap'
<i>ograg</i> 'wind yarn on frame'	<i>oʔadɨŋ</i> 'fight'
<i>olad</i> 'spread' (of cobra's hood)	<i>pwod</i> '(jump) across'
<i>om</i> 'hatch eggs'	<i>pod</i> 'jump over'; 'threaten'
<i>oys</i> 'harvest paddy'	<i>ruŋbog</i> 'wear/put on head'
<i>per</i> 'be spicy'	<i>sog</i> 'hold'
<i>puglay</i> 'bend wood'	<i>somo</i> 'know'
<i>ried</i> 'have wrinkles'	<i>somoqo</i> 'know/recognize s.o.'
<i>ronjug</i> 'dam stream'	<i>ton</i> 'wash'; 'stand'; 'offer'
<i>sambud</i> 'suck blood'	<i>sul</i> 'crow'; 'rub'
<i>sinig</i> 'fart'	<i>sunrol</i> 'greet'
<i>turag</i> 'piss'	<i>tipad</i> 'hide'
<i>umpar</i> 'blow fire'	<i>tugor</i> 'beat'
<i>ur</i> 'run'	<i>aborl</i> 'fight'
<i>uskob</i> 'hiccough'	<i>bagor</i> 'beat one's breast'
<i>sun</i> 'speak, ask, answer'	<i>bam</i> 'get'
<i>tun</i> 'wear on upper arm'	<i>bibal</i> 'marry'
<i>aledag</i> 'squeeze out water'	<i>butuy</i> 'fear'
<i>ay</i> 'be taboo'	<i>dɨyks</i> 'climb'
<i>ard</i> 'shave'	<i>dusog</i> 'wish, desire'
<i>wagbog</i> 'cut off head'	<i>gwidag</i> 'wash'
<i>bartug</i> 'cut wood'	<i>gel</i> 'offer'
<i>batur</i> 'spin fibre against thigh'	<i>gigeb</i> 'heat'
<i>besag</i> 'spread cloth'; 'go to bed'	<i>gijed</i> 'scratch'
<i>bibed</i> 'place yarn on loom warp'	<i>jɨnl</i> 'win'
<i>bokor</i> 'chat'	<i>jiral</i> 'fish w/trap'
<i>bubob</i> 'bark'	
<i>buktaq</i> 'pound fibre'	
<i>bulayɨn</i> 'ball yarn'	
<i>dɨwɨn</i> 'cook'	
<i>dwoyɨndar</i> 'cook something bought'	

qel 'arrive'
qerakjul 'palm-read'
geb 'burn'
gid 'be sticky'
goswi 'chirp, whistle'
gub 'graze'; 'tie (wood)'
gur 'rain'
gurañarjul 'look back'
igsam 'defecate'
jaql 'keep watch'
karabqem 'spoil'

(199) 'Labile' roots

<i>jul</i> (I) 'see, look'	(II) 'hang up'
<i>lug</i> (I) 'dig'	(II) 'set (of sun)'
<i>ruj</i> (I) 'bring'	(II) 'take'
<i>tul</i> (I) 'throw'; 'fix in ground'	(II) 'hang, tie'
<i>tur</i> (I) 'search for'	(II) 'sprout'

Note also: *raŋ* (I) 'cook' *raŋdal* 'cook' (II)

6.2.2 Plains Remo verb list

(200) I	II
<i>arlim</i> 'scratch'	<i>ipor</i> '(a man) said of a woman'
<i>al</i> 'breathe'	
<i>iksam</i> 'shit'	
<i>ukseŋ</i> 'cry' <i>ug</i> 'cry'	
<i>ur</i> 'run' <i>kumab</i> 'bathe'	
<i>urij</i> 'walk'	<i>kumbog</i> 'wash head'
<i>uskobam</i> 'hiccough'	
<i>entur</i> 'spin yarn against thigh'	
<i>oroï</i> 'not to be'	
<i>kaba qeŋ</i> 'get surprise'	
<i>kirol</i> 'shout'	
<i>kukub</i> 'cough'	
<i>kurob</i> 'belch'	
<i>gid</i> 'stick to something'	
<i>giral</i> 'fish w/trap'	<i>guisuy</i> 'wash feet'
<i>guiqag</i> 'wash'	<i>guimog</i> 'wash face'
<i>gundlay</i> 'roll down'	<i>gay</i> 'enter'
<i>gupaq</i> 'celebrate 1st birthday'	
<i>gur</i> 'rain'	
<i>gurag</i> 'spin thread'	
<i>gegeb</i> 'heat'	
<i>gebig</i> 'wear'	
<i>gosig</i> 'whistle'	
<i>gosil</i> 'wear loin cloth'	

<i>cokoti</i>	'mix w/hand'	
<i>jagl</i>	'keep watch'	
<i>jin</i>	'win'	
<i>jojor</i>	'descend'	
<i>jor</i>	'descend'	
<i>qais</i>	'climb'	
<i>qaʔur</i>	'drizzle'	
<i>qakumag</i>	'give birth (hum)'	
<i>qaʔnur</i>	'be cloudy'	
<i>qit</i>	'be' 'live'	
<i>qun</i>	'go, run away'	
<i>qusug</i>	'want'	
<i>qem</i>	'be(come)' 'do'	
<i>qegug</i>	'be hanged'	
<i>qel</i>	'cross by jumping'	
<i>qen</i>	'cook curry'	
<i>qendag</i>	'cook by mixing all vegetables'	
<i>qoqol</i>	'laugh'	
<i>qrig</i>	'lie down'	
<i>taʔmil</i>	'sneeze'	
<i>tar</i>	'become white'	
<i>tugor</i>	'beat'	<i>goy</i> 'be able'
<i>tuqag</i>	'piss'	<i>gois</i> 'die'
<i>tubog</i>	'tie/put on head'	
<i>turguq</i>	'be behind'	
<i>togqak</i>	'get up'	
<i>toŋ</i>	'pound'	
<i>nsaŋ</i>	'refuse'	
<i>pisel</i>	'slip'	
<i>puʔep</i>	'snap fingers'	
<i>per</i>	'be hot'	
<i>babl</i>	'think'	
<i>bibe</i>	'get drowned'	
<i>butuŋ</i>	'fear'	
<i>burug</i>	'swell'	<i>les</i> 'sit'
<i>bekor</i>	'chat'	<i>lakar</i> 'tire'
<i>bobob</i>	'bark'	<i>rag</i> 'tear (cloth/paper)'
<i>murag</i>	'wake up'	<i>mel</i> 'dance'
<i>rakt</i>	'be startled'	<i>lugqag</i> 'drown'
<i>lagbur</i>	'scrape w/spade'	<i>lemog</i> 'sleep'
<i>lor</i>	'vomit'	<i>sag</i> 'come'
<i>waiser</i>	'grumble'	
<i>wagbog</i>	'be ahead'	
<i>sinig</i>	'flatulate'	
<i>silip</i>	'sneeze'	
<i>sukur</i>	'warm self up'	
<i>suʔug</i>	'blossom'	
<i>sun</i>	'say'	
<i>suloŋ</i>	'sweat'	

susubog ‘comb hair’

selaj ‘mature’

Note: *kuy* ‘poke with’ *kūy* ‘scold’ (both = I)

wai (I) ‘call’ (II) ‘marry’

7 BRIEF ANNOTATED SAMPLE TEXT

7.1 The monkey tragedy

From Bhattacharya (1968:147–148), original line numbering retained. Note: In Bhattacharya’s transcription, a final consonant followed by an apostrophe, for example, *k'*, *p'*, indicates a pre-glottalized, unreleased stop, for example, ¹p¹, ²k¹. We have retained this convention in the following text. The influence of Indo-Aryan syntactically is pronounced in this text, for example, the complementizer/quotative/subordinator *ki* used before direct quotes.

- (i) *dɔkra dɪkri dɪ-gaʔ tɔ*
 old.man old, woman stay-PST.II.3 NARR.PRTCL
 ‘There was an (old) man and woman.’
- (ii) *biri gɔi-gɔi gɔy-ʔʔ tɔ*
 forest REDPL-cut cut-PST.I NARR.PRTCL
 ‘They cleared a forest (to cultivate it).’
- (iii) *gitin biri-bɔ saʔme kubete bulu-ga tɔ*
 that.CLOSE forest-LOC mandeya.corn a lot ripen-PST.II.3 NARR.PRTCL
 ‘A lot of mandeya corn ripened in the forest-field.’
- (iv) *gitin biri-bɔ saʔme su-sum-sa kubete piriʔ*
 that.CLOSE forest-LOC mandeya.corn REDPL.eat-PURP many bird
kukum gisakʔ sa-sap tɔ
 peacock monkey REDPL.COME NARR.PRTCL
 ‘Many birds, peacocks, and monkeys used to come to that forest to eat the corn.’
- (v) *gitin dɔkra biri-bɔ piriʔ kukum gisakʔ gubuʔ*
 that.CLOSE old.man forest-LOC bird peacock monkey pig
ɪnʔ-ɪnʔ-sa isa jag-bar ui.
 REDPL.chase-PURP daily watch-AUGM go
 ‘The man would go there daily to watch and drive away the birds, peacocks, monkeys, and pigs.’
- (vi) *gitin dɔkri mai dɔkra-palay kian isa*
 that.CLOSE old.woman 3SG man-FOR cooked.rice daily
ruŋ bebeʔ
 bring REDPL.give
 ‘The woman used to bring food for her husband daily.’

- (vii) *ruŋ-be-beʔ-sa atin gisaʔg-e gari-bo*
bring-REDPL-give-PURP that.far.one monkey-PL path-LOC
jaŋl-ɔ-seʔta a-dʒkri kiaŋ saʔay-ɔ su-sum
watch-PST.I-SS OBJ-woman cooked.rice scatter-PST.I REDPL-eat
'(she went in order to bring and give (him food)), some monkeys on the path watched her and scattered the food (in order) to eat it.'
- (viii) *dʒkri dʒkran-bo ui-seʔta mayn lɔge tɔŋg-ɔ*
Woman man-LOC/DIR go-SS 3SG-GEN ?side? stand-PST.I
'The woman went up to and stood by her husband.'
- (ix) *dʒkra sun-ɔʔ kiaŋ argu ruŋ-ɔ-nɔ.*
Man say-PST.I cooked.rice where bring-PST.I-2
'The man said "where is the food you brought?"'
- (x) *dʒkri sun-ɔʔ ki anda dʒkra nɔ-na kiaŋ*
woman say-PST.I COMP no man, you-GEN cooked.rice
gisaʔg-e saʔay-ɔʔ sum-ɔʔ
monkey-PL scatter-PST.I eat-PST.I
'The woman said "no, husband, monkeys scattered and ate your food."'
- (xi) *niŋ dʒkri-rem ma dɛm-t-iŋ*
I woman-person what do-NPST-1
'I am (just) a woman, what am I to do.'
- (xii) *dʒkra sun-ɔʔ ki niŋ mui buddi dɛm-t-iŋ*
man say-PST.I COMP I one trick do-NPST-1
'The man said "let me do a trick".'
- (xiii) *mar misiŋ ŋjur a-dʒkri sun-ɔʔ*
Again/another one.day morning OBJ-woman say-PST.I
nɔ gɔsig-seʔta biri-bo i-ya
you wear.cloth(by.men):PST.II-SS forest-LOC go-IMP
'The next morning he told his wife "dress like a man and go to the forest".'
- (xiv) *mar misiŋ ŋjur dʒkri tɔ mui kɔdi gɔsi-ga,*
again morning woman EMPH one dhoti wear-PST.II.3
mui taŋgiya mui kɔnd-bo ɔkser-ga
one axe one shoulder*-LOC/DIR hang.up-PST.II.3
'That same morning the woman put on a dhoti and placed an axe on her shoulder.' *(Bhattacharya glosses as 'shoulder' but the word does not appear in his lexicon; the usual word *tənarom*.)
- (xv) *mari dʒkri biri-bo ui-sa gari-gari ui-ga*
then woman forest-LOC go-PURP path-path go-PST.II.3
'The woman went by the path to go to the forest.'
- (xvi) *gisaʔg-e jul-ɔʔ-seʔta dʒkra ui-ga-ni dʒ*
monkey-PL see-PST.I=SS man go-PST.II-PROG QUOT
'The monkeys saw (her) and said "there goes the man."'

(xvii) *pɔɔk' dɔkri sag-na kiaŋ saɣay-ɔ sum-nay*
 Next.time woman come-DS cooked.rice scatter-PST.I eat-1PL
 'Next time the woman comes, let's scatter the food and eat it.'

(xviii) *dɔkra tɔ kiaŋ raŋ-ɔ? ma? dɔy-ɔ*
 Man EMPH cooked.rice cook-PST.I veg.curry cook-PST.I
 'The man cooked rice and veggy curry.'

ntra lɔbur-ɔ
 gruel prepare-PST.I
 'He prepared the gruel.'

mari kiyay tɔ ʃipni-bɔ kib-ɔ?
 then cooked.rice EMPH small.basket-LOC/DIR pour-PST.I
 'Then he poured the rice into a small basket.'

ntra patli-bɔ kib-ɔ?
 gruel small.pot-LOC/DIR pour-PST.I
 'He put the gruel in small pot.'

ma? mundʒi-bɔ sɔb-ga
 vegetable.curry earthen.pot-LOC hold-PST.II.3
 'The veggy curry was held in an earthen pot.'

mari mui dʒamuy-bɔ jɔgaɣl-ɔ mari dɔkri sugo
 then one large.basket-LOC collect-PST.I then woman like
tɛy-ɔ=seʔta gari gari ui-ga-ni
 carry.on.head-PST.I=SS path path go-PST.II-PROG
 'Then he put them in a large basket on his head like a woman and went down the path.'

(xix) *a-dɔkra gisaʔg-e jul-ɔʔ-seʔta saɣay-bar-sa ur-ɔ?*
 OBJ-man monkey-PL see-PST.I-SS scatter-AUGM=PURP run-CV
ui-ga tɔ
 GO-PST.II.3 NARR.PRTCL
 'The monkeys saw the man and ran off' to scatter his food.'

(xx) *gitin-a munaʔ-bai sensenuy a-dɔkra sɔb-ɔ? turgu*
 that.CLOSE-ATTR big-ADJ at.first OBJ-man seize-PST.I later
dʒau-dʒau-bai gisag sɔb-ɔ?
 REDPL-small-ADJ monkey seize-PST.I
 'The bigger one among them went first and caught hold of the man, and then the younger ones seized him.'

(xxi) *gitin dɔkra nsap-bɔ nsuʔg mui pɔɣig dʒi-ga?*
 that.CLOSE man waist-LOC knife one be.tucked.in AUX-PST.II:3?
 'There was a big knife concealed in the man's waist.'

(xxii) *gitin nsuʔg burɔŋ aten munaʔ-bai gisak pug-ɔ beɣ-ɔ?*
 that.CLOSE knife with that.one big-DET monkey gore-PST.I AUX-PST.I
 'The man gored the big monkey with that knife.'

- (xxiii) *pug-ɔ beɸ-ɔʔ-sa ɖau-ɖau gisaʔg-e a-munaʔ-bai*
 gore-PST.I AUX-PST.I-DEP REDPL-big monkey-PL OBJ-big-ADJ
gisag jul-ɔ
 monkey see-PST.I
 ‘He was stabbed and the younger ones looked at the big monkey.’
- (xxiv) *munaʔ-bai gisag sun-ɔʔ ki kɔʔn ɖɔkra sag*
 big-ADJ monkey say-PST.I QUOT DEIC man come:PST
 ‘The big monkey said: “This one is not the wife, the man has come.”’
- (xxv) *mari ɖau-ɖau-bai gisaʔg-e ɖɔɖɔy-seʔta ɖuŋ-ga*
 again small-COPY-ADJ monkey-PL jump-SS flee-PST.II.3
 ‘Then the younger monkeys jumped and ran away.’
- (xxvi) *ɖɔkra kiyan maʔ ntra ɔntur-ɔ-seʔta gisaʔg-en*
 man cooked.rice curry gruel leave-PST.I-SS monkey-PL[DEP]
teŋguʔ teŋguʔ iŋɔn-ɔ
 chase chase chase-PST.I
 ‘The man leaving the rice, vegetable, and the gruel chased them.’
- (xxvii) *gisaʔg-e semug bagbɔ ɖai-ga*
 monkey-PL tree on.top.of climb-PST.II.3
 ‘The monkeys climbed on the top of a tree.’
- (xxviii) *dɔkra semug aluŋ ui-seʔta a-gisaʔ-ge bagbɔ*
 man tree down go-SS OBJ-monkey-PL on.top.of
juɸ-ɔʔ-seʔta puɸa-le niŋ-na saʔme sum-ɔʔ-seʔ
 look-PST.I-SS (abusive)-PL I-GEN mandeya.corn eat-PST.I-SS
sukay-tɔ-pe ki
 escape-NPST.I-2PL QUOT
 ‘The man going below the tree looked up at the monkeys (and said), “You wretched creatures, will you escape eating my *mandeya*?”’
- (xxix) *ɖɔkra jul-ɔ jul-ɔ a-gɔŋ-seʔta mayn kiyay-ntra-maʔ*
 man see-PST.I see-PST.I NEG-CAP-SS he:GEN cooked.rice-gruel-curry
buŋ-ɔ-ki-n-bɔ ui-ga
 put-PST.I-PRF-DEP-LOC/DIR go-PST.II.3
 ‘The man waited there for some time, but then unable to wait more he returned to the place where rice, gruel, and vegetables were laid by him.’
- (xxx) *kiyay ntra mari tɛy-ɔ-seʔta biri-bɔ ui-ga*
 cooked.rice gruel again carry.on.head-PST.I-SS field-LOC go-PST.II.3
 ‘He then returned to the field carrying the rice, gruel, on his head.’
- (xxxi) *dɔkri jul-ɔ seʔta ma ɖeɖem ɖen-nɔ ɖɔkra ɖɔ*
 woman see-PST.I SS what REDPL:do AUX-2 man QUOT
seʔta sun-ɔʔ
 SS say-PST.I
 ‘Seeing (him) the wife (said), “What were you doing, husband?”’

- (xxxii) *sa dɔkra sun-ɔʔ anɔa saʔay-bar-na gisaʔg-e*
 then man say-PST.I no scatter-AUGM-DEP monkey-PL
a-mui dɔkra gisaʔg bus-lɔ beʔ-ɔʔ-niŋ
 OBJ-one.NONHUMAN man monkey stab-EMPH AUX-PST.I-1
 ‘Then the man said: “Nay, I stabbed the male one of the monkeys who came to scatter (the cooked rice).”’
 (NB: -AUGM-*bar* functions like reduplication with loan verb stems)
- (xxxiii) *gɔy-ta ʔu burɔ-ta ʔu*
 die-NPST.II or live-NPST.II or
 ‘Whether he will die or survive (no one can say).’
- (xxxiv) *turgu gitin kiah ntra maʔmbayɔŋ*
 afterwards that.CLOSE cooked.ricegruel vegetable.curry two.persons
dɔkra dɔkri baʔa-g-seʔ sum-ɔʔ
 man woman distribute.food.after.cooking-PST.II-SS eat-PST.I
 ‘After that the husband and wife shared the rice, gruel, and vegetable and ate.’
- (xxxv) *dɔkra a-dɔkri sun-ɔʔ ɔy saʔmele sɔrla:ye*
 man OBJ-woman say-PST.I DISC mandeya.corn cucumber
bɔda-le gaʔaŋ gulay bulu-g diʔ-ta
 boda.grain-PL beans all/many ripen-PST.II AUX-NPST.II
 ‘The man said to his wife: “Dear, the mandeya, cucumbers, boda grains, and jurunga beans, all have matured (in our field).”’
- (xxxvi) *mui dɔnɔkʔ dɔg-naŋ*
 one trap make-1DL
 ‘Let us make a trap.’
- (xxxvii) *tebe dɔkra dɔkri mbayɔŋ dɔnɔkʔ tiyar dɛdɛm tandɔm-ɔ*
 then man woman two.persons trap ready REDPL:do begin-PST.I
 ‘Then both husband and wife began to make a trap.’
- (xxxviii) *dɔnɔkʔ tiyar dɛŋ-ga*
 trap ready become-PST.II.3
 ‘The trap was ready.’
- (xxxix) *biri-bɔ-n gaʔaŋ puŋek sugɔ ruŋ-ɔʔ-seʔta dɔnɔg-bɔ tul-ɔʔ beʔ-ɔʔ*
 field-LOC-DEP beans ??? like bring-PST.I-SS trap-LOC tie-CV AUX-PST.I
 ‘They brought a bunch of jurunga beans (like *puŋek*) from the field and tied it inside the trap.’
- (xl) *tul-ɔʔ beʔ-ɔʔ seʔta dɔkra mar dɔkri diɔ-gari iʔ-ga*
 tie-CV AUX-PST.I SS man and woman house-path go-PST.II.3
 ‘After that the husband and wife went home.’
- (xli) *dɔkra dɔkri iʔ-ki-n turgu biri-bɔ gisaʔg-e*
 man woman GO-PLUP-DEP afterwards forest-LOC monkey-PL
saʔme su-sum-sa iʔ-ga
 mandeya.corn REDPL-eat-PURP go-PST.II.3
 ‘The husband and wife being gone the monkeys went to the field to eat mandeya.’

- (xlii) [missing from text]
‘All of them saw the mandeya crop, but were not attracted to it.’
- (xliii) [missing from text]
‘They went to eat the mandeya inside the trap.’
- (xliv) *sum-ti-ŋ dɔ seʔta titi bɔrɔŋ sɔb-ɔ-seʔ jikl-ɔ*
eat-NPST-1 QUOT SS hand with hold-PST.I-SS pull-PST.I
‘Saying “I will eat” they held the beans with their hands and pulled.’
- (xlv) *kɔl piʔ iʔ-ga*
mechanism break go-PST.II.3
‘The mechanism broke.’
- (xlvi) *gisaʔg-e ɖɔgɔy-bar a-gɔŋ-ga*
monkey-PL jump-AUGM NEG-CAP-PST.II.3
‘The monkeys were unable to jump out.’
- (xlvii) *gulay gisaʔg-e gɔy-ga*
all monkey-PL die-PST.II.3
‘All of them died.’
- (xlviii) *mar misiŋ njur ɖɔkra ɖɔkri raŋbipʔ*
that morning morning.gruel man woman cooking.food
ɖem-ɔʔ seʔta biri-bɔ ui-ga
do-PST.I SS forest-LOC go-PST.II.3
‘Next morning the man and his wife went to the field after taking their morning meals.’
- (xlix) *gulay biri-bɔ bul-ɔ-seʔta ɖɔkra a-ɖɔkri sun-ɔʔ*
all forest/field-LOC roam-PST.I-SS man OBJ-woman say-PST.I
‘The man going round the whole field said to his wife.’
- (l) *suɔ̃ guis-ɔ ɖik-lɔ ɖɔ seʔta sun-ɔʔ*
fire kindle.fire.by.blowing-PST.I stay-EMPH QUOT SS say-PST.I
‘Stay here kindling fire’
- (li) *ɖɔkra ɖɔnɔg-bɔ ui-ga*
man trap-LOC go-PST.II.3
‘Then he went to the trap.’
- (lii) *gulay gisaʔg-e sapay-ɔta*
all monkey-PL be.pressed-PST.I:NPST.II
‘All the monkeys have been pressed underneath.’
- (liii) *ɔ ɖɔkri siriʔ-lɔ*
DISC woman quickly-EMPH
‘O wife, come quickly, come quickly’

- (liv) *niŋ-na sa?me sum-ɔ?ta gisa?g-e kiril-ɔta*
 I-GEN mandeya.corn eat-PST.I:NPST.II monkey-PL become.fat-PST.I:NPST.II
 ‘They have eaten my mandeya and grown fat.’
- (lv) *ruŋ-se?ta sur-naŋ mari ma? dʒy-ɔ-se?ta sum-naŋ*
 bring-ss scorch-IDL then curry cook-PST.I-SS eat-IDL
 ‘Let’s take them and scorch them, and then let’s cook curry and eat.’
- (lvi) *dʒkra mar dʒkri mui mui dʒn-ɔ se?ta ruŋ-ga*
 man and woman one one carry.on.shoulder-PST.I SS take-PST.II.3
 ‘The man and his wife took them one by one on their shoulders.’
- (lvii) *suʃ-bɔ ruŋ-se?ta sur-ɔ?*
 fire-LOC/DIR take-ss scorch-PST.I
 ‘and taking them to the fire scorched them’
- (lviii) *dʒkra seli gɔgɔɾ-ɔ?*
 man meat cut-PST.I
 ‘The man cut the meat (with axe).’
- (lix) *dʒkri ma? dʒy-ʃ*
 woman curry cook-PST.I
 ‘The woman cooked curry.’
- (lx) *dʒkra seli tɔgtɔg-ɔ?*
 man meat REDPL:separate -PST.I
 ‘The man separated the meat (from the bones).’
- (lxi) *dʒkri seli baʒal-ɔ*
 woman meat distribute-PST.I
 ‘The wife distributed the meat (between the two).’
- (lxii) *baʒal-ɔ-se?ta mbayɔŋ dʒkra dʒkri tireleg-se?ta sum-ɔ?*
 distribute-PST.I-SS two.persons man woman together-ss eat-PST.I
 ‘When it was distributed both husband and wife (sat) together (and) began to eat.’
- (lxiii) *dʒkra sun-ɔ? nɔ baʒa-rem kub sɔb*
 man say-PST.I you distribute-person very.much hold/catch
dʒit-nɔ dʒ se?ta dʒkra sun-ɔ?
 AUX:NPST-2 QUOT SS man say-PST.I
 ‘The man said, “You, who did the distribution, have taken more (meat).”’
- (lxiv) *dʒkri sun-ɔ? nɔ tɔ jul-ɔ?nɔ niŋ mɔri kubetek*
 woman say-PST.I you NARR.PART see-PST.I-2 I how very.much
ruŋ-niŋ nɔn mɔ? wag i?ga ki dʒkra
 take-1 you:GEN eye crack AUX-PST.II.3 VOC man
 ‘The wife replied, “But you saw, how have I taken more? Have your eyes burst, O husband?”’

- (lxv) *gitin kata sun-ɔʔ-ki-n dapre dɔkra maʔ*
 that.CLOSE utterance say-PST.I-PRF-DEP suddenly man vegetable.curry
dɔy-ɔ-ki-n ŋkuĩ brɔŋ a-dɔkri bug-ɔ-beʔ-ɔʔ
 COOK-PST.I-PRF-DEP earthen.pot INS OBJ-WOMAN beat-CV-AUX-PST.I
 ‘When this was uttered the man suddenly struck his wife with the pot in
 which the curry was cooked.’
- (lxvi) *dɔkri gɔi-ga*
 woman die-PST.II.3
 ‘The woman died.’
- (lxvii) *dɔkra sun-ɔʔ ma palay bug-ɔ-beʔ-ɔʔ-niŋ dɔkri gɔy-ga*
 man say-PST.I what for beat-CV-AUX-PST.I-1 woman die-PST.II.3
 ‘The husband said, “why did I strike? The wife is dead.”’
- (lxviii) *dɔkra bisar-ɔ ɔkɔʔna biri-bɔ-na taas diʔ-ta*
 man think-PST.I so.much forest-LOC-GEN/DEP field.work be-NPST.II
a-dɔkri cucare bug-ɔ-beʔ-ɔʔ-niŋ
 OBJ-WOMAN in.vain beat-CV-AUX-PST.I-1
 ‘The man then thought (in his mind), “There is so much work in the
 field; I beat her for nothing.”’
- (lxix) *cih mɔr dɛm-ti-ŋ be dɔ seʔta gitin gisag-na*
 fie how do-NPST-1 DISC QUOT SS that.CLOSE monkey-GEN
siksaj mui diʔ-ga
 bone one be-PST.II.3
 ‘“Fie upon me! What will I do now?” At that time a piece of bone of one
 of the monkeys was lying there.’
- (lxx) *kɔʔn gisag seʔsa tɔ a-dɔkri*
 DEIC monkey for.the.sake.of NARR.PARTCL OBJ-WOMAN
ɔ-gɔys-ɔ-niŋ dɔ-seʔta a-siksaj gɔʔe lat bug-ɔ-beʔ-ɔʔ
 CAUS-die-PST.I-1 QUOT-SS OBJ-bone ?? kick hit-CV-AUX-PST.I
 ‘“It was for these monkeys that I killed the woman,” having said this he
 gave a kick to the bone.’
- (lxxi) *gitin siksaj teksuŋ-bɔ gai-ga*
 that.CLOSE bone leg-LOC enter-PST.II.3
 ‘The bone (i.e. its poison) ascended through his leg.’
- (lxxii) *a-dɔkra bisan dɔai-ga*
 OBJ-man poison to.spread(poision)-PST.II.3
 ‘The poison spread through the man (i.e. his whole body).’
- (lxxiii) *tebe dɔkra lɔna gɔi ui-ga*
 then man also die go-PST.II.3
 ‘The man also died.’

NOTES

- 1 For example, unlike other AVCs in Remo, the negative and the causative elements occur on the lexical verb not the auxiliary, which in the case of the negative is usually what the scope of the negative operator is over, that is, the action is not ongoing.

- 2 Curiously, because Remo has lost object morphology in the verb, the first conjugation or 'transitive/active' second singular imperative forms in Hill Remo are the only ones which appear in a mono-morphemic, mono-moraic form, a restriction against which, it has been argued (Anderson 2004, Anderson and Zide 2002) may have triggered the need for an intransitive imperative suffix in Proto-(South)-Munda in the first place. Note that Plains Remo allows no such forms, as explained later in the chapter.
- 3 The *-ta* element in Remo is probably some kind of emphatic. It may be cognate with Gta? *-ka* (see below) usually glossed 'only'. The shorter variant *se?* occurs in the three texts in Bhattacharya (1968) only with a plural (same) subject. Whether this apparent distribution is meaningful and non-random requires further research. Note also the possible relation of the same subject marker to the clausal connective *-sa* 'and then' in Remo.
- 4 Note that in the examples in this sentence, the switch reference marker attaches not to a stem marked with a past tense (or 'participle') marker, but rather with the simultaneous action marker *-lɔ* glossed 'while' in the interlinear analysis.

BIBLIOGRAPHY

- Abbi, A. (1992) *Reduplication in South Asian Languages*, Delhi: Allied Publishers Ltd.
- Anderson, Gregory D.S. (2001) 'A new classification of Munda: Evidence from comparative verb morphology', *Indian Linguistics*, 62: 27–42.
- (2003) 'Dravidian influence on Munda', *International Journal of Dravidian Linguistics*, 32(1): 27–48.
- (2004) 'Advances in Proto-Munda reconstruction', *Mon-Khmer Studies*, 34: 159–184.
- (2006) *Auxiliary Verb Constructions*, Oxford: Oxford University Press.
- (2007) *The Munda Verb. Typological Perspectives*, Trends in Linguistics Studies and Monographs 174, Berlin: Mouton de Gruyter.
- (this volume, Chapter 13) 'Gta?'.
- Anderson, Gregory D.S. and John P. Boyle (2002) 'Switch-reference in South Munda', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 39–54.
- Anderson, Gregory D.S. and Felix Rau (this volume, Chapter 8) 'Gorum'.
- Anderson, Gregory D.S. and N.H. Zide. (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Bhattacharya, Sudhibhushan (1968) *A Bonda Dictionary*, Poona: Deccan College.
- Dryer, Matthew (1986) 'Primary objects, secondary objects, and antitativity', *Language*, 62: 808–845.
- Elwin, Verrier (1950) *Bondo Highlander*, Bombay, New York: Oxford University Press.
- Fernandez, F. (1968) *A Grammatical Sketch of Remo: A Munda Language*, PhD dissertation, University of North Carolina.
- (1983) 'The morphology of the Remo (Bonda) verb', *International Journal of Dravidian Linguistics*, 12(1): 15–45.
- von Furer-Haimendorff, C. (1943) 'Avenues to marriage among the Bondos of Orissa', *Man in India*, 23: 168–172.
- (1950) 'Megalithic ritual among the Gadbhas and Bondos of Orissa', *Journal of the Royal Asiatic Society of Bengal*, 9: 149–178.
- Griffiths, Arlo (this volume, Chapter 12) 'Gutob'.

- Nanda, Bikram Narayan (1994) *Contours of Continuity and Change: The Story of the Bonda Highlanders*, New Delhi; Thousand Oaks: Sage Publications.
- Parkin, Robert (1991) *A Guide to Austroasiatic Speakers and Their Languages*, Oceanic Linguistics Special Publication 23, Honolulu: University of Hawaii.
- Ramachandra Rao (1981) *The Sound System of Remo*, MPhil thesis, Osmania University Hyderabad.
- Zide, Norman (1965) 'Gutob-Remo vocalism and glottalised vowels in Proto-Munda', *Lingua*, 14: 43–53.

GUTOB

Arlo Griffiths

1 INTRODUCTION

Gutob (Gad[a]ba) is a South Munda language spoken in Orissa and Andhra Pradesh (and among migrant workers in northern West Bengal and Assam), by an unknown but probably rather limited number of speakers.*

1.1 Name of the language

The name of the language, Gutob, as used by its speakers when speaking it, is identical with their self-designation as a social group (*pace* Parkin 1991:31), and may be cognate with the ethnonym and language name Gta? (Anderson, this volume).¹ To my knowledge, Ramamurti (1938) was the first among linguists to refer to this language as ‘Gutob’.

Previous work (by Indian authors: Mahānti 1956, Bhaskara Rao 1969, Gaura 1991, Subba Rao 1992) has tended to prefer the (Desia) Oriya name ‘Gadaba’ (/gadba/ ~ /godba/ <gādabā>). This practice seems less appropriate not only on the grounds that it is preferable to employ a community’s own name for its language, but also for the reason that, in this particular case, the name ‘Gadaba’ is potentially confusing (Parkin 1991:31): it has been used in the literature (Burrow and Bhattacharya 1962–1963, Bhaskara Rao 1980, 1998) also to refer to an unrelated Dravidian language spoken by a community called Ollar in Koraput. In view of the likely Munda etymology of the name Gutob (see Note 1), application of the term ‘Gadaba’ – which to all appearances is derived from the former – to this Dravidian language and its speakers would appear to be secondary. According to Bhattacharya (1956:2), the word ‘Ollar’ ‘is usually [?] derived from the Gadba (i.e. Gutob) word *ola* [/olag/ A.G.], meaning “leaf,” and this derivation which may be taken to be an example of the linguistic phenomenon called “folk etymology,”² is associated with a belief that Ollar women previously used to put on leaves instead of clothes’. Burrow and Bhattacharya (1962–1963:46) add: ‘The name used by the [Dravidian-speaking] Gadbas near Salur of themselves is Konḍekor, which means “mountaineers”.’

1.2 Speakers of the language

Members of the numerous other communities (Berger 2002) sharing the living space of the speakers of Gutob (hereafter: the Gutob-Gadbas) refer indiscriminately to them and speakers of Ollari as ‘Gadbas’. Even though there are no practical restrictions on intermarriage (Berger 2002:70, n.13, *pace* Pfeffer 2001a:104), there is an ideal of endogamy, and the two groups are clearly separate, inhabiting different villages. While the latter group has been described ethnographically by

Thusu and Jha (1972), information on the Gutob-Gadbas remained scattered (e.g. von Fürer-Haimendorf 1943, Izikowitz 1969) until the German ethnologist Georg Pfeffer started a series of field trips in the 1980s, resulting in a number of important publications (Pfeffer 1991, 1997, 1999, 2001a, 2001b). His student Peter Berger has spent 21 months doing ethnographic fieldwork among the Gutob-Gadbas from 1999 to 2001: with the publication of Berger's PhD thesis on their ritual and society (2004 [published 2007]), and his other writings (2000, 2001–2002, 2002, 2007b, forthcoming), a relative abundance of ethnographic information is now available. There are also some recent contributions by scholars who are not professional ethnographers, of which Das 1999 and Patnaik 1992 may be cited here.

The only drawback (from our point of view) of the ethnographic work to date, notably that of Pfeffer and Berger, is that it has relied indirectly, via interpreters (Pfeffer), or directly (Berger) on the Desia language for collection of ethnographic data in general, and in particular for information on oral traditions. Berger makes some important sociolinguistic remarks (2007a:24, n.18): 'im Bereich der Rituale (z.B. der Invokationen) ist das Desia die dominante Sprache (was bereits von Fürer-Haimendorf in Bezug auf den Bondo [the speakers of Remo, A.G.] bemerkte, 1943b:168, fn.1)'. Although invocations of gods, spirits, and demons are only rarely done in Gutob, preliminary investigations by myself, partly undertaken in fruitful cooperation with Berger, have revealed rather extensive oral literature (stories, songs) in Gutob, especially among the older generations. In everyday situations, all but the youngest generation still appear comfortable in this language, but the number of children who no longer learn Gutob (whether as second language or at all) seems to be increasing rapidly.

Bilingualism in the local lingua franca, the Desia dialect of Oriya (Gustafsson 1973a/b, Dasgupta and Bhattacharya 1975, Mahapatra 1985, Gustafsson 1989) is universal among the Gutob-Gadbas, and indeed it is so among all the local 'tribal' and 'scheduled-caste' communities (Pfeffer 2001b:779). The primary direction of language-shift is, and presumably has been for a considerable amount of time, towards Desia, while those (mostly male) members of the younger generations who have enjoyed some level of education have also come in contact with the language of schooling, which is standard Oriya.³

1.3 Location of speakers

Speakers of Gutob inhabit villages scattered over the southern parts of Orissa and the adjoining northern districts of Andhra Pradesh. The highest concentration of Gutob-Gadbas – according to the 1995 redrawing of Orissa District boundaries – is found in the Lamptaput block of Koraput District. This appears to be the core area of Gutob-Gadba settlement. From this core area, speakers of the language appear in recent centuries to have migrated to the plains of Andhra Pradesh (Patnaik 1992:4f., also Burrow and Bhattacharya 1962–1963:46), and also to Rayagada District, where anthropologist Roland Hardenberg and I have seen them living among the Kondhs in a village near the market place Majiguda, not far from Kalyansinghpur.

1.4 Number of speakers

Due to the confusion about whom the name 'Gadaba' refers to (sections 1.1–1.2, see also Thusu and Jha 1972:1ff.), it is nearly impossible to get reliable data on the

number of speakers. The 1961 census, according to Manoharan (2001) the most reliable census to date for linguistic-demographic data, clearly states:

Though lately work on a section of Gadaba has been reported to have revealed Dravidian affiliations yet for want of a comprehensive work on Gadaba language, it has been decided ... to conform to the Linguistic Survey of India Classification. Gadaba has therefore been included here in the Munda Branch of the Austro-Asiatic Sub-Family. (Mitra 1964:clxxxii)

This means that its figure of 40,193 ‘Gadaba’ speakers conflates (at least) two entirely unrelated languages. According to Manoharan (2001:129), the most recent available census (1991) lists approximately 28,000 ‘Gadaba’ speakers, again conflating (at least) two linguistic communities.

There are no data on the relative strengths of the Gutob- and Ollar-Gadba populations, nor – more importantly – on the level of retention of the native languages among these two communities, as opposed to a switch towards Desia Oriya, or Telugu in Andhra Pradesh. Basing myself on the above figures, and on my impression (developed during an extensive field trip through large tracts of the southern part of Koraput District) that Ollar-Gadbas are relatively more numerous than Gutob-Gadbas, I had come to the estimate that there are approximately 10,000–15,000 speakers of Gutob.⁴ But after these words had been written, the following even more pessimistic estimate came to my attention (Rajan and Rajan 2001b:lf.):

Gadabas are concentrated mainly in the Koraput District of Orissa. The total population of all Gadaba in Orissa is 56,911 according to the 1981 census of India. The recent survey of the Asha Kiran Society (1998) puts Gutob as 5,000. This low number is because of two major hydroelectric power projects in the district. As a result of these hydroelectric projects many Gadaba people were evacuated and resettled at different places in Koraput. Thus they became a minority, and they gradually shifted to the Indo-Aryan Desia language. Only forty villages in Lamtaput still maintain their Gutob-Gadaba language.⁵

1.5 Classification/taxonomic issues

Gutob’s nearest relative is Remo, spoken in the immediate vicinity of the core Gutob area by the Bonda or Remo tribe (Anderson and Harrison, this volume): the similarities with Remo in every field of grammar are great. In fact, Bhattacharya (1968:xxx) affirms: ‘In the beginning we were of the view that Gutob and Bonda [= Remo] are two dialects of the same language. But we were convinced later that they are different enough to be treated as two languages.’ Bhattacharya’s judgment has found general acceptance among linguists and appears indeed to be supported by the data (e.g. the two verbal systems are quite distinct). On the relationships of Gutob further afield in the Munda language family, see now Anderson 2001, who emphasizes some previously unrecognized grammatical similarities with Kharia (Peterson, this volume).

1.6 Known dialect divisions

It can be stated with some confidence that at least two dialects of Gutob exist: I shall call these ‘Koraput Gutob’ and ‘Andhra Gutob’, and shall assume the former

to represent a 'standard' form of the language. Some phonological differences *within* Koraput Gutob (/boɖoŋ/ :: /buɖoŋ/ 'with', /uigi/ :: /uizi/ 'went', /uiloŋ/ :: /iloŋ/ 'will go', /oʔn/ :: /oʔon/ 'child') may after further research turn out to correspond with a two- (or three-) way division of the core area of Gutob in Koraput: east of the Kolab reservoir (villages Tikrapada, Deptahanjar), west of the Kolab reservoir (Kangrapada, Guneipada, Jalahanjar, Pipalput, Alangpada, Raipada etc.). Perhaps dialectically distinguishable from the previous two is the somewhat removed Onkadili area (Gadbapada), not yet – to my knowledge – visited by any linguist interested in Gutob.⁶

The level of mutual intelligibility of Andhra Gutob with Koraput Gutob is uncertain. The Andhra dialect differs from the latter in phonology (see Bhaskara Rao 1969), having undergone strong influence from the locally dominant Dravidian language Telugu. Data on other fields of grammar are much less reliable, being available only in the linguistically somewhat naive study by Subba Rao (1992). Differences from Koraput Gutob are especially striking in the lexicon, where besides influence from Telugu, one notices conservatism, for example, in the retention of native numerals. Below this last stratum of loan vocabulary, evidence for a layer of Indo-Aryan (presumably Desia Oriya) loans, notably /lokune:n/ 'people' (Subba Rao 1992:19), confirms the assumption made on non-linguistic grounds that the Gutob-Gadbas in Andhra Pradesh form an offshoot from the core habitat in Koraput.

Whether the language of Gutob-Gadbas in other areas, such as Rayagada District, belongs dialectically with one of the above groups, or whether one or more additional dialect groups exist, remains an open question.

1.7 Literary status

No indigenous written literature exists. Virtually none of the Gutob-Gadbas in Orissa are literate. With attitudes toward the value of education, especially of women, being rather negative among the Gutob-Gadbas, and with the actual educational facilities offered by the State government being extremely poor, only few have pursued their educational career long enough to have learned the written form of Oriya and its script. According to the *Ethnologue*, the 'literacy rate in second language', that is, (standard) Oriya, perhaps more properly called 'third language', was 6.53% in 1977. Oriya script would lend itself well to rendering Gutob, with the exception of the glottal stop /ʔ/, for which the Oriya 'visarga' sign <ḥ> might be adopted. For previous attempts to render Gutob in Oriya script, see Mahānti (1956), Gaurā (1991), and the *Gutob-Gadaba Language Learner's Guide*, published by the Asha Kiran Society in 2001.

1.8 History of research, prospects

Although the existence of the language has been known since the middle of the nineteenth century, and although this was, along with Sora, the only Munda language of Southern Orissa to be included in the *Linguistic Survey of India* (Konow 1906: 229–238),⁷ Gutob is still a relatively poorly known language. The small booklet in Oriya by Mahānti (1956) gives some idea of the nominal and verbal morphology, and some classified lexical data, but is presented in an inconsistent and imprecise Oriya

transcription. The work by Gauṛa (1991) was done with the collaboration of one native speaker and – as is clear from their name Khemdu (the name of the Bear clan, which has no members among the Gutob-Gadbas) – two non-native speakers. Although it is basically written in Oriya, it has the advantage (for linguists not conversant with Oriya) of using both IPA and Oriya transcriptions, but does not distinguish native and borrowed morphology, and especially phonology/lexicon. The work contains a small collection of texts (in Oriya script only).

The most serious work has been done from the 1960s onwards under the auspices of the Chicago *Munda Languages Project*. This has resulted in quite extensive typed text collections (DeArmond 196?, Zide and Das 1965), regrettably without a (good) translation or grammatical analysis as well as an incomplete draft of a verb lexicon (Zide and Das 1963). These materials have been kindly put at my disposal for use in this chapter by Zide and DeArmond. Besides the texts, there are numerous grammatical, and especially comparative studies by Zide and his students (e.g. DeArmond [no date], 1976), which however give little insight into the overall grammatical system of the language, and its lexicon. The same problem holds true for the comparative studies by Bhattacharya (1975), relying on his own field data for Gutob.

Sarma (2001) presents some sociolinguistic information on the Andhra Pradesh Gutob-Gadbas. A Tamil couple of the local (Christian) Asha Kiran Society, just south of Lamptaput, Jamuna and Herold Rajan, have been working in the Lamptaput area since 1996; Peter Berger kindly provided me with their publications 2001a and 2001b.⁸

In the following, I refer to my database – work in progress – which has taken as its starting point the texts recorded and transcribed by Norman Zide and Richard DeArmond for the mentioned *Munda Languages Project* at the University of Chicago in the 1960s.⁹ References with the structure ‘4.5’ refer to Zide’s texts, and those with the structure ‘C.5’ to DeArmond’s. I hereby gratefully acknowledge both scholars’ generosity in placing their material at my disposal. The data are interpreted here in accordance with my own results reached during three (brief) periods of fieldwork in the village Jalahanjar in 2001, 2002, and 2004. It may be noted that word and morpheme boundaries follow the transcriptions in the Zide and DeArmond texts: no attempt is made to distinguish suffixes, clitics, and postpositions.

2 PHONOLOGY

Unfortunately, a clear picture does not emerge from the several publications which have been devoted to Gutob phonology by earlier researchers (Zide 1965:44, 1972:512, Gauṛa 1991:3–15). Rajan and Rajan (2001a) present a sketch of their interpretation of the phonology, which offers a clearer picture, although I do not agree with their analysis in all respects. My own fieldwork up to this time has not concentrated on phonological issues. The following paragraphs are, therefore, rather tentative, and several points will most likely need to be revised and supplemented in light of future fieldwork.

2.1 Vowel inventory

The vowel inventory of Gutob is as follows. The phonemes are given along with their most common allophones.

- | | |
|-------------|---------|
| (1) i [i,ɪ] | u [u] |
| e [e,ɛ] | o [o,ɔ] |
| a [ɑ,ʌ,a] | |

The following two minimal pairs will suffice as evidence for contrast between /i, e/ and /u, o/:

- (2) /pid/ ‘to remove thatch/tiles’ vs. /ped/ ‘to blow’
 (3) /sun/ ‘to throw’ vs. /soŋ/ ‘to sell’

Rajan and Rajan (2001a) assume contrast between /a/ and /ʌ/, but their examples in my analysis all contain the single phoneme /a/.¹⁰ Only loan vocabulary causes some problems for the above picture, with such (semi-)minimal pairs as follows (De. stands for Desia and Gu. for Gutob in the following examples):

- (4) /laz/ [la:dʒ] (De.) ‘shame’ vs. /laj/ [lac] (Gu.) ‘who?’
 (cf. Desia /niman/ [nima:n] ‘good’ etc.)
 (5) /dos/ [dɔs] (De.) ‘10’ vs. /dos/ [dos] (De.) ‘sin’

The vowels may appear combined in the following diphthongs, those appearing only in loan vocabulary being given in parentheses:

- | | | | | | |
|-----|--------|------|--------|----|-------|
| (6) | Vi | iV | Vu | uV | Vo? |
| | a /ai/ | /ia/ | (/au/) | — | /ao/? |
| | o /oi/ | /io/ | (/ou/) | — | — |
| | e /ei/ | /ie/ | (/eu/) | — | — |

The only known example of the combination /uV/ in native vocabulary does not seem to be monosyllabic, and is therefore not treated as a diphthong here: *ui* ‘to go’ is a nonreduplicating, therefore bisyllabic root (see DeArmond 1976:215 and section 3.2.10.1). Given the limited knowledge about Gutob phonemics and phonetics, I am hesitant to make further statements about the language’s diphthongs. It may be more appropriate to treat the close vowels in these diphthongs as approximants (/y, w/).

2.2 Suprasegmental phenomena (tone, register)

These phenomena have not yet been investigated in a detailed and systematic manner. Tone is known not to be phonemic, as is vowel length. Nasalization of vowels can occur under certain morphophonological conditions (see section 2.6), and can in rare cases become phonologically distinctive: contrast [rĩã] ‘bring!’ with *lia* (De.) ‘puffed rice’, and *sĩ* ‘sun’ with *siŋ* ‘to wind’. Note also the word *mĩõ?* ‘again’, which is morphologically opaque. The language seems to show (remnants of) a system of glottalized vowels (see Zide 1965:49–53). I have noted the words *laʔŋ* ‘tongue’, *gaʔŋ* ‘?’, *guʔŋ* ‘hunt’, *naʔŋ* ‘last year’, *uʔŋ* ‘to swell’, *oʔn* ‘child’, *uʔn* ‘four’ (obsolete), *aʔl* ‘bamboo’, *soʔl* ‘oil’, *paʔr* ‘to dawn’, and *tiʔr* ‘to sprout’. The phenomenon needs further study: attribution of the glottal stop to the vowel rather than to the coda is somewhat arbitrary for the time being.

2.3 Consonant inventory

The following table shows the inventory of obstruents in Gutob, along with their most common allophones. Phonemes that seem to be restricted to loan vocabulary are given in parentheses.

(7)		voiceless	voiced
	labial	p [p, pʰ]	b [b, pʷ]
	dental	t [t, tʰ]	(d)
	retroflex	(ʈ)	ɖ [ɖ, ʈ, ʈʰ]
	palatal	—	j [j, is, cʰ]
	velar	k [k, kʰ]	g [g, Ø, kʰ]
	dento-alveolar	s[s, ts, tʃ]	z [dz, dʒ]
	glottal	h, ʔ	—

Note that the allophones of the phonemes /s/ and /z/ are probably in free variation, and that /z/ occurs only rarely in words that cannot readily be identified as a borrowing from another language, for example, *zu* ‘to see’. /p/ can be fricativized, that is, nearing [ɸ] or [f], in initial position.

In final position, only unreleased plosives are permitted, which means the distinction between voiced and voiceless is neutralized here.¹¹ /d/, in native vocabulary, here merges with /ʔ/.

From the structural point of view, it seems that /j/, the unreleased palatal stop ([cʰ]) which occurs only in final position, belongs together with /s/, which occurs only in initial position, in native vocabulary.

In intervocalic position, labials → [b], /d/ → [t], /j/ → [is], and velars → /g/ (optionally Ø: /sala(g)-oʔ/).

The status of /h/ is marginal in Gutob, as it is in Desia.¹² In Gutob, it seems to be found – besides in such rare loanwords as *honu* [ɔnu, hɔnu] ‘monkey’ or *hundji* ‘village shrine’ – only in distal deictic demonstratives: *hu(nu)* ‘that (over) there’ and in interjections like *hu*, *āhā*.

The following sonorants are found in Gutob:

(8)	nasals	m, n, ɲ, ŋ
	approximants	r, l

The language shows a number of clear cases of free /n/~l/ variation. For example, the root *non* ‘to chase’ has a variant *lon*.

/ŋ/ is dropped in intervocalic position with concomitant nasalization of the bordering vowels, for example, *riŋ-oʔ* [riŋoʔ] ‘brought’ (contrast *riŋ-gi* [riŋgi] ‘took’).¹³

2.4 Syllable structure and phonotactics

The vast majority of the language’s syllables have the structure C₁VC₂, in which either C₁ or C₂, or both, may be Ø. Very rarely do we find syllable-initial clusters in native vocabulary, for example, *sin.droŋ* ‘medicine’: in such C₁C₂VC₃ syllables, C₂ is always /r/. The special case of C₁VʔC₂ words has been discussed in section 2.2. Syllable-initial and final clusters are only slightly more common in loan words, for example, *druka* ‘tiger’, *bund* ‘part of a plant stem’. Rajan and Rajan (2001a:26–35) give some further details.

2.5 Intonation/stress

Zide (1965:44) remarks, ‘Any syllable of CVC shape, and any morpheme-final syllable in a nonaffixal morpheme is – by definition – stressed.’ I have no data on intonation

or stress to confirm or contradict this statement. See Bhattacharya (1968:xxi and xxiii) on the closely related Remo language.

2.6 Morphophonology

Both the middle and the active past tense suffixes (see section 3.2.3–3.2.7) involve certain morphophonological changes. The quality of the vowel in the former (*-gu* or *-gi*) depends on the shape of the verb root. The MID-PST *-gV* takes the shape *-gi* after root-final /j/ (*goj-gi* ‘died’, *buj-gi* ‘lost’),¹⁴ and after the roots *piŋ* ‘to come’, *riŋ* ‘to take’ (but see *a-ri(ŋ)-gu-nu* ‘who has not been taken’, i.e. ‘spinster’), and *ui*¹⁵ ‘to go’. Apparently, the allomorph *-gu* occurs in all other cases (*guj-ti-gu* ‘washed his hands’, *qu-gu* ‘was’, *log-gu* ‘fell’, and *bil-gu* ‘got drunk’).

Roots ending in *-a* show elision of the vowel of the active past tense suffix *-o?*: *eta* (De.) ‘to think’, PST *eta-?* ‘thought’; the same would seem to hold true (optionally) of borrowed roots which normally end in *-ei*.¹⁶

Vowel harmony in word formation is a productive process in this language (see section 3.1.10 and cf. Zide 1965:46f.). It seems to be this process which governs the distribution of the allomorphs /dɛi/ and /dɔi/ of the third singular inalienable possessive marker: the second form seems to occur only after words with /o/. Vowel harmony triggered by the OPT suffix is encountered in the most important irregular verb, namely *qu(k)* ‘to be (located)’ (example 63). The presence of the final consonant *-k* in that same root is subject to further morphophonological rules:

qu before consonants and before *-o?*, in CAUS *ob-qu-o?* ‘made stay’
duk in IMP before *-a*: *duk-a*, *dik* in OPT before *-e*: *dik-e*

Another important irregular verb is *qem* ‘to become/to make’:

qem before HAB, ACT-NPST (*qem-to* ‘makes’, *qem-tu* ‘will make’) and vowel-initial suffixes, for example, *qem-o?* ‘made’
qen/qeŋ before the stem in reduplication and before MID-PST: *qenqem* ‘to make/become’, *qenŋu* ‘became’
qe before MID-NPST *-loŋ*: *qeloy* ‘will become’

Loss of the root-final nasal is also seen before the MID-NPST suffix in roots with final /ŋ/: *riŋ* ‘to take’ *ri-loŋ* ‘will take’, *piŋ* ‘to come’ *pi-loŋ* ‘will come’. Whether it is lost in the same way in all such roots is not yet known.

Two roots are known where a velar alternates with /ks/ before vowel-initial suffix, namely *pig* and *log* (e.g. *piks-o?* ‘broke in half’ and *obloks-o?* ‘caused to fall’).

3 MORPHOLOGY

3.1 Nominal morphology

3.1.1 Number

Gutob has two grammatical numbers: singular (unmarked) and plural, marked by the suffix *-nen* (PL). The explicit marking of plural number seems in many cases to be optional. A noteworthy usage of *-nen* is to form elliptic plurals: for example, *ioŋ-nen*

‘mother and those with her’. Unlike other South-Asian languages, Gutob does not employ number to express honorific status, a category which seems to be irrelevant in this language (*pace* Rajan and Rajan 2001b:45).

3.1.2 Case

According to the tentative analysis adopted here, case suffixes are always attached directly to the bare nominal, and are in this way distinguished from postpositions (see section 3.1.9), which can take the genitive case. There are three morphological cases in Gutob:

- (a) Subject case: unmarked.
- (b) Genitive/Possessive/Attributive case: marked with *-nu*. Postpositions (except *(pu)lai*) seem to take the genitive case of animate nouns and the unmarked object case of inanimate nouns.
- (c) Object case: unmarked, or marked with a suffix *-(pu)lai* (formally a postposition, though it very rarely occurs with *-nu*)¹⁷ on nominals and pronouns, and a prefix *o-* on pronouns. The same affixes are used to distinguish the indirect (or more human/animate) from the direct object.

The following example contains a combination of object-case markers:

- (9) *o-nij o? zana, oqʉoʎn-lai zana ura?,*
 OBJ-1SG EMPH known boy-OBJ known NEG
gikiŋ-pulai zana ura?, o-laj zana ura?.
 wife’s.elder.brother-OBJ known NEG OBJ-who known NEG
 ‘Only I know, the boy doesn’t know, [my] brother in law doesn’t know, nobody [but me] knows.’ [2.248]

The next example shows the direct object (*onob*) twice unmarked, both in indefinite and in definite use; it shows the two postpositional OBJ markers used in constructions parallel to the reduplicated infinitive (see section 3.2.8).

- (10) D: *mono? onob de?na, tuno? moq-gu ui-loŋ-nij.*
 where girl QUOT there getup-MID.PST go-MID.NPST-1SG
 M: *onob loŋei-lai o? ... ki maŋ paiŋi deŋ-deŋ?*
 Girl fuck-OBJ EMPH ... OR what work REDPL-DO
 D: *onob loŋei-pulai o? ura? mĩō? maŋ paiŋi deŋ-deŋ?*
 Girl fuck-OBJ EMPH not else what work REDPL-DO
 D: ‘Noticing a girl somewhere, I will get up and go there.’
 M: ‘To fuck the girl, that is. Or to do what?’
 D: ‘Just to fuck the girl, what else would I do?’ [C.113–119]

Direct and indirect object can be differentiated by marking the latter (according to Rajan and Rajan 2001b:14 only if both are animate):

- (11) *nij ramu-lai guso? beq-oʎ-nij*
 1SG Ramu-OBJ dog give-ACT.PST-1SG
 ‘I gave the/a dog to Ramu.’ (Rajan and Rajan 2001b:14)

3.1.3 Person

For the third person, Gutob can mark inalienable possession on kinship terms (though importantly, not on body parts). The suffix involved is *-dei/-doi*.

- (12) *utu pap-nu lagire remol-dei-lai mo? kana-gu*
 that sin-GEN due.to husband-3POSS-OBJ eye go.blind-MID.PST
ui-gi kimboj-dei boiri-gu ui-gi.
 AUX(go)-MID.PST wife-3POSS go.deaf-MID.PST AUX(go)-MID.PST
 'Due to that sin, her husband's eyes became blind, his wife became deaf.' [12.2]
- (13) *nom maŋ roza-nu oduo?n-doi-ki, die? tu onob sala-?*
 2SG Q king-GEN son-3POSS-Q QUOT that girl ask-ACT.PST
 'Are you the king's son, perhaps, the girl asked.' [15.108]

3.1.4 Definiteness

There seem to be no markers for definiteness in Gutob other than the optional (?) OBJ marking of definite objects (see example 19, also section 3.1.2). Reference to unspecified, non-definite nouns is made by means of *muiro?* or *ek*, respectively the native and the borrowed word for 'one' (see section 3.1.8), or by means of indefinite pronominals.

- (14) *maj muiro? gamsa mono?-nu-ki tor-o?-su riŋ-o?*
 3SG one towel where-GEN-Q search-ACT.PST-CONJ AUX(bring)-ACT.PST
ekdom barogaŋa dem-o?-su odi?-gu du-tu
 immediately 12.knots make-ACT.PST-CONJ wear-MID.PST AUX(be)-PRS
 'After he had found himself a towel from somewhere, and immediately had made 12 knots, he wore it.' [10.6]
- (15) *muiro? oduŋ di?to goj-gi ui-gi-su*
 one boy QUOT die-MID.PST AUX(go)-MID.PST-CONJ
quba deŋ-gu.
 ghost become-MID.PST
 'A boy, they say, died and became a ghost.' [7.1]

3.1.5 Class/gender

Bhattacharya states (1976:190):¹⁸

It is usually believed that Munda gender is a two-group inflectional and concordant type based on the concepts of animacy and inanimacy. The rigid pattern which is found in Sa[ntali] is considered to be the original Munda gender-type. But this pattern of animate and inanimate gender is not found in the five South Munda languages, Saora, Parengi, Gutob, Bonḍa and Diḍeyi, and it has become very dim in the Central Munda languages, Juang and Kharia.

It is possible, however, that a distinction between animate and inanimate nouns does play a marginal role in Gutob morphosyntax, for example, in the selection between the allomorphs *dig* and *di?ke* 'from' (see section 3.1.9).

Bhattacharya (1976:195) distinguishes ‘compounded sex-based gender’ and ‘inflected sex-based gender’.¹⁹ The former are combinations of sex-linked words with sexually indeterminate nouns. Thus, when a distinction between male and female is to be made clear in Gutob, it can be expressed by using the (Desia) words *andra* ‘male’, *gadra* ‘ram, male animal’, *mai* ‘female’: for example, *andra girem* ‘tom cat’, *andra kirtag* ‘stallion’, *gadra gime?* ‘billy goat’, *gadra menda* ‘ram’, *mai guso?* ‘bitch’, and *mai kirtag* ‘mare’ (see also *podru-taj* ‘male calf’). Or it can be expressed morphologically, with suffixes (-a masculine, -i feminine), in Desia borrowings: *puḍralpuḍri* ‘male/female buffalo calf’, *buḍal/buḍi* ‘old man/woman’.

3.1.6 Pronouns

3.1.6.1 Personal pronouns

The following table shows the personal pronouns of Gutob:

(16)		SG	PL
	(1)	<i>niḡ</i>	<i>nei/naj</i>
	(2)	nom	pen
	(3)	<i>maj</i>	<i>majnen</i>

These forms are identical with the pronominal enclitics marked on verbs (see example 60), except that the third person singular is there unmarked (or Ø-marked), and the third person plural is marked only by the plural morpheme *-nen*. That this morpheme is simply a marker of number and not a person marker is made clear from collocations with 1PL and 2PL pronominals:

(17)	<i>pen</i>	<i>eran</i>	<i>ḡem-oʔ-na</i>	<i>ḡem-nen,</i>	<i>niḡ-nu</i>	<i>lioḡ</i>
	2PL	how	do-ACT.PST-CV	do-PL	1SG-GEN	wet.rice.field
		<i>sui-tu-niḡ.</i>				
		plow-ACT.NPST-1SG				
		‘You do whatever you do, I shall plow my wet rice field.’ [B.6]				

Comparative evidence from Gta? (Zide 1968:349) confirms that pronominals *nei* and *naj* were originally used to mark the distinction between inclusive and exclusive 1PL. In the field, I have occasionally felt it would be possible to prove that traces of this distribution still persist, but the database does not provide any clear evidence. It does, however, provide evidence for their merger:

(18)	<i>uraʔ,</i>	<i>kebe</i>	<i>naj</i>	<i>ar-omtur-nei.</i>
	no	when	1PL	NEG-leave-1PL
	‘No, we will never leave it!’ [B.13]			

The two forms seem to have developed a new distribution, where use of *-naj* is largely restricted to hortative function, in combination with imperative verb forms, and *-nei* is used elsewhere. See the following instructive examples:

(19)	<i>nei</i>	<i>pi-piḡ</i>	+ <i>el</i>	<i>niḡ-nu</i>	<i>bobreḡ-nen</i>	<i>boiragi-lai</i>	<i>bug-naj</i>
	1PL	REDPL-COME	+ time	1SG-GEN	brother-PL	sadhu-OBJ	beat-1PL
	<i>ḡieʔsu</i>	<i>non-oʔ</i>		<i>ri-riḡ-nen</i>		<i>ḡu-gu.</i>	
	QUOT	chase-ACT.PST		REDPL-AUX(take)-PL		AUX(be)-PST	
	‘At the time of our coming [when we came], my brothers were chasing after [him], saying, “Let’s beat the sadhu”.’ [15.166]						

- (20) *qo, i-a-naj nei-nu bo?*
 come.on GO-IMP-1PL 1PL-GEN LOC
 ‘Come on, let’s go to our place.’ [10.148]

Still, there are cases where *naj* is used outside such a hortative context:

- (21) *ura? ni qisel, o-naj-sa kete boros qey-gu sa?mel*
 no ni friend OBJ-1PL-also some year become-MID.PST millet
ka?ei-o? beq-o?, keroŋ ar-buhu-to, taire...
 be.wasted-ACT.PST AUX(give)-ACT.PST paddy NEG-ripen-NEG.PST so
 ‘No, friend. It’s been a number of years now that our millet got wasted, the paddy did not get ripe, so ...’ [D.53]

3.1.6.2 Interrogative, relative, and indefinite pronouns

In native vocabulary, we find six basic interrogative forms, which can all function as relative pronominals; loans from Desia provide all other interrogatives.

- (a) *maŋ* ‘what?’, whence *maŋdem* ‘why?’
 (b) *mono?* ‘where?’
 (c) *umbo?* ‘in what direction?’, whence *umbo?qi?ke* ‘from where, from what direction?’
 (d) *unqoi* ‘when, which day?’, whence, apparently, *ar-unqoi* NEG-which.day ‘next year’ (= ‘not any day soon?’)
 (e) *eran/laren/lemran* ‘how?’
 (f) *laj* ‘who?’

maŋ can be used both as a substantive and as an adjective:

- (22) *uloŋ riŋ-o?-na maŋ dem-to-pen?*
 straw bring-ACT.PST-CV what DO-HAB-2PL
 ‘What do you do when you get the straw?’ [1.85]
 (23) *maŋ din-e pen puza dem-o??*
 what day-LOC 2PL puja DO-ACT.PST
 ‘On what day did you do the puja?’ [2.271]

Note further the use of this word to introduce questions (cf. the identical use of Hindi *kyā* and Oriya *kaṇa*):

- (24) *maŋ ser-gu + me?qiy-gu-su lo?ei-to-pen?*
 Q sing-MID.PST + dance-MID.PST-CONJ fuck-HAB-2PL
 ‘Do you sing, dance, and then fuck?’ [C.85]

In a common usage corresponding to Desia *kis-ʔa* ‘what?’, *maŋ* is found combined with *-ʔa*, seemingly the Oriya ‘article’ discussed at some length by Neukom and Patnaik (2003:22ff.):²⁰

- (25) *nom maŋʔa seb-tu, o-nij ar-kupei-a teŋgia.*
 2SG what slaughter-ACT.NPST OBJ-1SG NEG-cut-NEG.NPST axe
 ‘What will you slaughter, the axe will not cut me.’ [B.76]

Indefinites are derived from these by adding the morphemes *-sa* ‘and’, *-ki* Q:

- (26) *umboʔ-sa umboʔ-naj moq-gu i-a.*
 where-and where-1PL get.up-MID.PST AUX(go)-IMP
 ‘Let’s get up and go somewhere!’ [5.392]
- (27) *laj-sa tenu kondek goi-oʔ beq-oʔ-na goi-tu-nij*
 who-and that a.bit cut-PST AUX(give)-PST-CV cut-ACT.NPST-1SG
iaŋ-su djeʔ sun-oʔ-nij
 IRR-CONJ QUOT say-ACT.PST-1SG
 ‘If anyone cuts that little piece [of land], I should cut it, that’s what I said.’ [5.23]

An example of the use of *-ki* has been given above in section 3.1.4 (example 14). Here is another one:

- (28) *maŋ-ki deʔ sun-oʔ-nom, u samo sun.*
 what-Q QUOT say-ACT.PST-2SG that story tell
 ‘Whatever you said, tell that story.’ [5.24]

But interrogatives can also function as indefinite pronouns without such an addition:

- (29) *laj-nu imi gibir bai, laj-nu imi gusoʔ bai, ...,*
 who-GEN name pig brother who-GEN name dog brother
tu lok-nen sobu kimboj riŋ-oʔ-nen.
 those people-PL all wife bring-ACT.PST-PL
 ‘Someone’s name was Pig Brother, someone’s name was Dog Brother, ...
 those people all brought a wife.’²¹ [11.15]

Negative indefinite pronouns (‘nobody’, ‘never’, etc.) are simply formed by the combination of interrogative pronoun and a negation, just as in Oriya (see Neukom and Patnaik 2003:100–104). Cf. example (9), and the following:

- (30) *maŋta sasta uraʔ.*
 what cheap NEG
 ‘Nothing is cheap.’ [A.64]

The only proper indefinite pronoun is the Desia borrowing *kisi* (Oriya *kichi*) which corresponds in the following example to Gutob *maŋta*:

- (31) *goj-gi ui-gi-na tu oʔn-lai kisi*
 die-MID.PST AUX(go)-MID.PST-CV that daughter-OBJ anything
milei uraʔ oron + bostor
 be.available NEG meal + clothes
 ‘When they had died, that daughter didn’t have anything, food or clothes.’ [10.4]

3.1.1.7 Demonstratives and other deictics

In the analysis adopted here, Gutob has a three-way demonstrative system (Zide 1991:353f.):²²

- | | | | | |
|------|------------|-----------|--------------|---------|
| (32) | | Proximate | Intermediate | Remote |
| | main bases | e- | u- | tu- |
| | derived | eke | | utu/otu |
| | expressive | | | ha-/hu- |

(33) About the beginning of time:

eke zu-oʔ-na uraʔ, otu zu-oʔ-na uraʔ.
 here see-ACTPST-CV NEG there see-ACT.PST-CV NEG

'If you looked here: no(thing). If you looked there: no(thing).' [6.2]

To the three main bases, which can occur as free-standing forms, additional morphemes can be added.

-toŋ: the semantic specifics are unclear, but this suffix seems to be added primarily to deictic bases with reference to humans (see Zide 1972:509). For example, *e-toŋ-nen*, *u-toŋ-nen*, *tu-toŋ-nen* (also to the third person marker: *o-maj-toŋ-nen*) – 'these, those, those there, they'.

-noʔ: indicates place, for example, *e-noʔ*, *u-noʔ*, *tu-noʔ*. This is presumably the same morpheme as in *monoʔ* 'where?', and one might presume – in view of the other known cases of /l/~/n/ interchange (see section 2.3) – that there is a connection with *eloʔ*, *uloʔ*, *tuloʔ* '(come) here, there, yonder', *huloʔ* '(look) there'. Rajan and Rajan (2001b:16), on the other hand, interpret *-loʔ* as 'generic' and *-noʔ* as 'specific' marker.²³

-nu: meaning unclear, *e-nu*, *u-nu*, *tu-nu*, 'this, that, that there'. Zide (1991:353) calls this a 'nominalizer postbase', which suggests that he may have thought of a connection with the *-nu* suffix, one of whose main uses is in nominalization.

I am not yet sure where to place in this system the forms *aloʔ* '??', *teloʔ* 'there (?)', and *utun* 'this' or 'that' (?). The second occurs, for example, in the following:

(34) *kaligai-nu ig-taŋ riŋ-oʔ-su ispor + mapru-lai*
 black.cow-GEN dung-cow bring-ACT.PST-CONJ Shiva-lord-OBJ

teloʔ riŋ-oʔ-niŋ
 there(?) bring-ACT.PST-1SG

'Having brought the dung of the Black Cow, I brought it there for the Lord Shiva.' [6.8]

3.1.8 Numerals (including cardinal, ordinal, distributive, collective, and classifiers)

Except for the numeral *muiroʔ* 'one', and possible vestiges of the inherited word for 'two' such as in *bumuiay* 'two brothers' (from *buiay* 'brother'), numerals borrowed from Desia have completely ousted native numerals in present-day Gutob. In the 1930s, Ramamurti was still able to record the following native numerals (1938:xviii–xx).²⁴

1	<i>muiroʔ</i>	7	<i>gi:l</i>
2	<i>umbar, mar, umbar, ummar</i>	8	<i>tam-gi</i>
3	<i>iŋen iʔgen, iʔgen-roʔ</i>	9	<i>tim-gi</i>
4	<i>uʔu:n (uʔn)</i>	10	<i>gol</i>
5	<i>mallai</i>	11	<i>gol-mui</i>
6	<i>tir</i>	12	<i>gol-mba:r</i>

However, in his letter from Koraput to F.B.J. Kuiper of 31 October 1951, Izikowitz already claims that he could find only the numerals 1–3: ‘Funny enough they can only count to 3 in their own language, the rest goes in Oriya. 1,2,3 is *mūiro*, *nībār* and ‘igen.’ is the glottal stop.’ Norman Zide’s fieldwork in the 1960s still yielded these same three numerals (Zide 1978:51). Subba Rao (1992:17) states for Andhra Gutob: ‘Gadaba language has numerals only up to five. Counting above five is done by way of addition and multiplication.’ He records: one = /mu:ỹu/, two = /mba:ru/, three = /igge:nu/, four /u:nu/ (or /pu:nja/, see just below), five = /moley/ – these five items confirm Ramamurti’s data.

Beyond the above mentioned word for ‘twelve’ we have no information on native counting up to 20 (and onwards).²⁵ As for Desia, we need not recount the whole list of numerals (for which, see Gustafsson 1989:1007ff.), but may point out only the following expressions, *goṭek* ‘one unit = one’ (next to *ek*), *zoḍek* ‘one pair’ (next to *dilḍui* ‘two’), *punzek* ‘one heap = four’ (next to *sar*, see Zide 1978:51) and *di-punza* ‘two heaps = eight’ (next to *aṭ*), which nicely exemplify the same system that is also pervasive for higher numbers, namely the use of multiples of *koḍi* ‘score’.

As to ordinals, *aṭtulagtu* ‘early, ahead, in front’ can be used in the meaning ‘first’. At present, I have no information about strategies for making higher ordinals – the addition of *-o:* to the cardinal as described by Subba Rao (1992:19) for Andhra Gutob resembles nothing known to me from Koraput. Information is also not available on distributives and collectives.

The language does not have an extensive set of numeral classifiers, and those words which do seem to function in such a way tend to be borrowings from Desia:

- humans: *-ḍan*, as in *mūḍan* ‘alone’, ‘one man’ (apparently only with ‘one’)
 humans: *lok* (De.)
 young people: *rasi* (De. *rasi* ‘herd?’)
 cattle (‘a head of cattle’): *munḍ* (De.)

A frequently heard element after numerals, numeral adverbs like *ketek* ‘how many?’ (De.), or classifiers is the suffix *-laka* (which I tentatively gloss here with EMPH, not pretending that this throws any light on the morpheme’s function):

- (35) *muiro?* *ḍuma-pulai* *umar-laka* *boṅtel* *tol-to-nen* *mui-laka*
 one ghost-OBJ two-EMPH buffalo tie.up-HAB-PL one-EMPH
buḍa *ḍor-to-nen* *mui-laka* *ḍuma-pulai*
 cattle kill-HAB-PL one-EMPH ghost-OBJ
 ‘For one ghost, they tie up two water-buffaloes; they kill one [head of] cattle for one ghost.’ [F.3]
- (36) *dinke* *ketek-laka* *beḍ-beḍ-nen* *ḍu-gu?*
 per.day how.much-EMPH REDPL-give-PL AUX(be)-MID.PST
 ‘How much did they give per day?’ [18.7]
- (37) *koḍe-ṭa* *ḍuma* *ḍeṅ-gu-na* *salis* *munḍ-laka* *boṅtel*
 twenty-ART ghost become-MID.PST-CV forty head-EMPH buffalo
tol-to-nen.
 tie.up-HAB-PL
 ‘If there are 20 ghosts, they tie up 40 buffaloes.’ [F.4]

3.1.9 Adpositions

Gutob has a considerable number of postpositions, and no native prepositions. Postpositions differ from case markers in that they can be added to the nominal marked for genitive case. The distribution of cases where GEN suffix *-nu* does and does not intervene has not yet been worked out, but may depend on animacy. Below, I list the most important postpositions in roughly (Roman) alphabetical order. I have attempted to illustrate their meanings using examples both with GEN-marking (where available in the database) and without.

ali ‘near’:

- (38) *and̩ei-gu piy-gi tu onob-nu ali tu oʔn-lai*
 return-MID.PST AUX(come)-MID.PST that girl-GEN near that child-OBJ
salag-oʔ: ‘umboʔ-nom ui-nom?’
 ask-ACT.PST whither-2SG go-2SG
 ‘When he had come back near that girl, she asked that boy, “where are you going?”.’

aluŋ ‘inside’:

- (39) *tubog aluŋ ɖu-to ki tobnaŋ ɖu-toʔ*
 earth inside be-HAB or above be-HAB
 ‘Is it in the earth, or above?’ [2.367]

buɖoŋ/bodoŋ ‘with’:

- (40) A man talking about problems after having sex:
koʃte toʔ-gu-niŋ, soʔl + siʔl buɖoŋ
 with.difficulty pull.out-MID.PST-1SG oil + ECHO with
amroj-gu-niŋ suŋ-oʔ
 smear-MID.PST-1SG AUX(throw)-ACT.PST-1SG
 ‘I pulled it out with difficulty. I smeared it with oil etc.’ [C.28]
- (41) *aʔ niŋ amrika saibo-nu buɖoŋ nondpur niŋ ɖu-tu,*
 now 1SG America saheb-GEN with Nandpur 1SG be-NPST
aʔ ‘puri i-a-naj’ ɖieʔ sun-oʔ-nen-ni.
 now Puri go-IMP-1PL QUOT say-ACT.PST-PL-*ni*
 ‘Now I am with the America gentlemen [in] Nandpur; now they are saying, “Let’s go to Puri”.’ [E.6]

boʔ ‘at, in, to’ (LOC) and *ɖiʔke* ‘from’:

- (42) *niŋ ui-gi-niŋ ɖien boʔ sonek, su tu ɖiʔke*
 1SG go-MID.PST-1SG house LOC a.moment CONJ that from
mĩõʔ niŋ piŋ-gi-niŋ
 again 1SG come-MID.PST-1SG
 ‘I went home for a while, and then I came [back] from there again.’ [4.5]
- (43) *muiroʔ ioŋ-nu ɖiʔke, muiroʔ apoŋ-nu ɖiʔke zonom + zat*
 one mother-GEN from one father-GEN from birth + ECHO
naj ɖem-oʔ-naj ɖu-tu ze, nom goj-gi-na niŋ maŋ
 1PL do-ACT.PST-1PL be-ACT.NPST SO, 2SG die-MID.PST-CV 1SG what

ḍu-gu-nu leka, niḡ goj-gi-na nom maḡ
 AUX(=be)-MID.PST-GEN account 1SG die-MID.PST-CV 2SG what

ḍu-gu-nu leka?
 AUX(=be)-MID.PST-GEN account

‘From one mother, from one father we have been born, so if you die, what is to become of me, and if I die, what is to become of you?’ [5.361]

For the latter postposition, we also find *ḍig*, perhaps limited to inanimate nouns:

- (44) *maj obsei-o? sun-o?-nu sādi ḍalni ḍig ḍi?to*
 3SG show-ACT.PST say-ACT.PST-GEN ancestral acre from QUOT
mala + ṭoila ḍi?to koḍok + kuḍuk kaṭ + kuṭ gorṇḍei + girṇḍei
 wood + ECHO QUOT hoe.work + ECHO cutting + ECHO trimming + ECHO
ḍem-o? suḡ-o? ḍi?to
 do-ACT.PST AUX(throw)-ACT.PST QUOT
 ‘He got the hoe-work, the cutting, the trimming done on the wood from the ancestral acre which had been shown [and] mentioned.’ [5.34]

kuruy ‘towards’/‘at’, see example (73) and the following:

- (45) *gi?ḍaḡ-siḡ kuruy oḡ-o? ḍu-gu ḍie?na tu ḍi?ke*
 back-house at listen-ACT.PST AUX(=be)-PST QUOT that from
tur-gu piḡ-gi ḍi?to.
 leave-MID.PST AUX(=come)-MID.PST QUOT
 ‘When he had heard it at the back of the house, he came out from there.’ [5.155]

lagire ‘due to’: see example (12).

munay ‘like, by way of’:

- (46) *gaṭio? tur-to ... niḡ-sa pi-loḡ-niḡ zu-o?-nu*
 a.lot come.out-HAB 1SG-too come-MID.NPST-1SG see-ACT.PST-GEN
munay muson
 like one.day
 ‘A lot is displayed. I too will come one day for seeing.’ [A.41]

o?ḍon ‘near’:²⁶

- (47) *tu gikil-nu o?ḍon barogaṭia ui-gi.*
 that tiger-GEN near B. go-MID.PST
 ‘Barogatia went near that tiger.’ [10.53]

orbon ‘near’:

- (48) *niḡ maj-nu orbon piḡ-gi-niḡ*
 1SG 3SG-GEN near come-MID.PST-1SG
 ‘I came close to him.’ [N. Zide/B.P. Das, field data (notebook N. morphology)]

pulai ‘in order to, for’ (OBJ), see section 3.1.2, and:²⁷

- (49) *niḡ sun-tu-niḡ onob-rasi-pulai*
 1SG say-ACT.NPST-1SG girl-CLSSFR-OBJ
 ‘I will tell the girls.’ [A.129]

sumoŋ ‘in front of’:

- (50) *die?su, tu o?n odorpođ deŋ-gu, ioŋ-đei-nen-nu*
 QUOT that child *odorpođ* become-MID.PST mother-3POSS-PL-GEN
sumoŋ boro? đu-tu.
 in.front cry AUX(=be)-PRS
 ‘And so, that child became *odorpođ*, and is crying in front of its mother and
 the rest.’ [15.62]

tobnaŋ ‘above’: see example (39).

3.1.10 Derivation

The most common means in Gutob to derive nouns from nouns or other parts of speech is the use of the nominalizing suffix *-kaŋ*, glossed NMLZ in this description (after Rajan and Rajan 2001b:2f. and 17).

- (51) *buron-kaŋ riŋ-to-pen ki usor-kaŋ?*
 live-NMLZ bring-HAB-2PL or be.dry-NMLZ
 ‘Do you bring green or dry wood?’ [A.11]
- (52) *lok, maŋ lok runđei-loŋ-nen? uŋgom-kaŋ to?*
 people Q people gather-MID.NPST-PL village-NMLZ DISC
 ‘People, will people gather? Villagers, you mean?’ [2.280]

There are several other native processes of nominal derivation clearly evident in the lexicon, though none of them seems to be productive in the current form of the language (this needs to be verified).

3.1.10.1 Deverbatives/nominalizers (-Vn-)

There are many examples of the infix *-Vn-*, which is used to derive nouns from verb roots.²⁸ It is infixed before the initial or only vowel of such roots. The quality of the vowel is determined by vowel harmony: the infix takes the shape *-in-* before front vowel and *-un-* before central or back vowel. In vowel-initial roots, the infix becomes in effect a prefix, but here the form *an-* occurs before /a/ in the root. The following list contains all examples I have found thus far.

- | | |
|---|--|
| (53) <i>an-ab</i> ‘husk’ (<i>ab</i> ‘to husk’) | <i>s-in-id</i> ‘bird-lime, gum’ (<i>sid</i> ‘to
bird-lime’) |
| <i>b-in-i?tir</i> ‘saliva’ (<i>bi?tir</i> ‘to spit’) | <i>s-un-ag</i> ‘loan’ (<i>sag</i> ‘to tear’) |
| <i>b-un-aj</i> ‘writing, arithmetic,
embroidery’ (<i>baj</i> ‘to decorate’) | <i>s-un-ar</i> ‘comb’ (<i>sar</i> ‘to comb’)
(<i>rug</i> ‘to open’) |
| <i>b-un-o?</i> ‘ladder’ (<i>boklbo?</i> ‘?’) | <i>s-un-og</i> ‘broom’ (<i>sog</i> ‘to brush’) |
| <i>g-in-ir</i> ‘fishing net’ (<i>gir</i> ‘to fish’) | <i>s-un-oi</i> ‘plowshare’ (<i>sui</i> ‘to plow’) ²⁹ |
| <i>g-in-ira?</i> ‘spindle’ (<i>gira?</i> ‘to spin’) | <i>t-in-el</i> ‘threshing-floor’ (<i>tel</i> ‘?’) ³⁰ |
| <i>g-un-ug</i> ‘hatchet’ (<i>gug</i> ‘to peck’) | <i>t-un-ol(+bo?)</i> ‘(head)band’
(<i>tol</i> ‘to tie’) |
| <i>in-iŋ</i> ‘rope’ (<i>iŋ</i> ‘to hang’) | <i>z-un-uŋ</i> ‘clothesline’ (<i>zuŋ</i> ‘to
suspend’) |
| <i>p-in-eđ</i> ‘flute’ (<i>peđ</i> ‘to blow’) | |
| <i>r-un-aŋ</i> ‘medicine’ (<i>raŋ</i> ‘to concoct?’) | |
| <i>r-un-ug</i> ‘doorway’ | |

I find one case of a derivation which also seems to involve reduplication (see Peterson on Kharia in this volume): *s-in-ij-sij* ‘spool’ from *sij* ‘to wind’. A possible example of the common Austroasiatic *-Vm-* infix may be *gumul* ‘borer (insect)’ from *gul* ‘to bore’.

3.1.10.2 The animacy marker *gV-*

There is a sizeable set of words (comprising mainly animal names) formed with a prefix *gV-*, where the quality of the vowel again depends on that of the underlying root. The following is a list of all animal names belonging to this formation currently known to me:

- (54) *gi-bir/gi-big* ‘pig’ *gi-kil* ‘tiger’ *gi-li?* ‘rabbit’ *gi-me?* ‘goat’
gi-si ‘louse’ *gi-sij* ‘chicken’ *gu-bon* ‘bear’ *gu-ga?* ‘crow’
gu-laj ‘bull, ox’ *gu-sa?* ‘monkey’ *gu-so?* ‘dog’

The basic lexical element, forming the second syllable, can be encountered in other combinations in compounds (see below), but cannot occur as free-standing form: we may therefore assume that derivation with the *gV-* prefix was necessary at a stage of the language where the phonological system did not allow monosyllabic words. My explanation of the name Gutob as *gu-tob* (from *tubog* ‘earth’) has been given in Note 1. It does not seem impossible that also *guloj* ‘stranger, member of the Rona caste’ is derived with the same prefix from *suloj* ‘far away’. If these words have been correctly analyzed, it would mean that the semantic contribution of the prefix is wider than merely the denotation of animal names, and includes also humans: in other words, it may be taken as a marker of animacy. Depending on comparative linguistic evidence which is not available to me, such kinship terms as *gikiy* ‘wife’s elder brother (?)’ may also turn out to belong to this formation.

3.1.10.3 A prefix *sV-*?

Not having access to any evidence which would disprove or confirm such a hypothesis (suggested already by Anderson and Zide 2002), I also add here a list of words which may be taken to have been formed with a prefix *sV-*, again showing vowel harmony between the first and second syllable. The meaning of this prefix is not clear.

- (55) *silen* (only in *silen-qa?* ‘sweat’) *silej* ‘long’ *silim* ‘?’
simin ‘day-time, day’³¹ *sisaj* ‘bone’³² *subul* ‘sweet’
sukug ‘gourd’ *sujmol* ‘seed/pit’ *suloj* ‘far away’
sumol ‘pure’ *suram* ‘antelope’ *suloj* ‘stomach’
sulob ‘tree’ *sumoj* ‘in front of’ *susuj* ‘leg’³³

3.1.10.4 Nominal combining forms

This section must be read in conjunction with section 3.2.11, for the processes involved in deriving nominal combining forms for incorporation into nominal compounds are the same as those for incorporation into verbs. Although one may describe most combining forms as reduced monosyllabic counterparts of full free-standing forms, historical arguments in some cases suggest rather that the free-standing forms are expanded versions of roots (= combining forms); there are also cases where an etymological connection between full and combining form is not evident at all.³⁴

- (56) *raj-saj* ‘cremation ground’ *raj* ‘?’ *si-saj* ‘bone’
ig-li ‘ear-wax’ *ig* ‘excrement’ *litir* ‘ear’
utob-sij ‘chicken egg’ *utob* ‘egg, ball’ *gi-sij* ‘chicken’
iʔ-taj ‘cow-dung’ *ig* ‘excrement’ *kitaj* ‘cow’
giʔdaj-sij ‘back of the house’, with *sij* CF of *dien* ‘house’

3.1.11 Adjectives

The language does not seem to have a separate part of speech ‘adjective’. Many lexical items used for nominal attribution can take verbal morphology, and can by that criterion be treated as verbs (Zide 1985:97).

According to my field data, comparatives and superlatives are expressed in typical South Asian fashion by means of the ‘ablative’ postposition *qiʔke*: *X qiʔke* ADJ ‘ADJ-er than X’ and *sobu qiʔke* ADJ ‘the ADJ-est of all’. The database at my disposal, however, contains no example sentences with this construction.

3.1.12 Adverb(ial)s

Rajan and Rajan (2001b:19 [section 3.3]) distinguish ‘two types of adverbs, the first one occurs as a single word and the other occurs as an onomatopoeic word and it can occur more than [*sic*] once in the VP’. The following are their two examples:

- (57) *sida* *ui-gi*
straight go-MID.PST
‘He went straight ahead.’ (Rajan and Rajan 2001b:19)
- (58) *maj ka ka ka dɛŋ-gu-su* *maj laɖu som-oʔ*
3SG ka ka ka become-MID.PST-CONJ 3SG ladu eat-ACT.PST
‘It (the crow) cried *ka ka ka* and ate the ladu.’ (Rajan and Rajan 2001b:19)

3.2 Verbal morphology

As in other Munda languages, the distinction between nouns and verbs is not very strict in Gutob. Take the following example, where the borrowed word corresponding to the Oriya noun *mahājana* ‘important man’ is inflected as a verb:

- (59) *mazon-gu-nom*
important.man-MID.PST-2SG
‘You became an important man.’ [10.159]

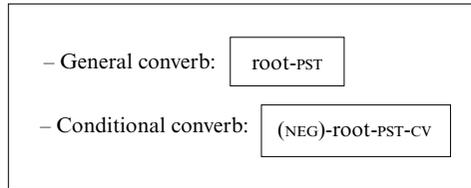
The morphology of the Gutob verb can be described with the help of a few simple canonical shapes:

BOX 12.1: POSITIVE FINITE VERB FORMS

(CAUS-)root-suffix(TENSE/VOICE/ASPECT/MOOD)-SUBJECT.PRONOMINAL

REDPL-FOOT-SUBJECT.PRONOMINAL

BOX 12.2: CONVERBS



3.2.1 Subject

Subject is marked by (clitic) pronominals generally placed in the verb phrase; the rules for the exact placement of the pronominal seem to be quite free, or if they are not, they have as yet defied attempts at analysis (Zide 1997:308ff.). These pronominals have the same shape as personal pronouns (section 3.1.6.1), except that the third person singular is Ø-marked and third person plural (optionally) marked by the general plural marker *-nen*.

(60)	SG	PL
1	niŋ	nei/naj
2	nom	pen
3	Ø	-nen

I have the impression that repeated subject marking is a feature of expressive discourse:

(61)	<i>niŋ</i>	<i>ta</i>	<i>sobu</i>	<i>paiŋi</i>	<i>niŋ</i>	<i>dem-oŋ-niŋ</i>	<i>beŋ-beŋ-niŋ</i>
	1SG	DISC	all	work	1SG	do-ACT.PST-1SG	REDPL-AUX(give)-1SG
	'I am doing all the work for you.' [5.57]						

3.2.2 Object types

Encoding of objects pronominally in the verb, as is seen in various other Munda languages, plays no role in Gutob.

3.2.3–3.2.7 Tense, aspect, mood, orientation/directionality, valence/voice

In Gutob, the meanings of the same TAM-morphemes are different in negative verb forms from their meanings in the corresponding positive forms. As Zide and Anderson (2001:537, n. 4) have observed:

[t]he negative in Gutob is quite complex. Very few categories use the same markers in the negative as the positive. Thus, in the positive, *-to* marks a customary present [here called 'habitual', A.G.], but in the negative it marks past tense: *ser-to* sing-CUSTOMARY.PRESENT 'sings' vs *ar-ser-to* NEG-sing-NEGATIVE.PAST 'didn't sing' (N. Zide field notes).

The following table shows the usages of the various morphemes:³⁵

(62)

	Positive		Negative	
	Middle	Active	Middle	Active
NPST	-loŋ	-tu	-a	-Ø
PST	-gV	-oʔ		-to
IMP	-a	-Ø	-gV	-oʔ
HAB		-to		—
OPT		-e		-e

The basic auxiliary is *ɖu(k)* ‘to be’, which has the following rather irregular positive paradigm:

(63)	<i>ɖu-tu</i>	PRS	‘is’
	<i>ɖu-loŋ</i>	FUT	‘will be’
	<i>ɖu-gu</i>	PST	‘was’
	<i>ɖuk-a</i>	IMP	‘be!’
	<i>ɖu-to</i>	HAB	‘is’
	<i>ɖik-e</i>	OPT	‘may be’
	<i>ɖu</i>	INF	‘(to) be’

Note also the following negative forms: *uraʔ* ‘is not’, *arɖuto* ‘was not’, *arɖuka* ‘will not be’, *arɖike* ‘may not be’. With the above suffixes, and these forms of the auxiliary, we can now build a paradigmatic overview. Shown here are all the possible finite forms (with unmarked third singular subject) of the intransitive verb *ser* ‘to sing’ and the transitive verb *som* ‘to eat’.

(64)

		Positive		Negative	
		Middle	Active	Middle	Active
1	INF	s eser	s omsom	s eser uraʔ	s omsom uraʔ
2	NPST	s erloŋ	s omstu	s arsera	s arsom
3	PST	s ergu	s omoʔ	s arserto	s arsomto
4	HAB	s erto	s omto	s eser uraʔ	s omsom uraʔ
5	IMP	s era	s om	s arsergu	s arsom oʔ
6	OPT	s ere	s ome	s arsere	s arsome
7a	PRS PRF	s ergu ɖutu	s omoʔ ɖutu	s sergunu uraʔ	s omoʔ nu uraʔ
7b				s ergu ɖu uraʔ	s omoʔ ɖu uraʔ
8a	PST PRF	s ergu ɖugu	s omoʔ ɖugu	s eser arɖuto	s omsom arɖuto
8b				s ergu ɖu uraʔ	s omoʔ ɖu uraʔ
8c				s sergunu uraʔ	s omoʔ nu uraʔ
9	FUT PRF	s ergu ɖuloŋ	s omoʔ ɖuloŋ	?	?
10	HAB PRF	s ergu ɖuto	s omoʔ ɖuto	?	?
11	PRS PROG ₁	s eser ɖutu	s omsom ɖutu	s eser uraʔ	s omsom uraʔ

CONTINUED

		Positive		Negative	
		Middle	Active	Middle	Active
12	PST PROG	s eser d̥ugu	d̥ugu	s eser ura? d̥ugu	s omsom ura? d̥ugu
13a	FUT PROG	s eser d̥uloŋ	s omsom d̥uloŋ	s eser ura? d̥uloŋ	s omsom ura? d̥uloŋ
13b				s eser ar̥duka	s omsom ar̥duka
14	HAB PROG	s eser d̥uto	s omsom d̥uto	s eser ura?	s omsom ura?
15	IMP PROG	s eser d̥uka	s omsom d̥uka	s eser ar̥d̥ugu	s omsom ar̥d̥ugu
16	OPT PROG	s eser d̥jike	s omsom d̥jike	s eser ar̥d̥jike	s omsom ar̥d̥jike
17	PRS PROG ₂	s erguni	s omo?ni	s eser ura?	s omsom ura?

A construction with the pattern REDPL-root-PST-PR.S is of uncertain status. In my field notes I have written *zuzuo?nen* = *zuzunen d̥utu* [REDPL-look-3PL AUX: PRS.PROG₁] ‘they are looking’. Examples in Norman Zide’s text collection, however, suggest rather a preterite (perfect) meaning:

- (65) *asam ui-gi-nu lok, mayta sun-tu-nom*
 Assam go-MID.PST-GEN people what say-ACT.NPST-2SG
luga + paŋa ... o ... taŋka + tusa die?-nu
 clothes + ECHO, oh, money + ECHO QUOT-GEN
munaj riŋ-riŋ-o?-nen kaŋ asam
 like REDPL-bring-ACT.PST-PL kaŋ Assam
ui-gi-nu lok riŋ-o?-nen.
 go-MID.PST-GEN people bring-ACT.PST-PL
 ‘The people who went to Assam, what will you say, clothes, ... , oh, ... things
 like money they have brought, the people who went to Assam, they brought.’
 [18.16]

The negative forms of the FUT.PFT and the HAB.PFT are not yet known (due to gaps in field data). In the following paragraphs, the positive forms are taken as the basis of discussion; for examples of negative forms, see section 3.2.9.

3.2.3 Tense

Gutob has two basic tense categories, ‘non-past’ and ‘past’, the former being morphologically divided into a ‘present’ and a ‘future’ only in the case of the verb *du(k)* ‘to be’ (see just above). For the use of the NPST, see example (66). Examples of the PST are strewn throughout this sketch.

3.2.4 Aspect

Among the simple verb forms in the table (rows 4–5), the reduplicated root dubbed ‘infinitival present’ (with a roughly ‘continuous’ meaning) and the form in *-to*,

dubbed ‘habitual’, express aspectual values. The other aspect categories, ‘perfective’ and ‘progressive,’ are expressed with complex constructions. A selection of forms is exemplified below. First contrast ‘infinitival present’ with ‘non-past’:

- (66) A: *maŋdem mĩõ? nom piŋ-gi be?*
 why again 2SG come-MID.PST POL
- B: *niŋ ta sobu paĩti niŋ dem-o?-niŋ*
 1SG DISC all work 1SG do-ACT.PST-1SG
beq-beq-niŋ
 REDPL-AUX(give)-1SG
- A: *nom maŋdem piŋ-gi-nom?*
 2SG why come-MID.PST-2SG
- B: *niŋ to zaiŋa nom mon dem-o?-na,*
 1SG DISC whatever 2SG mind make-ACT.PST-CV
unu niŋ dem-o? beq-tu.
 that 1SG do-ACT.PST AUX(give)-ACT.NPST
- A: ‘Why have you come again?’ B: ‘I am doing all the work for you.’
 A: ‘Why have you come?’ B: ‘Whatever you have in mind, I will do it for you.’ [5.56–59]

Past forms followed by the auxiliary ‘to be’ yield perfectives. Here are examples of a ‘habitual perfective’ and a ‘present perfective’ form:

- (67) *simra-gu o? qu-to-nen, diŋsel.*
 become.silent-MID.PST EMPH AUX(be)-HAB-PL friend
 ‘[Such women] really keep silent, friend.’ [A.119]
- (68) *kinða? tol-o?-nen qu-tu, eran dem-o??*
 river dam-ACT.PST-PL AUX(be)-PRS how do-ACT.PST
 ‘They have dammed the river, how so?’ [A.85]

Reduplicated forms followed by the same auxiliary ‘to be’ yield progressives:

- (69) *ura?, naik + barik sobu su-sun-nen qu-tu.*
 no headman + messenger all REDPL-say-PL AUX(be)-PRS
 ‘No, the headman and the messenger, they are all saying so.’ [B.24]
- (70) *so din bo? dos taŋka, ura?-na agar taŋka-laka riŋ-riŋ-nei*
 6 day LOC 10 rupee NEG-CV 11 rupee-EMPH REDPL-take-1PL
qu-gu, ranđe, tubog o? ob-tur terepete.
 AUX(be)-PST you.know earth EMPH CAUS-emerge with.difficulty
 ‘In six days we were taking ten rupees, if not eleven rupees, you know, digging up earth with difficulty.’ [18.8]

The language also has an alternative construction for the ‘present progressive’ (calqued on the Desia construction described by Mahapatra 1985:75), making use of a borrowed morpheme *-ni*: root-PST-(PR.S)-*ni* (row 17 in example 64). See, *sun-o?-nen-ni* in example (41), and *ob-gu-ni* ‘is biting’ in (77).

3.2.5 *Mood*

Two of the simple forms (rows 5 and 6) express modality, here dubbed ‘imperative’ and ‘optative’, respectively:

- (71) *muiro? kui naj to-o?-su tu kui bo?*
 one well 1PL dig.out-ACT.PST-CONJ that well LOC
ḍuḍuḍ-gu goj-gi i-a-naj
 drown-MID.PST die-MID.PST AUX(GO)-IMP-1PL
 ‘Let’s dig out a well, drown ourselves in it, and die!’ [9.8]

- (72) *nor zonom ḍem-e-nen*
 man birth do-OPT-PL
 ‘May men be born!’ [6.14]

The (clitic?) particle *iaḡ* (IRR) marks counterfactuality:

- (73) *niḡ-nu kuruy ḍa? ḍu-gu-na, niḡ ḍa? i?-tu-niḡ iaḡ*
 I-GEN way water be-PST-CV I water drink-ACT.NPST-1SG IRR
 ‘If I had water, I would drink water.’ [DeArmond n.d.:9]

The precise function of the (clitic?) particles *ḍoḡ* and *ḍej* has to be worked out with further field data. The currently available text collection provides no clarity, but it seems their grammatical behavior and function fall within the domain of modality:

- (74) *maj mĩḍ? bin-kaḡ beḍ-tu ḍej.*
 3SG again other-NMLZ give-ACT.NPST ḍej
 ‘He might (?) give something else.’ [1.178]
- (75) *ito? ḍe?na pi-loḡ-niḡ ḍoḡ.*
 thus QUOT come-MID.NPST-1SG ḍoḡ
 ‘In that case I can (?) come.’ [A.103]

3.2.6 *Orientation/directionality*

These concepts seem to play no morphological role in the grammar of Gutob, their expression being achieved through auxiliary verb constructions as discussed in section 3.2.12 below.

3.2.7 *Valence/voice*

Through a double set of TAM-markers, Gutob morphologically distinguishes two ‘voice’ classes, called ‘active’ and ‘middle’ in this description; these two classes correspond rather closely with semantic transitivity and intransitivity, respectively.

It is noteworthy that semantically transitive verbs, which in other contexts take ACT inflection, take MID endings when reflexively combined with body parts of the speaker as object(s). See example (40) above, and:

- (76) *kunḍiḡ-sa ra?-gu majki ura? ze.*
 vagina-and tear-MID.PST wife not that
 ‘And my wife₁ had torn her₁ vagina, indeed.’ [C.36]

- (77) *maj aŋti-lai ob-gu-ni.*
 3SG finger-OBJ bite-MID.PST-*ni*
 ‘He₁’s biting his₁ finger.’ (Rajan and Rajan 2001b:17)

The connection of MID inflection with semantically transitive verbs is not limited to reflexivity, but is also found in the case of reciprocals, which explains why the CAUS-marking in the following example does not trigger ACT inflection (*pace* Zide 1985:98):

- (78) *imi ob-oŋ-gu-nen*
 name CAUS-hear-MID.PST-PL
 ‘They made each other hear names.’ [11.14]

3.2.8 (Non-)finiteness

The language does not seem to enforce a very strict distinction between finite and non-finite verb forms; at least there are no positively marked non-finite forms. In a non-third person context, the absence of pronominal marking can be used as a criterion for non-finiteness; but even the ‘(conditional) converb’ marked with *-na* (see section 4.7), which would seem to be one of the most clearly non-finite categories, fails this criterion in a rare case like the following (a partial repetition from example 66), where the *-na* phrase is marked with pronominal *nom* (though not as suffix):

- (79) *niŋ to zaiŋa nom mon dem-o?-na, unu niŋ*
 1SG DISC whatever 2SG mind make-ACT.PST-CV that 1SG
dem-o? beŋ-tu.
 do-ACT.PST AUX(give)-ACT.NPST
 ‘Whatever you have in mind, I will do it for you.’ [5.59]

The negative copula has a form *ura?-gu* (the precise difference in usage from *arđuto* still needs to be worked out) that seems to be found only in direct or nearly direct combination with *-na*:

- (80) *niŋ ura?-gu de?-na-sori o-nom sob-o?*
 I not.be-PST QUOT-CV-*sori* OBJ-2SG carry.away-ACT.PST
ui-loŋ iaŋ be die? di?to
 AUX(go)-MID.NPST IRR POL QUOT QUOT
 ‘Had I not been there, he would have carried you away.’ [5.374]
- (81) *a? da? ura?-gu-na mono?-nu keroŋ de-loŋ?*
 now water not.be-PST-CV where-GEN paddy become-MID.NPST
 ‘If there is no water now, from where will the paddy come?’ [D.29]

Using the above criterion of absent person marking, the ‘general converb’ (formally identical with a third person singular PST, see above) is the only other certain non-finite category known to me at present. See the general converb *đuđuŋgu* in example (71) above, and

- (82) *su, niŋ majnen bođoŋ đu-gu ui-gi-niŋ.*
 CONJ 1SG 3PL with be-MID.PST go-MID.PST-1SG
 ‘Then, I went after having stayed with them.’ [3.9]

- (83) *teŋgia riŋ-o?* *seb-tu-niŋ.*
 axe take-ACT.PST slaughter-ACT.NPST-1SG
 'I will take an axe and slaughter [you].' [B.79]

The bare or reduplicated form of the verb (3.2.10.1) serves, among other functions, to derive a form of the verb here called 'infinitive': see examples (19) and (70).

3.2.9 Negation

Gutob makes use of the following negative copulae and negative prefixes:

- (84) 'not' 'not yet'
 copula *ura?* *oroj*
 prefix *ar-* *mor-*

The copula *ura?* can at least partially be inflected as a verb (*ura?gu* 'was not'), but this inflected form seems to occur only in non-finite constructions (see examples 80 and 81). For *oroj*, of which no inflected forms occur in the available data, see the following:³⁶

- (85) ... *ebke andei aka niŋ oroj*
 up.to.now come.back EMPH 1SG not.yet
 '... up to now, I haven't come back yet / am not coming back yet (?)' [D.78]

It seems likely that *mor-* has a fuller semantic load than *ar-* (perhaps: 'not yet, not even'):

- (86) After repeated promises:
oŋ mama, nom be?-tu-nom ki ura?
 listen father.in.law you give-ACT.NPST-2SG or not
nom-nu onoo?n-pulai? nom mor-beq-be? munay de?na
 you-GEN daughter-OBJ you NEG-REDPL-give like QUOT
o-nom tja gilei-tu-niŋ
 OBJ-2SG standing swallow-ACT.NPST-1SG
 'Listen, father-in-law, will you give her or not, your daughter? If you say that you will not yet/still not give her, I will swallow you right away ...' [5.71–72]
- (87) *ito? emran du-loŋ, gujda?-gu mor-suŋ-o?-na?*
 this.way how be-FUT wash.in.water-MID.PST NEG-AUX(throw)-ACT.PST-CV
 'How can it be like this, if you don't even wash it?' [C.53]
- (88) *oŋ, apay-nu samo mor-sob-o?-na mō? laj-nu*
 listen father-GEN word NEG-carry-ACT.PST-CV then who-GEN
samo sob-tu, a? uŋgom bitre muiro? naik
 words carry-ACT.NPST now village inside one headman
du-loŋ, naik boq + san-nu samo
 be-FUT headman big + small-GEN words
ar-sob-o?-na laj-nu samo sob-tu-nen?
 NEG-carry-ACT.PST-CV who-GEN words carry-ACT.NPST-PL
 'Listen, if he does not even accept father's words, then whose words will he accept? Now there will be one headman within the village; if the headmen (?)

do not accept the words of [the villagers] big and small, whose words will they accept?' [2.257]

The following is an example of a negative 'infinitival present' (table row 1) or present progressive (11/17) – there is a possible interpretation as negative perfective (see also example 85).

- (89) D: *mindij qa? emran deŋ-gu, qisel, nei-nu kuruj*
 this.year water how become-MID.PST friend we-GEN path
mulke pi-pij ura?
 totally REDPL-come NEG

M: *mulke pi-pij ura?, qisel, mindij ziuna pulai*
 totally REDPL-come NEG friend this year living for
boqe kosto deŋ-gu ui-gi.
 very difficult become-MID.PST AUX(go)-MID.PST

D: 'How was the rain this year, friend? On our side, nothing at all has come / is coming'

M: 'Nothing has come / is coming at all, friend – it has become very hard to live this year' [D.3–4]

- (90) *si tu-tur oroj.*
 sun REDPL-come.out not.yet
 'The sun has not risen yet / is not rising yet.' [A.G. field notes]

Negative non-past (2):

- (91) *ito?-kaŋ-nen de?na nij pi-loŋ, ura?-na, loŋ-loŋ-nu-nen*
 this way-NMLZ-PL QUOT 1SG come-NPST not.be-CV REDPL-curse-GEN-PL
de?na nij ar-pij-a.
 QUOT 1SG NEG-come-NEG.NPST
 'If they are of such type, I will come; if not, if they are [of the type] that curse, I won't come.' [A.120]

Negative past (3):

- (92) *lej-gi-nen-su-sori utu o?n di?to 'eke pi-loŋ-nij, aba*
 sit-MID.PST-PL-CONJ-sori that child QUOT here come-MID.NPST-1SG father
pi-loŋ-nij aba' die?su maj-nu ioŋ o?duo?n-?ei-lai
 come-MID.NPST-1SG father QUOT he-GEN mother son-3POSS-OBJ
di?to ar-baŋ-to. ar-baŋ-to-su maj ioŋ-?ei
 QUOT NEG-send-NEG.PST NEG-send-NEG.PST-CONJ 3SG mother-3POSS
boqoŋ ito? qien bo? du-gu.
 with like.this house LOC be-MID.PST
 'They sat there and that child said, "I will come here to father, I will come to father," but his mother did not allow her son to go. She did not allow [him] to go and in this way he stayed at home, with his mother.' [5.66–67]

Negative habitual (4):³⁷

- (93) *dek-te gonɔa?-nen ura?*
 watch-CV urinate-PL NEG
 '[Women] do not urinate while [someone] is watching.' [2.563]

Negative imperative (5):

- (94) *nom mator o-nij maŋ ar-ɔem-o? die?su di?to*
 you just OBJ-1SG what NEG-do-(NEG)ACT.IMP QUOT QUOT
maj sun-o?
 3SG say-ACT.PST
 '“Just don't do anything to me,” he said.’ [5.76]
- (95) *āhā ar-beɔ-o?-naŋ e mama. tirgig sata*
 yes NEG-give-(NEG)ACT.IMP-1SG hey father.in.law later truth
beɔ-tu-nom be die? di?to-su tur-gu ui-gi
 give-ACT.NPST-2SG POL QUOT QUOT-CONJ leave-MID.PST AUX(go)-MID.PST
di?to.
 QUOT
 ‘“Yes, let's not do [any] giving now, father-in-law. Honestly, you will give [her] later,” he said and took off.’ [5.139]

3.2.10 Derivation

There are two productive processes of verbal derivation in Gutob, namely reduplication and causative pre-/infixation.

3.2.10.1 Reduplication

Reduplication is found only with monosyllabic roots/stems and can be full or partial.³⁸ The precise constraints governing reduplication have not yet been sorted out, but the place of articulation of the root-final consonant does seem to play a role (root-final consonants that are found in both fully and partially reduplicating roots are printed in bold-face):

- C_1VC_2 roots (a): REDPL = root [C₂ can be **b, d, g, l, m, n, ŋ, ɲ**]
 C_1VC_2 roots (b): REDPL = C₁V [C₂ can be: **b, g, j, r, l, ?**]
 CV roots: REDPL = root
 VC roots: REDPL = root
 CV₁V₂ roots: REDPL = CV₁
 V₁V₂ roots: REDPL = root

3.2.10.2 Causative pre-/infix

Causative stems are derived from simple verb roots by means of the CAUS morpheme *ob-/ob-*. In the case of certain (all?)³⁹ disyllabic roots with vowel /u/ in the first syllable (and some with /i/), it is infixated into the root, replacing the /u/ (or /i/). Examples are as follows:

- buri?* ‘to be full’ *b-ob-ri?* ‘to fill’
buron ‘to live’ *b-ob-ron* ‘to cause to live’

<i>butoj</i> ‘to be scared’	<i>b-ob-toj</i> ‘to scare’
<i>bulu</i> ‘to be ripe, ripen (INTR)’	<i>b-ob-lu</i> ‘to make ripe’
<i>puḍal</i> ‘to break (INTR)’	<i>p-ob-ḍal</i> ‘to break (TR)’
<i>tunon</i> ‘to stand’	<i>t-ob-non</i> ‘to erect’
<i>ḍuḍig</i> ‘to sleep’	<i>ḍ-ob-ḍig</i> ‘to put to sleep’
<i>birim</i> ‘to bow’	<i>b-ob-rim</i> ‘to cause to bow’

In all other cases, the same morpheme *ob-* is prefixed to the unchanged simple root, occasionally with assimilation of its /b/ to the root-initial consonant. Examples:

<i>gir</i> ‘to learn’	<i>ob-gir</i> ‘to teach’
<i>som</i> ‘to eat’	<i>ob-som</i> ‘to feed’
<i>sarda</i> ‘to be joyful/pleased’	<i>o-sarda</i> ‘to make joyful, to please’
<i>moḍ</i> ‘to get up (ITR)’	<i>o-moḍ</i> ‘to get up (TR)’

Sometimes, an apparently underlying simple root does not seem to exist, for example, *ob-sej* ‘to show, to guide’ < **sej* ‘to see (?)’.

3.2.11 Noun incorporation and combining forms

Noun incorporation is not a productive process in Gutob, but there are many examples which show that it must have been in a former stage of the language. Nouns incorporated in verb stems appear in their combining form (section 3.1.10.4) and can assume the role of object, instrument or sometimes intransitive subject. Examples are:

sog-til ‘to glean grain fallen on the ground’ < *sog* ‘to sweep’, *t-in-il* ‘threshing-floor’, *iḥ-ḍom* ‘to defecate’ < *iḥ* ‘excrement’, *ḍom* ‘?’; *guj-tom* ‘to wash the mouth’ < *guj* ‘to wash’, *tumog* ‘mouth’.

3.2.12 Auxiliary verb constructions

Under this heading, I treat two constructions:

general converb + inflected AUX

infinitive + inflected form of *goyḍem*

3.2.12.1 General converb + inflected auxiliary verb

Auxiliary verb constructions of this type are structurally parallel to sequences of a general converb with a finite verb form (section 3.2.8),⁴⁰ but auxiliary constructions are characterized by extensive semantic bleaching of the second, fully inflected member of the complex construction. The following is a list of the verbs that figure as auxiliaries in this construction. Note that several auxiliaries are functionally somewhat diverse:

(96)	Verb	Lexical meaning	Functions as auxiliary verb
	<i>suŋ</i>	‘throw’	sudden motion; completive
	<i>riŋ</i>	‘take’, ‘bring’	self-benefactive
	<i>ui</i>	‘go’	translocative/itive; completive
	<i>piŋ</i>	‘come’	cislocative/ventive
	<i>beḍ</i>	‘give’	benefactive; completive

<i>sarei</i>	‘finish’	telicizer
<i>lagei</i>	‘apply’	inceptive

suy: see example (87) and

- (97) *ura?*, *ar-kupei-a* *o-nij*, *teŋgia pulai pigs-o?*
 no NEG-cut-NEG.NPST OBJ-1SG axe OBJ split-ACT.PST
suy-tu-nij.
 AUX(throw)-ACT.NPST-1SG
 ‘No, it will not cut me: I will break the axe in two.’ [B.80]

rij: according to Hook (1991:186f.), ‘No auxiliary plays the role played by *le* (homophonous with ‘take’) in Hindi.’ This statement, made with specific regard to *Gta?*, but apparently intended also to hold for *Gutob*, is certainly false. Consider the following counter-examples:

- (98) *o-nij mastor ai-o?* *rij-o?*.
 OBJ-1SG master call-ACT.PST AUX(take)-ACT.PST
 ‘A master₁ called me to him₁’ [3.1]
- (99) *moḍo + moḍo gisiŋ naj sob-o?* *rij-Ø*.
 big + big chicken we buy-ACT.PST AUX(take)-ACT.IMP
gime? naj sob-o? *rij-Ø* *zatra ḍen-ḍem*
 goat we buy-ACT.PST AUX(take)-ACT.IMP festival REDPL-do
lai ḍie? sarloŋ-gu-su, *zar gor-ke ta-ke*
 for QUOT talk-MID.PST-CONJ whose house-OBJ that-OBJ
tur-gu ui-to-nen.
 go.out-MID.PST AUX(go)-HAB-PL
 ‘“Let’s buy a really big chicken, let’s buy a goat, for holding the festival”: in this way they discuss, and appear at whoever’s house.’ [8.4]⁴¹

ui: see the preceding example, and

- (100) *muiro? oḍug ḍi?to goj-gi ui-gi-su*
 one boy QUOT die-MID.PST AUX(go)-MID.PST-CONJ
ḍūba ḍeŋ-gu.
 ghost become-MID.PST
 ‘A boy, they say, died and became a ghost.’ [7.1]
- (101) *nom zu-o?-na kaba-gu ui-loŋ-nom,*
 you see-ACT.PST-CV become.amazed-MID.PST AUX(go)-MID.NPST-2SG
ḍisel.
 friend
 ‘When you see it, you will become astonished, friend.’ [A.92]

piŋ:

- (102) *ui-gi-na ḍḍei-gu pi-piŋ*
 go-MID.PST-CVB return-MID.PST RED-AUX(come)
ḍe-loŋ?
 become-MID.NPST
 If we go, will it be possible to get back? [A. 77]

beq: see example (66), and

- (103) *maj sun-oʔ-su-sori tirgig mĩōʔ tu noj + el*
 he say-ACT.PST-CONJ-*sori* afterwards again that night + time
iq-oʔ-nen, som-oʔ-nen, iq-oʔ-nen, mĩōʔ quđiʔ-gu
 drink-ACT.PST-PL eat-ACT.PST-PL drink-ACT.PST-PL again sleep-MID.PST
beq-oʔ-nen.
 AUX(give)-ACT.PST-PL
 ‘He said [this], and after that, they drank again during that night, they ate,
 they drank, and they went to sleep again.’ [5.46]
- (104) *đaʔ riŋ-oʔ-su ispor + parboti-lai điʔto ai-oʔ*
 water bring-ACT.PST-CONJ Shiva + Parvati-OBJ QUOT call-ACT.PST
 ‘elo, riŋ-oʔ niŋ beq-oʔ’ đieʔ điʔto.
 come.here bring-ACT.PST 1SG AUX(give)-ACT.PST QUOT QUOT
 ‘Having brought water, he called to Shiva and Parvati: “Come here, I’ve
 brought [it] for you”.’ [6.11]

The verbs *lagei* and *sarei*, finally, are both borrowings from Desia:

- (105) *maŋta lai nom piŋ-gi lagei-gu-nom đu-tu*
 what for 2SG come-MID.PST AUX(apply)-MID.PST-2SG AUX(be)-PRS
nom unđam lok, mama, đieʔ điʔto.
 2SG old.man person father.in.law QUOT QUOT
 ‘For what reason have you started coming? You’re an old man, father-in-
 law.’ [5.111]
- (106) *lai đoŋ-oʔ sarei-gu-na gisiŋ goʔ-to-nen.*
 rice COOK-ACT.PST AUX(finish)-MID.PST-CV chicken cut.up-HAB-PL
 ‘When they have finished boiling the rice, they cut up the chicken.’ [8.12]

3.2.12.2 Infinitive form + inflected form of *goŋ/đem*

The modal verb *goŋ* ‘to be able’ differs from the above in that it can be combined rather with (infinitival) reduplicated forms, in the same way as the auxiliary *đu(k)* ‘to be’ (section 3.2.3), although it can also be combined with general converbs (example 109).

- (107) *niŋ kolikata ui-gi đeʔna pi-piŋ ar-goŋ-a niŋ*
 1SG Calcutta go-MID.PST QUOT REDPL-COME NEG-can-NEG.NPST 1SG
etai-bar-nen đu-tu.
 WORRY-INF-PL AUX(be)-PRS
 ‘If I go to Calcutta, they worry I will not be able to come back.’ [D.46]
- (108) *zu-zu niŋ goŋ-uraʔ.*
 REDPL-see 1SG can-NEG
 ‘I can’t see.’ [A.G. field notes]
- (109) *boiragi tiŋ-oʔ ar-goŋ-to.*
 hermit shoot-ACT.PST NEG-can-NEG.PST
 ‘The hermit was not able to shoot at [it].’ [10.39]

The construction with (infinitival) reduplicated form is also found for the verb *qem* ‘to become’. See example (102) and the following:

- (110) *usor tubog mīō? sun-sun qe-loŋ?*
 dry earth then REDPL-build become-MID.NPST
 ‘Then will dry earth allow [for a house] to be built?’ [1.100]

3.3 Expressives

Gutob does not show a system of expressives in the sense of a separate part of speech, not derived from another morphological class (Diffloth 2001:263ff.). The rare ‘expressive’ use of the sound [h] has been mentioned in sections 2.3 and 3.1.7. I treat here the echo formations which are a pervasive feature of the language. The following is a good example:

- (111) *laj o-maj-toŋ-nen saka ura?-su, majnen qoŋgor + par*
 who OBJ-3-toŋ-PL aid not-CONJ 3PL hill + mountain
koqok + kuqok qem-o?-su buron-nen qu-tu
 hoe + ECHO do-ACT.PST-CONJ live-PL AUX(be)-PRS
a?so-nen qu-tu.
 live-PL AUX(be)-PRS
 ‘As nobody helped them, they were living by doing hoework⁴² on the hills.’ [5.5]

The example illustrates three predominant patterns: (i) echo proper: repetition of the first word with phonological changes (*koqok + kuqok*); (ii) pairing of two native words with identical or related meaning (*buron + a?so*); and (iii) pairing of two borrowed words with identical or related meaning (*qoŋgor + par*). The following example illustrates (iv) pairing of a native with a borrowed word with identical or related meaning (*luqo + bokrei*):

- (112) *u qi?ke piŋ-gi-su kete lok boqoŋ luqo +*
 there from come-MID.PST-CONJ some people with joke +
bokrei ui-gi-su, anond qeŋ-gu-su-nei, qu-gu-nei.
 banter go-MID.PST-CONJ joy become-MID.PST-CONJ-1PL stay-MID.PST-1PL
 ‘After we had come from there, and had gone around cutting jokes with several people, and had become happy, we stayed [home?].’ [4.3]

4 SYNTAX

4.1 Syntax of the simple sentence

4.1.1 Typological features

Constituent order in Gutob is generally subject-object-verb. Adjectives, numerals, demonstratives and quantifiers precede the head noun, adverbs precede the head they modify. The internal structure of noun phrases can be represented as follows (after Rajan and Rajan 2001b:28):

Noun phrase → (QF)/(DEM) (ADJ₁) (ADJ₂) NOUN

There is evidence from the (frozen) system of nominal composition that constituent order was different in an earlier stage of the development of this language (Note 34 in section 3.1.10.4).

4.1.2 Alignment

Alignment in Gutob is of the ‘nominative–accusative’ type.

- (113) *niŋ o-maj-toŋ-nen ob-gir-oʔ-niŋ-su, majnen gir-gu-nen*
 1SG OBJ-3-toŋ-PL CAUS-learn-ACT.PST-1SG-CONJ 3PL learn-MID.PST-PL
 ‘I taught them and they learned.’ [3.3]

4.2 Complex sentence structure

Mention may be made here of a kind of participle that can be formed by compounding *el* ‘time’ with another word. See examples (19) (*nei pipiŋ + el*), (103) (*noj + el*), and the following:

- (114) ... *utu keroy bulu-gu-nu-el qiʔto ois-oʔ-nen*
 ... that rice ripen-MID.PST-GEN-time QUOT harvest-ACT.PST-PL
qiʔto
 QUOT
 ‘... when that rice had ripened, they harvested.’ [5.289]

4.3 Coordination

Rajan and Rajan (2001b:29) list several examples of asyndetic coordination (see my example 44). In addition, Gutob seems to show several particles of coordinate conjunction, notably *-sa*, *mīōʔ*, and *qoŋ*, but the exact syntactic status of each of them needs further investigation. See also the conjunctive marker *-su* glossed CONJ in this sketch (section 4.7).

- (115) *zona aqo-sa nei-nu oʔ, lioŋ-sa nei-nu*
 corn orchard-and 1PL-GEN EMPH wet.rice.field-and 1PL-GEN
oʔ, lamgo-sa nei-nu oʔ.
 EMPH dry.field-and 1PL-GEN EMPH
 ‘The corn orchard is ours, and the wet field is ours, and the dry field is ours.’
 [B.35]
- (116) *memor mīōʔ niŋ mīōʔ sonia aqʂali boʔ*
 ward.member again 1SG again Sonia Adsali LOC
log-gu-nei beq-oʔ.
 fall-MID.PST-PL AUX(give)-ACT.PST
 ‘The ward member, I and Sonia fell down to Adsali.’ (Rajan and Rajan 2001b:19)
- (117) *maŋ qem-tu-niŋ napaŋla lok dieʔsu tu onooʔn-qoi-lai*
 what do-ACT.NPST-1SG unable person QUOT that daughter-3POSS-OBJ
qoŋ oqʂuoʔn-qoi-lai sun-oʔ.
 and son-3POSS-OBJ say-ACT.PST
 ‘“What shall I do? For I am an incapable person,” she said to her daughter and to her son.’ [5.219]

Disjunction is marked with the word *ki*. See examples (10), (51), and the following:

- (118) *oʔqon kinqʂaʔ ki suloyʔ*
 close river or far
 ‘[Is] the river close or far?’ [A.25]

4.4 Complement clauses

Complement clauses are built with the quotative *dje?* (derived from *dj* ‘to speak’) or its extensions *d(i)e?na* and *dje?su*. Verbs of speaking are often repeated around the complement clause.

- (119) *lokön sun-o? nom-nu agia dje?na ispor + mapru*
 Lakshmana say-ACT.PST 2SG-GEN order QUOT Shiva + lord
niŋ ui-loŋ andari bon bo? kaligai-nu i?taŋ
 1SG go-MID.NPST dark forest LOC black.cow-GEN dung
riŋ-tu-niŋ dje? mapru sun-o?
 bring-ACT.NPST-1SG QUOT lord say-ACT.PST
 ‘Lakshmana then said: “Because it is your order, O Lord Shiva, I will go to the Dark Forest. I will bring the dung of the Black Cow,” said the Lord.’ [6.6]
- (120) *saukar zu-o?-su umbo?-nom ui barogaŋia dje?*
 merchant see-ACT.PST-CONJ whither-2SG go B. QUOT
salag-o?
 ask-ACT.PST
 ‘The merchant saw him and asked, “Where are you going, Barogatia?”’ [10.30]
- (121) *mono? onob dje?na, tuno? moŋ-gu ui-loŋ-niŋ.*
 somewhere girl QUOT thither get.up-MID.PST go-MID.NPST-1SG
 ‘Noticing a girl somewhere, I will get up and go there.’ [C.113]
- (122) *kaligai sun-o?, eno? lok ura?, bak ura? niŋ*
 black.cow say-ACT.PST here people not.be men not.be 1SG
aŋte eran deŋ-gu du-loŋ-niŋ dje?su kaligai
 alone how become-MID.PST be-FUT-1SG QUOT black.cow
sun-o?
 say-ACT.PST
 ‘The black cow said, “There are no people here, no men: how shall I stay alone?” spoke the Black Cow.’ [11.4]

Another complementizer-like quotative (*djito*) is treated under section 5.2. There are rare examples of complement clauses entirely lacking a complementizer: see example (107, before *etaibarnen dudu*) and example (134).

4.5 Relative-type clauses

Gutob shows two types of relative clauses. The first, making use of the GEN marker, seems to be inherited, while the other constructions, whether they use borrowed pronouns or native ones, make the impression of being calques on Indo-Aryan constructions. I refer back to the examples (28), (79), and (99), and add the following (see also sentence xxvii in the sample text under section 7):

- (123) *nei purbe dine di?ke den-ŋem-nu liŋ.*
 1PL olden day from REDPL-DO-GEN wet.field
 ‘It is a wet field which we work on from of old.’ [B.11]

- (124) *mor-ui-gi-nu lok kuḍu, o-maj lion milei*
 NEG-go-MID.PST-GEN person hunger OBJ-3SG wet.field be.available
ura?
 NEG
 'The man who has not gone is hungry; he does not have wet land.'
 [18.25]
- (125) *ura? nei-nu dadi on-o?-nu tunu.*
 no 1PL-GEN uncle plant-ACT.PST-GEN that
 'No, it is the one that our uncle planted.' [B.38]
- (126) *laj mara + mari ḍeṅ-gu-nen o-maj razi ḍem-to.*
 who beat + ECHO become-MID.PST-PL OBJ-3SG agreed make-HAB
 'He makes those who have fought with each other settle their dispute.' [I.4]

4.6 Subordinate clauses

These types of clauses are formed in various ways, making use of the quotative constructions exemplified above (section 4.4), and by the markers *-su* and *-na* that are discussed just below in section 4.7.

4.7 The markers *-su* and *-na*

It has been claimed in a recent article that Gutob appears to have a switch-reference system (Anderson and Boyle 2002:41–42).⁴³

In Gutob ..., the same subject marker *-su* attaches to a past form of the verb, but one lacking person inflection, in line with the generally redundant nature of subject person inflection in switch reference systems. ... The different subject marker in Gutob is *-na*. It also attaches to a past tense form of the verb, similarly lacking person inflection.

There are numerous exceptions to the claims made by Anderson and Boyle. These exceptions do not necessarily invalidate their generalizations entirely, but viewing the markers *-su* and *-na* primarily as opposite poles in one switch-reference system does not seem to me to be the most fruitful approach. In this sketch, I take *-su* as a marker of conjunction ('and'), while I take *-na* to mark predominantly conditional converbs.⁴⁴ The main difference between the two markers seems to me to lie in the association of *-na* with conditional clauses, an association not shared by *-su*. Against ostensible lack of person inflection before *-su*, see examples (92) (*lej-gi-nen-su-sori*), (113) (*ob-gir-o?-niṅ-su*), and further:

- (127) *niṅ piṅ-gi-na pen[-nu] uygom rati konḍek ḍuk-a-naj-su,*
 1SG come-MID.PST-CV 2PL[-GEN] village night a.bit be-IMP-1PL-CONJ
arke dine maskund i-a-naj.
 next day Machkund go-IMP-1PL
 'When I come, let's stay a while in your village at night, and then go to Machkund the next day.' [A.107]

Although the texts show a clear tendency for *-su* with same subject reference (examples *passim* in the present sketch), it is not limited to this. It can connect clauses both with

and without switch of subject. The following examples illustrate different subject reference:

- (128) *nij o-maj-toŋ-nen ob-gir-oʔ-nij-su, majnen*
 1SG OBJ-3-toŋ-PL CAUS-learn-ACT.PST-1SG-CONJ 3PL
gir-gu-nen
 learn-MID.PST-PL
 ‘After I taught them, they learned.’ [3.3]
- (129) *ɖu-gu-nei-su majnen ɖien-boʔ moɖ-gu piŋ-gi-nen.*
 be-PST-1PL-CONJ 3PL house-LOC get.up-MID.PST come-MID.PST-3PL
 ‘After we stayed [a while], they got up and came home.’ [4.4]
- (130) *nij piŋ-gi-nij-su o-nij sa + pa beɖ-oʔ-su*
 1SG come-MID.PST-1SG-CONJ OBJ-1SG tea-ECHO give-ACT.PST-CONJ
ob-lei-oʔ-nen.
 CAUS-sit-ACT.PST-PL
 ‘After I had come, they gave me tea etc. and made me sit.’ [4.6]

Uses of *-na* in temporally subordinate and conditional clauses with same subject reference, on the other hand, are numerous. See example (106) and the following:

- (131) *piŋ-gi zu-oʔ-na babu-nei, boɖoŋ-nei*
 come-MID.PST see-ACT.PST-CV babu-1PL with-1PL
algu + al-gu-su, sobuɖa zu-oʔ-nei,
 COPY + roam-MID.PST-CONJ everything see-ACT.PST-1PL
minɖi aɖ boʔ sobuɖa som-oʔ-nei
 yesterday market LOC everything eat-ACT.PST-1PL
 ‘When we came and saw, we, we went around in yesterday’s market with the Babus, and saw everything, ate everything.’ [4.2]
- (132) *ɖien boʔ lej-gi ɖu-gu-na miŋoʔ*
 house LOC sit.down-MID.PST AUX(be)-MID.PST-CV again
laj zu-tu-nen?
 who see-ACT.NPST-3PL
 ‘If they keep sitting at home, whom else will they see?’ [18.23]

The function of morphemes found in combination with *-su* and *-na*, such as *(-na)-ro*, *(-na)-sina*, and *(-sul-na)-sori*, has yet to be determined.

5 SEMANTICS/DISCOURSE

5.1 Semantics

Two topics that might be selected for treatment under this heading are ‘color terms’ and ‘kinship terms’. I must be very brief about both. Gutob has the cross-linguistically common basic set of three color terms, the commonly used terms being those borrowed from Desia: *dob* ‘white’ (Gutob *pileʔ*), *roŋ* ‘red’ (Gutob equivalent unknown),

kala ‘black’ (Gutob *iqej*). Regarding the topic of kinship terminology, I refer to Zide and Zide (1991), Pfeffer (1999), and Berger (forthcoming).

5.2 Discourse

Besides the quotative markers discussed in section 4.4, there is another particle, of related derivation, namely *qi?to*, presumably in origin meaning ‘it is said’, here also glossed QUOT for lack of precise understanding of its function. It is very common in Gutob narrative.

- (133) *opzir + el, iqa? qi?to iq-o?, lai som-o?,*
 morning + time gruel QUOT drink-ACT.PST cooked.rice eat-ACT.PST
ma? som-o?, kond bo? teŋgia qi?to qon-o?,
 curry eat-ACT.PST shoulder LOC axe QUOT carry-ACT.PST
mĩō? tur-gu ui-gi qi?to tu maj-nu mala bo?
 again leave-MID.PST AUX(GO)-MID.PST QUOT that 3SG-GEN patch LOC
 ‘In the morning, he drank gruel, he ate rice, he ate curry, he carried his axe on his shoulder, and took off again, to that patch of wood of his.’ [5.52]

There is a very common marker of emphasis, the particle *o?* (see Rajan and Rajan 2001b:40); see examples (9–10), (67), (70), and (115). The marker *laka* is used ‘to emphasize measurement’ (Rajan and Rajan 2001b:40); see examples (35–37) and (70). Perhaps related to *laka* is the particle *aka*, also used to emphasize the preceding word. All three markers are here glossed EMPH. For *aka*, see example (85) and the following:

- (134) *niŋ ta ripot qem-tu aka qi-qi?-niŋ qu-tu.*
 I DISC report make-ACT.NPST EMPH REDPL-say-1SG AUX(be)-PRS
 ‘I am saying that I will register a report.’ [B.30]

A striking feature of Gutob discourse is the very frequent occurrence of the particle *be*: see examples (66), (80), and (95). Rajan and Rajan (2001b:3 [section 4.2.1]) interpret it as a politeness marker, and this interpretation is tentatively followed here. Several other particles are borrowed from Indo-Aryan, such as the adversative or topicalizing particle *tolta* (Hindi *to*, Oriya *ta /tə/*; here glossed as DISC) found in examples (52) and (66).

6 LEXICON

As stated above, bilingualism is universal among the tribals of Koraput, and one may doubt whether nowadays Gutob is the first language of any speaker. Code switching from Gutob into Desia and back was an extremely common phenomenon already in the 1960s, as the following example of mixed Desia-Gutob phrases illustrates (the bold elements are Desia):

- (135) *noro-r murti nāĩ, dek-ba-ke sundor nāĩ,*
 man-GEN shape NEG watch-INF-OBJ beautiful NEG
kai-ba-ke mundur nāĩ, qeŋ-gu ui
 eat-INF-OBJ sweet NEG become-MID.PST AUX(go)
qem-to, pen ui-gi qe?na
 AUX(become)-HAB 2PL GO-MID.PST QUOT

qjen-nu paiṭi kimboj + remol pulai-sina zana.
 house-GEN work wife + husband to-only known

‘There’s not the shape of a man, it’s not beautiful to watch, there’s nothing sweet to eat. If you go, only the man and wife know the household work.’ [2.219]

Desia and Gutob morphology can be quite freely intermingled. See example (107) above, where the Desia infinitive marker *-bar* takes the place of the Gutob reduplication to form the present progressive *etai-bar-nen q̄u-tu* ‘they are worrying’ (the interlocutor answers with a more ‘Gutob’ form *eta-nen q̄u-tu*, meaning the same). See also the following example:

(136) *surjo-r saṭi uraṭ-ki, goso-r sitla uraṭ.*
 sun-GEN sunset NEG-or tree-GEN coolness NEG

‘The sun’s setting was not there, nor was the tree’s coolness there.’ [6.3]

Given such a sociolinguistic situation, it may come as no surprise that various fields of the inherited lexicon have been replaced wholesale by, or survive only marginally next to, borrowings from Desia. See the following statement of Zide (1985:97f.) concerning borrowings from Indo-Aryan:⁴⁵

Gutob (like Kharia and Gorum) has borrowed a great deal of vocabulary – including a great percentage of its verb stock – from IA (which, for Gorum and Gutob, means Desia). Of these borrowings, the great majority (but not adjectives or statives) takes the suffixes *-ei* and *-a*. At least thirty percent (of a not particularly conservative dialect) of the verb lexicon consisted of EI/A-taking verbs.

7 BRIEF ANALYZED TEXT: THE GOṬTER RITUAL

The following text was recorded in the 1960s by Richard DeArmond (it is labeled ‘F’ in my database). Thanks are due to Peter Berger for his help in translating the text. For anthropological accounts of the ritual, see Izikowitz (1969), Pfeffer (1991), (2001a), Berger (2007a:284–308). Berger’s forthcoming article listed in the bibliography contains the most detailed account, as well as some remarks on the possible meaning (and linguistic derivation) of the word *goṭter*.

(i) *godba-nen-nu kam-nu samo.*

Gadba-PL-GEN ritual-GEN story
 ‘A story about the ritual of the Gadbas’

(ii) *majnen-nu dadi, ani + puni, bai, ioṅ, goj-gi-na,*
 3PL-GEN grandfather ancestor+ECHO brother mother die-MID.PST-CV

goṭter togri-to-nen
 G. unroll-HAB-PL

‘When their grandfather, ancestor, brother, mother has died, they perform the *goṭter*.’

(iii) *muiroṭ q̄uma-pulai umar-laka boṅtel tol-to-nen, mui-laka*
 one ghost for two-EMPH buffalo tie.up-HAB-PL one-EMPH

buḍa q̄or-to-nen, mui-laka q̄uma-pulai.
 cattle kill-HAB-PL one-EMPH ghost for

‘For one ghost, they tie up two water-buffaloes; they kill one [head of] cattle for one ghost.’

- (iv) *koḍeṭa quma ḍeṅ-gu-na salis munḍ-laka boṅtel*
 twenty ghost become-MID.PST-CV forty head-EMPH buffalo
tol-to-nen.
 tie.up-HAB-PL
 'If there are 20 ghosts, they tie up 40 buffaloes.'
- (v) *sari banda-laka rukug ḍem-to-nen.*
 four banda-EMPH uncooked.rice make-HAB-PL
 'They provide four *banda* of uncooked rice.'
- (vi) *ozar, ḍeḍozar ṭaṅka korso ḍem-to-nen.*
 1,000 1,500 rupee expense make-HAB-PL
 'They make an expense of 1,000, 1,500 rupees.'
- (vii) *boṅtel sob-oṭ-su, munḍa boṭ tol-to-nen.*
 buffalo buy-ACT.PST-CONJ platform LOC tie.up-HAB-PL
 'When they have bought the buffaloes, they tie them at the platform.'
- (viii) *tebe goṭter baza baza-to-nen.*
 then G. rhythm drum-HAB-PL
 'Then they beat the *goṭter* rhythm.'
- (ix) *baza-ṭ-su meṭḍiṅ-to-nen.*
 drum-ACT.PST-CONJ dance-HAB-PL
 'And drumming, they dance.'
- (x) *ḍaṅ + ṭeṅ, ṭaṅgi + ṭeṅgia, kanḍa sob-oṭ-su, goṭter*
 stick-ECHO ECHO-axe sword carry-ACT.PST-CONJ G.
layboṭ boṭ tur-to-nen.
 field LOC leave-HAB-PL
 'They leave for the *goṭter* field, carrying wooden sticks etc., hatchets etc., swords.'
- (xi) *penḍom + ili id-oṭ-su, suloy + suloy-nu lok*
 beer + wine drink-ACT.PST-CONJ far + far-GEN people
piṅ-to-nen goṭter zu-zu.
 come-HAB-PL G. REDPL-see
 'People from very far off drink beer and wine, and come to see the *goṭter*.'
- (xii) *rana, gouḍu, mali, kumar, goren, sobu jati*
 R., cowherd, gardener, potter, G., all community
runḍei-to-nen.
 gather-HAB-PL
 'Ronas, Cowherds, Gardeners, Potters, Gorens (Harijans), all communities meet.'
- (xiii) *besi mara + mari, puza + ana ḍeṅ-gu*
 much fight + ECHO murder + ECHO become-MID.PST
ui-to-nen, bil-gu-su.
 AUX(GO)-HAB-PL get.drunk-MID.PST-CONJ
 'Being drunk, they get into a lot of fighting, murdering.'
- (xiv) *polis-sa piṅ-gi ḍu-to.*
 police-and come-MID.PST AUX(be)-HAB
 'The police are also there.'

- (xv) *polis-nu samo manei-nen ura?, bil-gu-su.*
 police-GEN words respect-PL NEG get.drunk-MID.PST-CONJ
 ‘Being drunk, they do not pay heed to the words of the police.’
- (xvi) *laj-nu samo mana + mani ura?, go?ter lajbo bo? bo?tel*
 who-GEN words respect-ECHO NEG G. field LOC buffalo
tol-o?-su.
 tie.up-ACT.PST-CONJ
 ‘Nobody’s words are heeded, after the buffalo has been tied on the *go?ter* field.’
- (xvii) *go?ter baza baza-?-su me?dij-to-nen, panzia-nen*
 G. rhythm drum-ACT.PST-CONJ dance-HAB-PL P.-PL
ke?u ma?i sol-gu-su.
 lime earth smear-MID.PST-CONJ
 ‘The Panzias smear [one another/themselves] with lime-soil, and dance while drumming the *go?ter* rhythm.’
- (xviii) *go?ter lajbo bo? dui ozar, tin ozar lok*
 G. field LOC two 1,000 three 1,000 people
run?ei-to-nen
 gather-HAB-PL
 ‘Two to three thousand people gather on the *go?ter* field.’
- (xix) *go?ter sarei-gu-na, panzia bo?tel sob-o? ui-to.*
 G. finish-MID.PST-CV P. buffalo carry-ACT.PST AUX(go)-HAB
 ‘When the *go?ter* is finished, the Panzias take away the buffaloes.’
- (xx) *run?ei-gu-nu lok sobu mo?-gu ui-to-nen.*
 gather-MID.PST-GEN people all rise-MID.PST go-HAB-PL
 ‘All the people who had gathered get up and go.’
- (xxi) *kilom lok arko dine potek ?u-to-nen.*
 guest people next day until be-HAB-PL
 ‘The guests remain until the next day.’
- (xxii) *dui din bozi som-o?-su laj-nu ?jen bo? maj*
 two day feast eat-ACT.PST-CONJ who-GEN house LOC 3SG
ui-to-nen.
 go-HAB-PL
 ‘While eating feasts for two days, they go to anybody’s house.’ [?]46
- (xxiii) *sar so ?anka-nu pa?ai sarei-to.*
 four hundred rupee-GEN cloth finish-HAB
 ‘400 rupees worth of cloth is used.’
- (xxiv) *salis pu?i rukug sarei-to.*
 forty pu?i uncooked.rice finish-HAB
 ‘40 *pu?is* of uncooked rice are used.’
- (xxv) *pondro pu?i sa?mel sarei-to.*
 fifteen pu?i ragi finish-HAB
 ‘15 *pu?is* of ragi (finger millet) are used.’

- (xxvi) *koḍe munḍ ki tiris munḍ boṅtel lajbo*
 one.score head or thirty head buffalo field
boʔ seb-oʔ-su zurei-to-nen.
 LOC slice-ACT.PST-CONJ tear-HAB-PL
 ‘They slice open 20 or 30 buffaloes in the field, and tear them open.’
- (xxvii) *suloj zurei-oʔ-su laj parla maj*
 stomach tear-ACT.PST-CONJ who could 3SG
sob-oʔ ui-to-nen.
 carry-ACT.PST AUX(GO)-HAB-PL
 ‘Those who are able to tear [out] the stomach, they take it away.’
- (xxviii) *majnen seli riṅ-gi ḍoṅ-oʔ, pila maizi*
 3PL meat take-MID.PST COOK-ACT.PST child wife
buḍoṅ som-to-nen.
 with eat-HAB-PL
 ‘They take the meat, cook it, and eat it together with children and wife.’
- (xxix) *zar ḍu-gu-nu boṅtel zetki, panzia sobu*
 whose be-MID.PST-GEN buffalo how.many P. all
sob-oʔ ui-to.
 carry-ACT.PST AUX(GO)-HAB
 ‘The Panzias take away all buffaloes, however many [they may be], and to whomsoever they might have belonged.’
- (xxx) *maj muiroʔ goʔter togri-oʔ-nu lok pulai beḍ-to.*
 3SG one G. unroll-ACT.PST-GEN person to give-HAB
 ‘They give it to a man who has performed the *goʔter*.’
- (xxxi) *muiroʔ buḍa ḍor-oʔ-su pans puṭi rukug*
 one cattle kill-ACT.PST-CONJ five puṭi uncooked.rice
ḍoṅ-oʔ-su sobu lok pulai bozi beḍ-to.
 cook-ACT.PST-CONJ all people to feast give-HAB
 ‘He kills one [head of] cattle, cooks five *puṭis* of rice, and offers a feast to all the people.’
- (xxxii) *goʔter baza baza-ʔ-su panzia ui-to*
 G. rhythm drum-ACT.PST-CONJ P. go-HAB
 ‘When the *goʔter* rhythms have been drummed, the Panzias go.’
- (xxxiii) *bozi som-oʔ sobu lok arko dine ui-to-nen.*
 feast eat-ACT.PST all people next day go-HAB-PL
 ‘When they have partaken of the feast all people go, the next day.’
- (xxxiv) *penḍom iḍ-oʔ ili iḍ-oʔ-su*
 beer drink-ACT.PST wine drink-ACT.PST-CONJ
bil-gu ui-to-nen.
 get drunk-MID.PST AUX(GO)-HAB-PL
 ‘Having drunk beer, having drunk wine, they get drunk.’
- (xxxv) *itoʔ ḍem-oʔ gaṭi ṭaṅka + poisa, keroy, irik + saʔmel*
 thus do-ACT.PST much rupee-money, paddy, irik + finger.millet,

sobu korsu dem-o? *suj-suj-nen,* *gadba*
 all expense make-ACT.PST REDPL-AUX(throw)-PL G.
zati-nu *lok.*
 community-GEN people
 ‘In this way, the people of the Gadba community come to make a lot of expenses
 of all kinds: money, paddy, *irik* millet (*panicum miliare*), and finger millet.’

NOTES

- * I am very grateful to the editor of this volume, to Gérard Diffloth, Frits Kortlandt, John Peterson, Tijmen Pronk, and Norman Zide for their reading and critique of part or the whole of a draft of this sketch which was completed in December 2004. This first draft lay untouched by me for more than two whole years. Given the fact that the final draft had to be prepared under pressure of time, I was in the end not able satisfactorily to address many of the important points of criticism of the mentioned readers. Nor was I able to incorporate, as I had first intended, possibly relevant data and analyses from two recently published (brief) sketches of Gutob that came to my attention after completion of my first draft: these are Mukherjee 2002 and Ghosh 2003. I thank Sven Grawunder who sent me a xerox of the latter publication in April 2007, as well as kristen De Joseph and Jasper May who assisted me in preparing this final draft.
- 1 Although the meaning of the name appears to be no longer transparent to the speakers who use it, it can be understood as a compound *gu-tob* ‘creature of the earth’, with *gV-* (see section 3.1.10.2) prefixed to *tob*, presumably a combining form (see 3.1.10.4) of *tubog* ‘earth’ (see *tom* in *guf-tom* ‘to wash the face’ :: *tumog* ‘face’). The combining form *tob* also appears to be found in *to?qur* ‘white ant’, whose Remo cognate (Bhattacharya 1968 #1497) is *tobqur*, and in Kharia *tobqa?* ‘mud’ = ‘earth-water’. Hence, the meaning of *gu-tob* in this analysis exactly parallels the meaning of Desia *ma?ia* ‘earth-people’, an important self-defining concept of Gutob-Gadba society discussed by Peter Berger (2007a:100, 105f.). On these grounds, I strongly prefer the given etymology above the derivation from the word for ‘egg’ (*utob*), as proposed to me by Norman Zide (personal communication, with reference to origin myths among speakers of North Munda languages). For a derivation from a word for ‘stream, river’, see Ramadas (1931).
 - 2 The frequently quoted association of the name Gad(a)ba with the Godāvarī river (e.g. Parkin 1991:31, Berger 2007a:183–189) seems likely to be another such case. Considering the etymological explanation of the name Gutob given in Note 1 above, and the likelihood of folk-etymological speculations on a name no longer understood, the historical value of indigenous traditions on the Gutob-Gadbas’ having come ‘originally’ from the Godāvarī river is not to be overestimated.
 - 3 Telugu is the predominant language among the Gutob-Gadabas who have descended into the plains of Andhra Pradesh. Sarma (2001:9) remarks about the use of Gutob in this habitat: ‘neither the older generation speak it among themselves, nor do they make use of it with the younger generation, making the language progressively defunct.’ See section 1.3.
 - 4 In a personal communication (27 August 2002), Peter Berger confirmed this estimate:

The Ollar are definitely much more numerous than the Gutob. The latter inhabit only a small stretch from Deptahanjar to Ongel, Onmail, Kangrapada etc. To the South are the Parenga, to the North the Ollar. I have not yet tried to make an estimate from my village surveys, but if we say there are 50 Gutob villages (probably there are less) in the Lamptaput/ Onukadilli area with an average of 200–300 people you get your numbers.

- 5 It is not entirely clear how one is to bring this estimate in line with the following statement, published by the same authors earlier in the same year (2001a:9 – quotation exactly as printed):

According to the 1981 census of India, the total Gadaba population is 56,913 in undivided Koraput district (Malkangiri, Rayagada, Koraput and Nowrangpur) Orissa. It is inclusive of Ollar Gadabas and Gotop Gadabas. In 1995 a Socio – linguistic survey done by Asha Kiran Society puts the number of speakers of Gutop – Gadaba in undivided Koraput Orissa, at 15,000 – 20,000. Majority of Gutop Gadabas live in Lamtaput block of Koraput District.

Without indication of any source, the most recent (15th) edition of the *Ethnologue* www.ethnologue.com/show_language.asp?code=gbj (accessed 4 May 2007) mentions 8,000 speakers, for the year 2000.

- 6 Rajan and Rajan (2001b:5), speak of four dialects within their ‘Gutob-Gadaba’, that is, our ‘Koraput Gutob’: Kinda Raji, Lamtaput dialect, Koraput dialect and Birong Raji. I do not know the meaning of the word *raji* (presumably not Desia *raji* ‘agreement’), nor whether the division into *kinda?* ‘river’ and *biroy?* ‘jungle’ dialects corresponds to an indigenous classification. The data for the Rajan and Rajan ‘write-up’ was collected from the Kinda Raji dialect spoken by the people of Tikkorpada Panchayat, Lamtaput Block, Koraput District, Orissa.’ See also Zide 1965:49 n13.
- 7 See the comment by Zide (1978:50):

Note the (previously recognized) confusion between Gorum (Parengi) and Gutob in the LSI. The numerals given in the lexical tables at the back of the Munda Section of Vol. IV include under Gadba (Bastar) what are Gadba (Gutob) forms. In parentheses following these are Gorum (Parengi) forms which are called Gadba (Vizagapatam).

EDITOR’S NOTE: The text sample returned under Gadba is likewise actually in Gorum.

- 8 The publisher’s address is Asha Kiran Society, Lamtaput – 764 081, Koraput, Orissa, India. Along with these publications aimed at a linguistic audience, a ‘Gutob-Gadaba Language Learner’s Guide’ has also been published, with basic sentences in Gutob (in Oriya script and romanization), in Oriya, and in English. Further language learning materials published later by the same society were brought to my attention by Felix Rau in 2005, but I have not yet found time to go through them attentively.
- 9 An electronic (searchable) form of the database is accessible online at hin.osaka-gaidai.ac.jp/griffithsa/ (accessed on 7 November 2007). Other texts, including those recorded by myself, are to be added in due course.
- 10 Their ostensible evidence for contrast is: /ʎay/ ‘rice’, /ʎay/ ‘for’, /sʎay/ ‘street’, /nayʔ/ ‘we (INC)’, and /layʔ/ ‘who’. I transcribe these words: /lai/, /lai/, /sai/, /naj/, and /laj/, respectively – there is one homonym in my analysis, viz. /lai/ ‘cooked rice’ or ‘for, OBJ’.
- 11 Previous work has tended to call unreleased plosives in final position ‘checked’ (Bhattacharya 1965–1966) or ‘preglottalized’. For example, Zide 1965:52 speaks of ‘the – phonetically preglottalised in final position – weakly voiced stops /b, d, j, g/’. The same author (1972:512) claims: ‘Gutob *guj-tom*- ‘to wash the mouth’, but *gui-mod*- ‘to wash the eyes’. The preglottalised *j* of the former is kept since no other postvocalic glottalised consonant follows it in the stem, whereas the glottalisation is lost in *gui-mod*- since the *j*-preceded preglottalized (postvocalic) *-d* in the same verbstem.’ I have no data to confirm the phenomenon described by Zide, and transcribe the first element of both words *guj*-.

- 12 In the analysis of Das Gupta and Bhattacharya (1975:11) and Mahapatra (1985), /h/ is entirely absent, but Gustafsson (1989) does record a very small number of words with initial /h/.
- 13 This is not a hard and fast rule: the nasal consonant seems to be preserved in disyllables, for example, [duruŋ-oʔ] ‘he sank (TR)’, [riŋ-riŋ-oʔ-nen] ‘they are bringing/have brought’ (see example 65).
- 14 The status of seeming exceptions *abaj-gu* ‘played’ (5.354) and *amroj-gu* ‘smeared’ (C.28) needs to be checked.
- 15 The suffix *-gi* tends to be pronounced as [dʒi] in the case of this verb.
- 16 If in such a case the transcribed texts also happened to miss the presumably intended/underlying glottal stop, and the suffix hence appears to vanish entirely, ostensible problems of grammatical analysis may emerge, such as the one hinted at by Hook 1991:192 n 10.
- 17 In fact the available data yield no compelling examples of *pulai* combined with the GEN and just one example of *-nu lai*.
- 18 Cf. this author’s broad definition of ‘gender’ as a ‘classification of nouns or other parts of speech on the basis of inflections or other grammatical features’ (ibid.).
- 19 The latter, Bhattacharya claims, is found only with reference to humans, but this claim is false: see *puḍralpuḍri* cited below.
- 20 Especially p. 26 on the obligatory usage of *-ta* after interrogative and other pronouns ‘in case of non-attributive use’.
- 21 The pairing of *laj* with *tu* could also be interpreted as a relative/correlative pairing, but the context seems to suggest otherwise.
- 22 As to the ‘expressive’ category, see Zide (1991:363): ‘A final word – on *h* in Munda demonstratives The recorded Gutob lexicon shows just one morpheme with *h*, *ha-* ‘yonder’. All forms with *ha* take a peculiar intonation, and the *a* is often expressively lengthened; thus, *haaanoʔ*, “waaay out there”. Note that Rajan and Rajan (2001b:16) posit a four-way system, but do not record the form with *h-*.
- 23 Anderson (p.c.) speculates that the second part of *monoʔ* (and *umboʔ*) might derive from *-n(u)-boʔ* (GEN-LOC).
- 24 Forms in parentheses are retranscribed following Zide 1978:51.
- 25 The mere existence of the forms recorded by Ramamurti shows that a radical change has occurred either on the Koraput or on the Andhra side. Regarding the latter, Subba Rao (1992:17) states: ‘For counting numbers from six to eleven, five is taken as the base. Mode of counting above five is based on additions to five. . . . Numbers higher than nine are made up both by multiplication and addition.’
- 26 Cf. also *o-pen oʔqon du-tu kiʔ* OBJ-2PL near be-ACT.NPST Q ‘Is it near you?’ [A.73], where *oʔqon* governs an objective case form. See also example (118).
- 27 As noted in section 3.1.2 (with Note 17) above, the status of this marker as a postposition seems uncertain, because clear cases of its combination with a GEN suffix are not available in the database.
- 28 See Peterson’s discussion of the cognate Kharia formation in this volume for a possible phonological explanation for the use of this infix.
- 29 The vowel /o/ in the last example (*sumoi*) remains unexplained. In addition, the words *munan* ‘manner’ and *tunom* ‘boiling’ may derive from verbs. The word *ninoŋ* ‘yoke’ does not show the expected vowel harmony, nor does a root *noŋ* ‘to yoke’ seem to be attested.
- 30 I do not know a root *tel* from which this verb could be derived. May we assume a connection with *boŋtel* ‘buffalo’, the animal used for threshing grains underfoot on the *tinel*, or with *soʔtil* ‘to glean grains fallen on the ground’?
- 31 I know one example of a possible *-Vm-* infix (*gumul* cited in section 3.1.10.1), so *simin* could perhaps be derived from *sī* (from **sinʔ*) ‘sun’.
- 32 The first syllable does not show the expected vowel, so this derivation is doubtful.

- 33 EDITOR'S NOTE: This is probably a reduplicated form, not a *sV- prefix, matching the reduplicated form for 'hand' (*titi*).
- 34 It is interesting to note that the order of elements in these compounds is the reverse of what one might expect in what is (currently) a 'modifier-head' language.
- 35 Zide (1985:101) states: 'the negative uses the same morphemes but redistributed arbitrarily, that is, I can find no logic to the "redistributed system".'
- 36 The tense-aspectual interpretation in these and several of the following examples of negated verb forms is uncertain.
- 37 See Mahapatra 1985:68 on the Desia 'adverbial durative' non-finite verbal suffix *-te*, here glossed as CV.
- 38 Stems derived by CAUS-derivation (section 3.2.10.2) are polysyllabic by definition, and therefore do not reduplicate; see example (70).
- 39 See Zide's statement (1985:94) concerning Kharia, but implicitly valid for South Munda languages more generally: 'Roughly ... monosyllabic stems take B^{px} as (probably) do bimorphemic stems and borrowed stems. The remainder – disyllabic monomorphemic stems – take (the great majority of them) B^{ix}' (where B^{px} and B^{ix} stand for the causative prefix and infix, respectively).
- 40 If my analysis of example (14) is correct, we have evidence that the compound can be interrupted by the CONJ marker (section 4.7).
- 41 The words *zar gorke take* in this example are Desia.
- 42 The interpretation of the echo formation *koḍok + kuḍuk* is not entirely certain, but it seems most likely that it is connected with the Desia words *koḍokbar* 'to work a field with a hoe' (see Gustafsson 1989:135) and *koḍki* 'hoe'.
- 43 EDITOR'S NOTE: In the cited article, the authors state there are two formally overlapping but functionally distinct systems that *-na* participates in, one typically in conversation (in texts and actual conversation) where *-na* functions as a subordinator marking conditional and temporally subordinate clauses and with possible same subject reference across the clauses, and one as a device of narrative discourse, where the patterning is one of a different subject marker. They do erroneously state that the *-su* element does not allow person marking which is clearly false.
- 44 It is to some extent parallel in function (and form!) to the 'conditional converb' in *-ile* of Oriya (Neukom and Patnaik 2003:251), based on the PST suffix *-il-*.
- 45 Zide's (1991) paper looks at possible grammatical influence from Dravidian. No evidence for direct lexical influence from Dravidian languages on Koraput Gutob is known to me, although there is considerable evidence for such influence on Gutob via Desia.
- 46 One wonders whether the text is correctly transcribed here. If *laj* can mean 'own', the CONJ marker *-su* can receive its usual translation 'after ...'.

REFERENCES

- Note: A complete survey of colonial (nineteenth-century) sources, some containing short word-lists, is available in Pinnow (1959:464ff.). These outdated sources are not repeated here.
- Anderson, Gregory D.S. (2001) 'A new classification of Munda: Evidence from comparative verb morphology', *Indian Linguistics*, 62:27–42.
- Anderson, Gregory D.S. and John P. Boyle (2002) 'Switch-reference in South Munda', in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 39–54.
- Anderson, Gregory D. S. and Norman Herbert Zide (2002) 'Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint', in M. A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian*

- Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Berger, Peter (2000) 'Gesellschaft, Ritual und Ideologie. Eine relationale Betrachtung der Gadaba des Koraput Distriktes, Orissa', *Mitteilungen der Berliner Gesellschaft für Anthropologie, Ethnologie und Urgeschichte*, 21:15–28.
- (2001–2002) 'Feeding the dead: Rituals of transformation among the Gadaba of Koraput', *ADIVASI*, 40–41:35–50.
- (2002) 'The Gadaba and the "non-ST" Desia of Koraput, Orissa', in G. Pfeffer and D.K. Behera (eds) *Contemporary Society: Tribal Studies*. Vol. V: The Concept of Tribal Society, New Delhi: Concept Publishing Company, 57–90.
- (2007a) *Füttern, Speisen und Verschlingen. Ritual und Gesellschaft im Hochland von Orissa, Indien*, Berlin: LIT.
- (2007b) 'Sacrificial food, the person and the Gadaba ritual system', in G. Pfeffer (ed.) *Periphery and Centre. Studies in Orissan History, Religion and Anthropology*, New Delhi: Manohar, 199–221.
- (forthcoming) 'Who are you, brother and sister? The theme of "own" and "other" in the go'ter ritual of the Gadaba', in P. Berger, R. Hardenberg, E. Kattner and M. Prager (eds) *An Anthropology of Values. Essays in Honor of Georg Pfeffer*, New Delhi: Longman.
- Bhaskara Rao, Peri (1969) 'Gadaba phonology', *Linguistic Survey Bulletin*, 14:5–8.
- (1980) *Koṅḍjekor Gadaba. A Dravidian Language*, Pune: Deccan College.
- (1998) 'Gadaba', in S.B. Steever (ed.) *The Dravidian Languages*, London and New York: Routledge, 328–355.
- Bhattacharya, Sudhibhushan (1956) *Ollari. A Dravidian Speech*, Delhi: Govt. of India, Department of Anthropology.
- (1965–1966) 'Glottal stop and checked consonants in Bonda', *Indo-Iranian Journal*, 9:69–71.
- (1968) *A Bonda Dictionary*, Poona: Deccan College.
- (1975) *Studies in Comparative Munda Linguistics*, Simla: Indian Institute for Advanced Study.
- (1976) 'Gender in the Munda languages', in P. Jenner, C.L. Thompson, and S. Starosta (eds) *Austroasiatic Studies Part 1*, Honolulu: University Press of Hawaii, 189–211.
- Burrow, T. and S. Bhattacharya (1962–1963) 'Gadba supplement', *Indo-Iranian Journal*, 6:45–51.
- Dasgupta, Dipankar and Shyamsundar Bhattacharya (1975) 'The Desiya of Koraput', *Indian Journal of Linguistics*, 2:1–24.
- DeArmond, Richard C. (196?) A collection of Gutob texts typed and mimeographed under the heading 'DeArmond's Texts', Unpublished.
- (1976) 'Proto-Gutob-Remo-Gtaq stressed monosyllabic vowels and initial consonants', in P.N. Jenner, C.L. Thompson, and S. Starosta (eds) *Austroasiatic Studies Part 1*, Honolulu: University Press of Hawaii, 213–227.
- (no date) [no title: paper on the verb in Gutob], Mimeographed.
- Diffloth, Gérard (2001) 'Les expressifs de Surin, et où cela conduit', *Bulletin de l'École française d'Extrême-Orient*, 88:261–269.
- Fürer-Haimendorf, Christoph von (1943) 'Megalithic ritual among the Gadabas and Bondos of Orissa', *Journal of the Royal Asiatic Society of Bengal*, 9:149–178.
- Gaura, Duḥkhiśyāma (1991) *Gādabā (vyākaraṇa, racanāvalī, śabdakoṣa)*, Bhubaneswar: Academy of Tribal Dialects and Culture.
- Ghosh, Arun (2003) 'Linguistic sketch of Gutob', in A. Ghosh *An Ethnolinguistic Profile of Eastern India. A case of South Orissa*, Burdwan: University of Burdwan, 107–131.
- Gustafsson, Uwe (1973a) 'Clause patterns in Kotia Oriya', in R.L. Trail (ed.) *Patterns in Clause, Sentence and Discourse in selected languages of India and Nepal II*, Norman, OK: SIL / Kathmandu: Tribhuvan University Press, 191–254.

- Gustafsson, Uwe (1973b) 'Kotia Oriya texts', in Ronald L. Trail (ed.) *Patterns in Clause, Sentence and Discourse in selected languages of India and Nepal III*, Norman: SIL / Kathmandu: Tribhuvan University Press, 111–212.
- (1989) *An Adivasi Oriya-Telugu-English Dictionary*, Mysore: CIL.
- Hook, Peter Edwin (1991) 'The compound verb in Munda: An areal and typological overview', *Language Sciences*, 13:181–195.
- Izickowitz, Karl Gustav (1951) Letter to F.B.J. Kuiper from Koraput, Orissa. Copy kept at the Kern Institute, Leiden.
- (1969) 'The Gotr ceremony of the Boro Gadaba', in S. Diamond (ed.) *Primitive Views of the World*, New York: Columbia University Press, 129–150.
- Konow, Sten (1906) *Linguistic Survey of India, Vol. IV. Munda and Dravidian Languages*. Reprint 1967, Delhi: Motilal Banarsidass.
- Kornel, Das (1999) *Tribal Cultural Heritage and Cult. The Gutob Gadaba of Orissa*, Bhubaneswar: Modern Book Depot.
- Mahānti, Gopinātha (1956) *Gādabā Bhāṣā*, Kāṭaka: Akhiḷa Bhārata Sarbasebā Saṅgha (Grāmasebaka Samabāya Press).
- Mahapatra, Khageswar (1985) *Desia. A Tribal Oriya Dialect of Koraput Orissa*. [= *ADIBASI* 25 (Nos 1–4)], Bhubaneswar: Tribal and Harijan Research-cum-Training Institute.
- Manoharan, S. (2001) 'Indian languages in censuses – A Study', *Indian Linguistics*, 62:117–140.
- Mitra, A. (1964) *census of India 1961. Volume 1. India Part II-C (ii) - Language Tables*, Delhi: Manager of Publications.
- Mukherjee, Sibasis (2002) 'Gadaba (Gutob)', in B. Mahapatra (ed.) *Linguistic Survey of India. Special Studies. Orissa*, Kolkata: Language Division, Office of the Registrar General, 299–314.
- Neukom, Lukas and Manideepa Patnaik (2003) *A grammar of Oriya*, ASAS [Arbeiten des Seminars für Allgemeine Sprachwissenschaft] Nr. 17, Universität Zürich.
- Parkin, Robert (1991) *A Guide to Austroasiatic Speakers and Their Languages*, Honolulu: University of Hawaii Press.
- Patnaik, Radha Krishna (1992) 'The Gadabas', in V. Subba Rao and R. Patnaik (eds) *Gadaba. The Language and the People*, Amaravathi: Sri Papayaradhya Sahiti Kendram, 2, iv, 60.
- Pfeffer, Georg (1991) 'Der intra-agnatische "Seelentausch" der Gadaba beim großen Lineageritual', in M. Laubscher (ed.) *Beiträge zur Ethnologie Mittel- und Südindiens*, München: Anacon, 59–92.
- (1997) 'The scheduled tribes of middle India as a unit: Problems of internal and external comparison', in G. Pfeffer and D.K. Behera (eds) *Contemporary Society: Tribal Studies. Professor Satya Narayana Ratha Felicitation Volumes*, Vol. I: Structure and Process, New Delhi: Concept Publishing Company, 3–27.
- (1999) 'Gadaba and the Bondo kinship vocabularies versus marriage, descent and production', in D.K. Behera and G. Pfeffer (eds) *Contemporary Society Tribal Studies. Vol. IV: Social Realities*, New Delhi: Concept Publishing Company, 17–46.
- (2001a) 'A ritual of revival among the Gadaba of Koraput', in B. Schnepel and H. Kulke (eds) *Jagannath Revisited: Studying Society, Religion and the State in Orissa*, Delhi: Manohar, 123–148.
- (2001b) 'Sprache und Religion bei den Stämmen Mittelindiens', in Dirk W. Lönne (ed.) *Tohfa-e-Dil. Festschrift Helmut Nespital. Band II: Kulturwissenschaften*, Reinbek: Verlag für Orientalistische Fachpublikationen, 769–782.
- Pinnow, Heinz-Jürgen (1959) *Versuch einer historischen Lautlehre der Kharia-Sprache*. Wiesbaden: Harrassowitz.
- Rajan, F. Herold and Jamuna Rajan (2001a) *Gutob-Gadaba Phonemic Summary*, Asha Kiran Society: Lamtaput, Orissa.

- (2001b) *Grammar Write-up of Gutob-Gadaba*, Lamtaput, Orissa: Asha Kiran Society.
- Ramadas, G. (1931) 'The Gadabas', *Man in India*, 11:160–173.
- Ramamurti, G.V. (1938) *Sora-English Dictionary*, Madras: Government Press.
- Sarma, C.R.S. (2001) 'Gutob Gadba: A sociolinguistic appraisal', Unpublished MS distributed at the National Seminar on Tribal Languages in India, Hyderabad: Osmania University, February 16–17.
- Subba Rao, Vavilala (1992) 'A grammatical analysis of a tribal language, Gadaba, as spoken in the village "Rajacheruvy Valasa", a Hamlet of Bobbili in Vijayanagaram district of Andhra Pradesh, [on the main title page: 'Language Analysis'], in V. Subba Rao and R. Patnaik *Gadaba. The Language and the People*, Amaravathi: Sri Papayaradhya Sahiti Kendram, ix, 101.
- Thusu, Kidar Nath and Makhan Jha (1972) *Ollar Gadba of Koraput*, Calcutta: Anthropological Survey of India.
- Zide, Norman (1965) 'Gutob-Remo vocalism and glottalized vowels in Proto-Munda', in G.B. Milner and J.A.E. Henderson (eds) *Indo-Pacific Linguistic Studies* (= *Lingua* 14), 43–53.
- (1968) 'Some comparative notes on Gata? pronouns', in J.C. Heesterman, G.H. Schokker, V.I. Subramoniam (eds) *Pratidānam: Indian, Iranian and Indo-European Studies Presented to F.B.J. Kuiper on his Sixtieth Birthday*, The Hague: Mouton, 348–358.
- (1972) 'Review of S. Bhattacharya, a Bonda dictionary', *Journal of the American Oriental Society*, 92:506–513.
- (1978) *Studies in the Munda Numerals*, Mysore: CIIL.
- (1985) 'Notes mostly historical on some participant roles in some Munda languages', in A. Zide, D. Magier, and E. Schiller (eds) *Proceedings of the Conference on Participant Roles: South-Asia and Adjacent Areas*, Bloomington: Indiana University Linguistics Club, 92–103.
- (1991) 'Possible Dravidian sources of some Munda demonstrative bases', in B. Lakshmi Bai and B. Ramakrishna Reddy (eds) *Studies in Dravidian and General Linguistics: A festschrift for Bh. Krishnamurti*, Hyderabad: Osmania University, 349–364.
- (1997) 'Gutob pronominal clitics and related phenomena elsewhere in Gutob-Remo-Gta?', in A. Abbi (ed.) *Languages of Tribal and Indigenous Peoples of India: The Ethnic Space*, Delhi: Motilal Banarsidass, 307–334.
- (1999) 'A note on a rhetorical-stylistic device in Gutob (Gadaba)', in K. Mahapatra (ed.) *Banaja-1999: Adibasi Exhibition Souvenir*, Bhubaneswar, 44.
- Zide, N. and G.D.S. Anderson (2001) 'The Proto-Munda verb and some connections with Mon-Khmer', in P. Bhaskara Rao and K.V. Subbarao (eds) *Yearbook of South Asian Languages and Linguistics*, Delhi: Sage Press, 519–540.
- Zide, Norman and B.P. Das (1963) *Gutob Verb Lexicon*, Unpublished.
- (1965) *Gutob Texts*, Unpublished Mimeo.
- Zide, N. and A.R.K. Zide (1991) 'A linguistic analysis of some South Munda kinship terms, I', in J. Davidson (ed.) *Austroasiatic Languages: Essays in Honour of H.L. Shorto*, London: SOAS, 43–59.

GTA?

Gregory D.S. Anderson

1 INTRODUCTION

The Gta? language is spoken by the Gta? people of southern Orissa, primarily in Malkangiri district and adjacent areas of Koraput as well. The people and language are known as Didayi [*diɖa:y[i]*] in India both in English and indigenous language vernacular publications. Their self-designation is generally *gta?* (with some groups self-identifying as Didayi, now as well), as in *gta?=re* ‘Gta? person’, with the first part (*gta?*) probably cognate with Gutob/Gadab[a] and the latter part (=re, the combining form of *remwa*) cognate with Remo (cf. *remo-sam* ‘Remo language’).

The final classification of Gta? remains an open question. The Hill Gta? variety is still too poorly attested to accurately assess even the prehistory of Gta? itself much less how it fits in with the rest of the cognate Munda languages. It has been classified as ‘Lower Munda’ by Bhattacharya (1975), the Gutob–Remo–Gta? subgroup of Koraput Munda (Zide 1969) and its own subgroup of South Munda (Anderson 2001a). Remo, in particular, seems to have had some not insignificant influence on Gta? over the past centuries (and Proto-Gutob–Remo interacted and shared developments with Proto-Gta?) rather than it necessarily forming a distinct subgroup with these two clearly related sister languages (Gutob and Remo).

Plains Gta? and Hill Gta? are the recognized varieties. They differ to a fair degree from what data is available. There are under 4,500 speakers of Gta? including both its varieties total.

Gta? remains an unwritten language. The confusing convention for using the grapheme {x} for [æ]/[ɛ] found in (Mahapatra *et al.* 1989) is not followed here.

From an ethnographic perspective, the Gta? are said to have the only polyphonic singing style attested among the people of India.

2 PHONOLOGY

To say that the phonology is unusual would be to put it mildly. In a sense, Gta? is similar to Georgian or Khoisan languages (or some Salish and other NW American languages) where clusters (or minor syllables) of an unusual sort clog syllable onset position. However, its manifestation is not at all as it is in these other languages. On top of this oddity, there is both a distinctly Munda and a distinctly South Asian ‘feel’ to the sound of the language. The use of glottal stops and pre-glottalized stops are hallmarks of Munda consonantism, while the retroflexion common in the language comes as no surprise given the language’s location at the juncture between Indo-Aryan and Dravidian.

- (3)
- | | | | | | |
|----------|------------|-----------|----------|----------|----------|
| <i>p</i> | <i>t</i> | <i>t̥</i> | <i>c</i> | <i>k</i> | <i>ʔ</i> |
| <i>b</i> | <i>[d]</i> | <i>d̥</i> | <i>j</i> | <i>g</i> | |
| <i>m</i> | <i>n</i> | | | <i>ŋ</i> | |
| | <i>s</i> | | | | <i>h</i> |
| | <i>l</i> | | | | |
| | <i>r</i> | <i>ɽ</i> | | | |

Note that *-g* is pre-glottalized in coda position in Gtaʔ. There is some phonotactic evidence in Gtaʔ that there may be a phonological set of prenasalized stops. A similar argument has been made about other Munda languages, for example, Santali (Anderson 2001b). This would add the following set to the those in (3) above: *mb nd [ŋd̥] [ɽj]* and *ŋg*.

2.4 Syllable structure and phonotactics

Gtaʔ is an unusual language from the perspective of syllable structure or phonotactics. It has an enormous number of ‘clusters’ found in word-initial position but a restricted number of consonants found in coda position. A small number of words with syllabic nasals and prenasalized stops may also be found.

- (4) Clusters attested in word-initial position

<i>bb</i>	<i>br</i>	<i>bn</i>	<i>bl</i>		<i>bs</i>	<i>bk</i>	<i>bt</i>	—										
<i>cc</i>		<i>cn</i>	<i>cl</i>						<i>cm</i>									
<i>d̥d̥</i>		<i>d̥n</i>								<i>d̥b</i>	—							
<i>gg</i>	<i>gr</i>	<i>gn</i>	<i>gl</i>	<i>gɽ</i>	<i>gs</i>		<i>gt</i>	<i>gm</i>	<i>gb</i>	<i>gd̥</i>	—	<i>gh</i>						
<i>hh</i>	<i>hr</i>	<i>hn</i>	<i>hl</i>	<i>hɽ</i>														
<i>jj</i>		<i>jn</i>																
<i>kk</i>	<i>kr</i>	<i>kn</i>	<i>kl</i>	<i>kɽ</i>	<i>ks</i>	—		<i>km</i>										
<i>ll</i>		<i>ln</i>	—		<i>ls</i>			<i>lm</i>	<i>lb</i>			<i>lg</i>	<i>lh</i>					
<i>m-m</i>	<i>m-r</i>	<i>m-n</i>	<i>m-l</i>					—	<i>m[-]b</i>									
<i>n-n</i>		—	<i>n-l</i>		<i>n-s</i>		<i>n-t</i>			<i>n[-]d̥</i>	<i>[n-g]</i>							
							<i>ŋ[-]k</i>				<i>ŋ[-]g</i>							
<i>pp</i>	<i>pr</i>	<i>pn</i>	<i>pl</i>	<i>pɽ</i>														
<i>rr</i>	—	<i>rn</i>			<i>rs</i>	<i>rk</i>			<i>rb</i>			<i>rg</i>						
<i>ss</i>	<i>sr</i>	<i>sn</i>	<i>sl</i>		—	<i>sk</i>		<i>sm</i>		<i>sb</i>	<i>sd̥ sg</i>							
<i>tt</i>	<i>tr</i>	<i>tn</i>	<i>tl</i>					—	<i>tm</i>	<i>tb</i>		<i>[tg]</i>	<i>th</i>					

In addition to these a small number of words may be found with the following clusters as well: *bt̥, lŋ, lj, gj, gc, n-c* and *m-p*.

- (5) *ŋku gnag-hwaʔ toʔ-ce ga-ge*
 tiger door-rope open-SS enter-PST
 ‘The tiger opened the door and entered.’ (Mahapatra and Zide n.d. 53)

This feature of Gtaʔ appears to be very similar to syllable structure constraints found in other Austroasiatic languages that are distant relations of Gtaʔ. However, rather than representing a retention of an archaic phonotactic feature directly

inherited from Proto-Austroasiatic lost in all other Munda languages, this feature of Gta? is more likely to be a pseudo-archaism, that is, a feature that looks to be archaic, but is, in fact, a secondary development by chance (?) mimicking an archaic feature. This feature consists of sesquisyllabic word-structure with an associated array of word-initial consonant ‘clusters’, see (6). Although, so-called sesquisyllabic words are found frequently in Mon-Khmer languages, their presence in Gta? arose through, among other processes, the loss of unstressed vowels in word-initial syllables. The specific form of the vowel found in such languages as Gutob or Remo is not predictable based on the Gta? form, but the reverse is largely true. Some words exhibiting this unusual CCV[V][C] structure that characterizes the language (in addition to the autonym Gta?) include

- | | |
|---------------------------------------|--|
| (6) <i>bba</i> ‘to pat’ | <i>bsa</i> ‘to grow long (of hair)’ |
| <i>bbo?</i> ‘snake’ | <i>bto?</i> ‘fire to dazzle’ |
| <i>bluy</i> ‘sister’ (said by female) | <i>bto?</i> ‘fear’ |
| <i>bkig</i> ‘pungent’ | <i>cca</i> ‘to lick’ |
| <i>bno?</i> ‘ladder of single bamboo’ | <i>clæ</i> ‘(be) long or tall’ |
| <i>bri</i> ‘uncultivated land’ | <i>cmu</i> ‘seed’ |
| <i>cnu</i> ‘promise, oath’ | <i>hli?</i> ‘bamboo shoot’ |
| <i>cri</i> ‘(become) deep’ | <i>hma?</i> ‘curry, vegetable’ |
| <i>drig</i> ‘listen’ | <i>hni</i> ‘village’ |
| <i>dqu</i> ‘stick’ | <i>hrwe?</i> ‘cow’ |
| <i>dno?</i> ‘trap’ | <i>jjo[?]</i> ‘put in’ |
| <i>ggæ</i> ‘chew’ | <i>jnir</i> ‘spindle’ |
| <i>gbug</i> ‘pig’ | <i>kmiy</i> ‘younger brother’s wife’ |
| <i>gce?</i> ‘cockroach’ | <i>kso?</i> ‘winnowing fan’ |
| <i>gra?</i> ‘mouse’ | <i>klir</i> ‘be blind’ |
| <i>gdi?</i> ‘frog’ | <i>kka?</i> ‘donkey’ |
| <i>gdo?</i> ‘female genitals’ | <i>knwe?</i> ‘wife’ |
| <i>ghæ?</i> ‘rope’ | <i>knqæŋ</i> ‘drum’ |
| <i>gia?</i> ‘be broken into pieces’ | <i>kri</i> ‘be fat’ |
| <i>glæ?</i> ‘ox’ | <i>kreso</i> ‘be hungry’ ‘hunger’ |
| <i>gmi?</i> ‘goat’ | <i>lli</i> ‘creeper’ |
| <i>gsu?</i> ‘dog’ | <i>lboŋ</i> ‘wet’ |
| <i>gni</i> ‘tooth’ | <i>lmir</i> ‘tasty’ |
| <i>gtaŋ</i> ‘dog-fly, tick’ | <i>lye?ria?</i> ‘be oily’ ‘to melt of oil’ |
| <i>hhia</i> ‘festival, worship’ | <i>lgo?</i> ‘neck’ |
| <i>lja</i> ‘shallow’ | <i>nsa?</i> ‘waist’ |
| <i>lse</i> ‘pus’ | <i>nso?</i> ‘banana’ |
| <i>mpæ</i> ‘child!’ | <i>ntoŋ</i> ‘gourd’ |
| <i>mmu</i> ‘nose’ | <i>nti</i> ‘hand’ |
| <i>mria?</i> ‘wake up’ ‘arise’ | <i>ntwig</i> ‘mushroom’ |
| <i>mni</i> ‘name’ | <i>pna?har</i> ‘big size broom’ |
| <i>mbar</i> ‘2’ | <i>ppa?</i> ‘to brush, flick’ |
| <i>nnaŋ</i> ‘to sound’ | <i>pqar</i> ‘to burst’ |
| <i>ncia</i> ‘bone’ | <i>pru-gug</i> ‘bare, uncovered, naked’ |
| <i>ncoŋ-ti</i> ‘right hand’ | <i>pni?</i> ‘flute’ |
| <i>mbæsia?</i> ‘left hand’ | <i>plig</i> ‘bird’ |
| <i>ndia?</i> ‘water’ | <i>præg</i> ‘after a moment’ |

<i>ndɾæʔ</i> ‘root’	<i>rbog</i> ‘to rot’
<i>ŋgaʔ</i> ‘rat, mongoose species’	<i>rgæiŋ</i> ‘rub body against wall (like animal)’
<i>nliɑ</i> ‘tongue’	<i>rkoʔ</i> ‘uncooked rice’
<i>ŋgog</i> ‘insect’	<i>rna</i> ‘loft, attic under roof’
<i>nhug</i> ‘ear’	<i>rsiŋ</i> ‘assemble, gather’
<i>nji</i> ‘3’	<i>rrua-kig</i> ‘winter’
<i>njiɾ</i> ‘morning’	<i>snaŋ</i> ‘house’
<i>ssig</i> ‘cold’	<i>srog</i> ‘sip, slurp’
<i>sbo</i> ‘taste’	<i>tɪ</i> ‘hand’
<i>sgir</i> ‘be warm’	<i>tre</i> ‘kendu tree’
<i>skoʔ</i> ‘gourd-ladle’	<i>tmwaʔ</i> ‘mouth’
<i>slaʔ</i> ‘tree’	<i>tmu</i> ‘tethering pot’
<i>smo</i> ‘front to front’	<i>t[u]gwa</i> ‘yesterday’
<i>sni</i> ‘sun’	<i>tlæ</i> ‘to stack’
<i>sne</i> ‘place of work’	<i>thwā</i> ‘stand’
<i>sniʔ</i> ‘horn mallet used to beat bark-fiber’	<i>tboʔ</i> ‘earth, ground’

There is an interestingly restricted set of consonants allowed in coda position word-finally. This includes *ʔ*, *ŋ*, *k*, *g*, *r*, *n*. One also finds *t* but only in loans and *t* as well, but only in demonstrative stems. In word-medial syllables, coda position also allows *c*, *l* and *m*.

Velar nasals are highly marked word-initially in Gtaʔ, almost only (with one or two lexical exceptions) found as a syllable nucleus or as a part of a prenasalized stop complex. Retroflex nasals only occur as allophones of /n/ before [ɟ].

- (7) Remo Gtaʔ
ŋkuy *ŋkui*
‘water pot’ ‘water pot’ (Bhattacharya 1975: 47)

2.5 Intonation/stress

Stress occurs in final position in Gtaʔ words, with some exceptions. For example, some morphemic combinations attract stress (the negative in combination with a second singular subject in various conjugations), while others disprefer it (the tense/aspect suffixes). Longer words have secondary stress patterns, but these also remain to be investigated.

- (8) *bʔo-k[e]-ne* *dapɾe* *heʔ-baŋ* *næŋ* *ljo* *ceʔmwa* *na-big-ɟe*
fear-T/A-NF immediately today-ABL I field grass 2-sow-or
n[ʌ]-á-big *ɟak-ce* *basoŋ-ke*
2-NEG-SOW QUOT say-T/A
‘Fearing, he said “from today will you or will you not sow my field?”’
(Mahapatra and Zide no date, I.26)

2.6 Morphophonology

There are a range of changes that individual morphemes undergo when used in combination with others. Vowel changes may be evidenced when comparing the forms

for the numerals in the second decade in Gta?. A *u* in the second syllable triggers *u*-vocalism in the element meaning ‘ten’ (*gu-*), while with a different following vowel, we find *go-*. Note that the *o*-vocalism is original in the stem.

- (9) *gonji* ‘13’ *gugu* ‘17’
gomal ‘15’ *gu?tur* ‘16’
gombar ‘12’ *gua* ‘10’
gommwij ‘11’ *gu* ‘7’
gosotij ‘19’ *tur* ‘6’

Some morphemes (when appearing in certain phonological environments or morphemic combinations) seem to accrue a nasal or appear in phonological variants, one of which has a prenasalized element or a floating nasal and one (or more) variant(s) that lack(s) it.

- (10) (a) *andji we-pe-ŋ-dij-ke dakce basoŋ-ke*
 where go-2PL-DO-T/A QUOT say-T/A
 ‘“Where are you going”, he says’ (Mahapatra and Zide no date, C.11)
- (b) *mæ-ne b-n-asa? mæ? jontu-dig we-la abaŋ*
 He-GEN residence/NOUN/residence what animal-ADD go-DS survival
nɖu mæ tmwa?-ha? ga-uwe-dij-ke
 NEG.COP he mouth-EMPH enter-REDPL:AUX:TLOC-AUX-T/A
 ‘Whatever animal went by his residence does not survive, but rather ends up in his mouth.’ (Mahapatra and Zide no date, H.2)
- (c) *b’ba-ræ basoŋ-ke nɖu andi-andi-ne si-si=re-hij at*
 father-3 say-T/A no COPY:where-GEN plough=man-PL there
gali wig-dij-e dakce-ka bari basoŋ-la-ndig a-mane-ke
 way go.home-AUX-FUT QUOT-only again say-DS-ADD NEG-agree-T/A
 ‘Her father disagreed, “no, ploughmen might be passing through on their way home.”’ (Mahapatra and Zide no date, I.11)
- (d) *bari me?-swa-ne diɖi-le hanɖa-ndæ pag=li? we-ke*
 then one-day=DEF how-INTERJ husband-3 break=shoot go-T/A
 ‘Then one day what happened but the husband went for shoots.’ (Mahapatra and Zide no date, G.7)
- (e) *ɖukri hanɖa-ndæ-ke hāwe pno? bæ-ke*
 Old.woman husband-3-CASE sharpen spear send-T/A
huŋ-dæ-ke hā beŋɖa=ha? bæ-ke
 child-3-CASE pluck okra=vegetable send-T/A
 ‘The old woman sent her husband to sharpen a spear and sent her child to pluck some okra.’ (Mahapatra and Zide no date, L.18)
- (f) *wig-ce bba-ræ-ke basoŋ-ke “nɖu aba næ? ljo*
 go/come-SS father-3-CASE say-T/A well father:VOC we:INC land
ja-ndæ-ha? kia bbig-dij-ke dakce bba-ræ-ke
 who-3-EMPH paddy REDPL:SOW-AUX-T/A QUOT father-3-CASE
basoŋ-ke”
 say-T/A
 ‘She came home and told her father “well daddy, is someone (supposed to be) sowing our field?”’ (Mahapatra and Zide no date, I.10)

3 MORPHOLOGY

3.1 Nominal morphology

The nominal morphology of Gta? constitutes a developed and relatively straightforward system from an inflectional perspective. Categories of number, case, and person are found; marking of definiteness may be realized in the morphological system of the language as well. The derivational system of the language, on the other hand, presents a very complicated picture, which is discussed only in brief in section 3.1.10.

Generally speaking, the issue of parts of speech is a complex one in Munda (see Evans and Osada 2005 for a recent discussion). Under the heading of nominal morphology, we recognize the following partially functionally determined (partially morphologically and/or syntactically determined) subcategories of nominals for Gta?: nouns, pronouns (including interrogatives, negative and indefinite), demonstratives, numerals (including classifiers), adpositions, adjectives or adjectivals and adverb[ial]s. It should be borne in mind that working out the exact nature of these subcategories, and what if anything can be used as diagnostic for determining individuated parts of speech in Gta?, remains an object of future research.

In terms of the formal or templatic structure of a noun in Gta?, a maximal one is hard to determine, and there are several different sub-templates. For personal pronouns it is

- (11) OBJ-PRONOUN-CASE
a-næ-ke k-mæ-hij a-mia?
 OBJ-WE-CASE DEM-3PR-PL NEG-KNOW
 ‘They do not know us.’ (Mahapatra and Zide no date, A.36)

For third person pronouns, the maximal template is

- (12) OBJ-3PR-DL/PL-CASE
a-mæ-pa-ke go+gsia? ηdu
 OBJ-3PR-DL-CASE child+monkey NEG.COP
 ‘They had no child[ren].’ (Mahapatra and Zide no date, M.2)

For demonstrative pronouns, the maximal template is

- (13) OBJ-DEM-3PR-PL-CASE
a-k-mæ-hij-ke mmwiy co{a-k[e]-ne ne-[m]bi? ɖakce bari basoŋ-ke
 OBJ-DEM-3PR-CASE one lame-T/A-DEP 1PL-give QUOT again say-T/A
 ‘‘Let’s give them that lame one’’, he said.’ (Mahapatra and Zide no date, A.49)

For animate nouns the template is

- (14) Noun-3-PL-CASE
huy-ɖæ-hij-ke
 child-3-PL-CASE
 ‘(to) his children’

while inanimate nouns, unless anthropomorphized in texts, tend to be uninflected. This template is expandable in the case of the Noun by expanding what functions in the templatic spot of the stem via juxtaposition compounding, derivational compounding, or simple derivation. So, for example, ‘Stone God’ is one word in Gta? and is an example of juxtaposition compounding, simple concatenation of two free forms of nouns with no coordinative particle, affix or clitic, for example,

- (15) *kitoŋ-bwar basoŋ-ke hāprig=ta iaŋ hāprig=ta bari*
 God=stone say-T/A shut=mouth mother:voc shut=mouth then
ʃātu=ta iaŋ ʃātu=ta
 open=mouth mother:voc open=mouth
 ‘The stone-god said, “shut your mouth, mother, shut your mouth” and then
 “open your mouth, mother, open your mouth.”’ (Mahapatra and Zide no
 date, A.21)

3.1.1 Number

There are two formally marked number categories for nouns in Gta?. In addition to the singular which has no formal marker, there is the dual marker *-pa* and the plural marker *hiŋ* (also realized as *hī*). The dual marker, historically, is the second person dual marker. The dual marker is commonly found on the rightmost element in a juxtaposed compound (with two full free forms of nouns), as in the following two examples:

- (16) (a) *griŋ-ŋku-pa ho-ba?lir-ke*
 Cat-tiger-DL RECIP-talk-T/A
 ‘The cat and tiger talked to each other.’ (Mahapatra and Zide no
 date, H.5)
 (b) *ʒokra-ʒokri-pa sapa jiwa-ke bagwe?-har-ke*
 Old.man-old.woman-DL all animal-CASE kill-PL-T/A
 ‘The old man and old woman killed all the animals.’ (Mahapatra and
 Zide no date, D.24)

It may also attach to the third person pronominal stem to encode ‘they two’.

- (17) *mæ-pa mba-ya hara-hari ʒiŋ-ce kōḍa pwetur-æŋ-har-ke*
 he-DL two-person defeat:ECHO AUX-SS mountain surround-shit-PL-T/A
 ‘The two of them trying to beat each other, went around the mountain
 shitting.’ (Mahapatra and Zide no date, II.12)

As the last two examples show, the *-pa* marker for dual is limited to nominal or pronominal stems, and is not repeated for a third dual subject inflection for the verb. Instead the number of the subject, if encoded at all formally (see the first example above for an instance of a sentence that lacks an overt subject marker in its verb), is encoded by a non-singular (or plural) verbal subject marker *-har-*. In other words, dual as a number category is only morphosyntactically activated for nouns, while verbs make a simple singular/non-singular distinction.

As aforementioned, the plural marker comes before the case suffix but after the person suffix in the maximal nominal template of Gta?. Thus we find examples with such structures as Noun-PL (18) Noun-3-PL (19) 1/2-Noun-PL (20) DEM-3PR-PL (21)

- (18) (a) *k-mæ-hij ke-la olæŋ-læʔ-k[e]-ne remwa-hij a-næ-ke*
 DEM-3PR-PL see-DS wander-AUX-T/A-DEP person-PL OBJ-WE-CASE
k-mæ-hij a-miaʔ
 DEM-3PR-PL NEG-know
 ‘They, (you see), are nomads, and they do not know us.’ (Mahapatra and Zide no date, A.36)
- (b) *nɖu næʔ des lag-neʔ-ɖo ɖæk ɖakce-ka bari*
 No WE.INC country away-1PL-run like.this QUOT-only again
kia-hij ho-baʔlir-ke
 paddy-PL RECIP-talk-T/A
 ‘“No let us run away from this country” the paddies conversed.’ (Mahapatra and Zide no date, C.7)
- (19) (a) *huŋ-ɖæ-hij basoŋ-ke iaŋ næ bari na ɖætte*
 Child-3-PL say-T/A mother:VOC we also/again you with
ne-paŋ-e
 1PL-come-FUT
 ‘The children said, “mother we also will come with you.”’ (Mahapatra and Zide no date, B.15)
- (b) *knweʔ-ræ-hij ɖig basoŋ-ke we-la-pe bir=tia*
 Wife-3-PL also say-T/A go-IMP-2PL gayil(wild.cattle)
tar-læʔ-ke
 come.out-AUX-T/A
 ‘The[ir] wives also said, “go, the wild cattle has come out.”’ (Mahapatra and Zide no date, A.3)
- (20) *næ=dre-hij*
 1PL-person-PL
 ‘our people’ (Mahapatra and Zide no date, F.11)
- (21) *k-mæ-hij oʔ-kæŋ ɖoŋɖ+kost ɖiŋ-ce*
 DEM-3PR-PL this.much punishment+trouble do-ss
bar=si+jir=si ɖɖiŋ-ɖiŋ-ke
 two=day+three=day REDPL:do-AUX-T/A
 ‘These ones were working with so much trouble the last two or three days.’ (Mahapatra and Zide no date, A.32)

Complex templates are found as well. This includes internally complex nouns that may be derived compounds, consisting of, for example, a combination of free form + combining form of nouns (22), or one of non-finite verb + combining form (23), or consisting of juxtaposed compounds as well, for example, Noun-Noun-PL (24)

- (22) *heʔ-dig tæŋ kitoy big-ne ceʔmwa ɖak-ce kitoy-ke*
 today-even that god sow-PRTCPL/NF/GEN grass COMP god-CASE
gtaʔ=re-hij gge-miaʔ-ke
 Gtaʔ-people-PL REDPL:worship-CUST-T/A
 ‘Even today the Gtaʔ people worship that god as he might [have] sow[n] grass.’ (Mahapatra and Zide no date, I.32)

- (23) *b'ba-ɾæ basoŋ-ke nɖu anɖi-anɖi-ne si-si=re-hij at*
 father-3 say-T/A no COPY:where-GEN plough=man-PL there
gali wig-dij-e ɖakce-ka bari basoŋ-la-ŋɖig a-mane-ke
 way go.home-AUX-FUT QUOT-only again say-DS-also NEG-agree-T/A
 'Her father disagreed, "no, ploughmen might be passing through on their way home."' (Mahapatra and Zide no date, I.11)
- (24) (a) *eɽ-ɾe ggo=re-hij koŋɖa-baŋ wig-dij-la*
 now-time REDPL:hunt=person-PL mountain-abl come.home-AUX-DS
sela-mbweɽ-hij kero kisaɽlo gag-la ggo=re-hij
 female-PL way loin.cloth tie-DS REDPL:hunt=person-PL
hni gga a-ia-har-ke
 village REDPL:enter NEG-CAP-PL-T/A
 'Now [if] the hunters return home, and the women have blocked their access by tying loincloths together, the hunters can't enter the village.' (Mahapatra and Zide no date, A.63)
- (b) *meɽ-swa-ne kia-manda-hij rsij-ke bari ho-baɽlir-ke*
 One-day-DEF/GEN paddy-group-PL gather-T/A again RECIP-talk-T/A
 'One day all the paddies assembled and spoke with each other.' (Mahapatra and Zide no date, C.4)

As in many Eurasian SOV languages, a noun appears in the singular following a numeral (in some instances with an intervening classifier but not in this example; for more see section 3.1.8).

- (25) *nɖia kuma-ce nji harke dij læɽ-la knweɽ-ɾæ gweɽ-we-ge*
 Water bath-ss three month do COP-DS wife-3 die-go/AUX-T/A
 '(after) she gave birth, three months passed and the wife died.' (Mahapatra and Zide no date, B.4)

Also, inanimate nouns often lack number marking altogether, even when clearly referentially plural.

- (26) *sela-mbweɽ-hij kero kisaɽlo gag-la ggo=re-hij*
 female-PL way loin.cloth[s] tie-DS REDPL:hunt=person-PL
hni gga a-ia-har-ke
 village REDPL:enter NEG-CAP-PL-T/A
 '(if) the women have blocked their access by tying loincloths together, the hunters can't enter the village.' (Mahapatra and Zide no date, A.63)

3.1.2 Case

The case system of Gta? is relatively complex. There are both grammatical or structural cases and postpositional or semi-cases. There is also a range of postpositional elements in Gta? discussed in section 3.1.9 which may be entering into the case system. Possession and object marking are particularly complex in terms of their realization and variation and are addressed separately below.

The unmarked form of the noun functions as a subject (ii/iii/v) or in as of yet undetermined contexts, verbal objects (direct/indirect) as well, even if animate and referential (i), and commonly so if inanimate in reference (ii/iii/iv).

- (27) (a) *coʔa=bir ne-biʔ-la k-mæ-hij a-næ-ke mæʔ mane-e be*
 Lame=gayil 1PL-give-DS DEM-3PR-PL OBJ-we-CASE obey-FUT Q
 ‘We will give them the lame gayil (but) will they obey us?’ (Mahapatra and Zide no date, A.35)
- (b) *mæ-hij int-baŋ janwai hni hātar-ce tarpa-ʔæŋ*
 3PR-PL that.side-ABL Janwai village abandon-ss ceremonial.drum
linʔi=kwi ale-tæn-ne sorte saʔ-ce at
 ceremonial.cooking.pot all-that-DEF/GEN entirely hold-ss there
jar-ce haʔaʔ-koŋʔa basa-har-ke
 climb.down-ss vicinity-mountain reside-PL-T/A
 ‘From that side they abandoned the village of Janwai, took up all their ceremonial drums and pots, climbed down and settled in the vicinity of the mountain.’ (Mahapatra and Zide no date, F.6)
- (c) *dæk eʔ-baŋ næŋ ljo ceʔmwa na-big-la a-na-ke bol*
 look today-ABL I field grass 2-NEG-SOW OBJ-you-CASE good
n-a-ʔij ʔæt ʔakce-ka bari kitoŋ-ke basoŋ-ke
 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god, “look, you sow grass in my field (again), then I will not be nice to you.”’ (Mahapatra and Zide no date, I.28)

Another potential function of an unmarked noun in a Gtaʔ sentence is to mark a location or goal of an action (there is no locative case form in the language).

- (28) *na ʔwa læ[ʔ]-ceʔ cili+haʔʔo na-we-raŋ-e ja tæŋ-sa*
 You home remain-ss meat+fish 2-go-bring-FUT who that-PURP
sle we-ge
 work go-FUT
 ‘(she said) “[if] you stay at home and [go] bring meat and fish; who will go to work for those things?”’ (Mahapatra and Zide no date, M.16)

Possibly the most straightforward of the case forms of Gtaʔ is the ablative in *-baŋ* which belongs to the class of postpositional or semi-case elements. It carries both a spatial motion away from meaning (i) as well as a more metaphorical extension of the ablative semantics to mean ‘from’ with a time referent (ii). It shows no real variation in terms of its realization, consistently appearing as *-baŋ*.

- (29) (a) *eʔ-ʔe ggo=re-hij koŋʔa-baŋ wig-ʔij-la*
 now-time REDPL:hunt=person-PL mountain-ABL come.home-AUX-DS
 ‘Now [if/when] the hunters return home from the mountains.’
 (Mahapatra and Zide no date, A.63)
- (b) *kitoŋ basoŋ-ke “na ljo eʔ-baŋ næŋ n-a-big*
 god say-T/A you field today-ABL I 1-NEG-SOW
ʔakce basoŋ-ke”
 QUOT say-T/A
 ‘The god says, “from today on I will not sow your field any longer.”’
 (Mahapatra and Zide no date, I.29)

In this latter function, the ablative is particularly common with interrogative and demonstrative stems.

- (30) (a) *mmwiy goʔæ=o basoŋ-ke oʔna-bay aʔsi-ge*
 One boy say-T/A when-ABL fever-T/A
 ‘One boy asked, “since when has he been sick?”’ (Mahapatra and Zide no date, M.12)
- (b) *mæ-hij int-bay janwai hni hātar-ce*
 3PR-PL that.side-ABL Janwai village abandon-SS
 ‘From that side they abandoned the village of Janwai.’ (Mahapatra and Zide no date, F.6)
- (c) *at-bay gbe gmiʔ-ke bʔo-læʔ-ke*
 there/then-ABL bear goat-CASE fear-AUX-T/A
 ‘since then bear has been afraid of goat.’ (Mahapatra and Zide no date, II.15)

Another functionally straightforward case form in Gta? is the purposive or designative case, usually rendered ‘for’ or ‘in order to’ in English translation. It marks the purpose for which an action is performed. It has several realizations, including *-sa*, *-nsa*, *-saʔ*, *-nsaʔ*, and *-sar*.

- (31) (a) *næŋ slweʔ-nsa konta ũ-we-ʔe cili+haʔʔo-nsa*
 I belly-PURP mountain 1-AUX/go-Q:RHET meat+fish-PURP
ũ-we-e
 1-AUX/go-FUT
 ‘Am I to go to the mountain work in order to fill our bellies or in order to eat fish and meat?’ (Mahapatra and Zide no date, M.17)
- (b) *eʔ mæʔ-saʔ næʔ neʔ-we-e ɖakce ja-ʔig*
 today what-PURP we.INC 1PL.INC-go-FUT QUOT who-ADD
a-we-ke
 NEG-go-T/A
 ‘“For what will we go today?” (they said) and no one came.’ (Mahapatra and Zide no date, D.29)
- (c) *mæʔ-sar n-a-big ɖakce mæ basoŋ-ke*
 what-PURP I-NEG-SOW QUOT he say-T/A
 ‘“Why shouldn’t I sow?” He said.’ (Mahapatra and Zide no date, I.20)
- (d) *gbe basoŋ-ke arei clæ+nlug kaʔi+ɖiraŋ na mæʔ-nsaʔ*
 bear say-T/A ho long+ear stick+horn you what-PURP
kala-haʔ raʔ=sæ raʔ=sæ ppaŋ na-miaʔ-ke
 daily go.ahead=obstruct COPY REDPL:COME 2-CUST-T/A
 ‘Bear said, “whoa there long-ear stick-horn why do come first every day and block (me)?”’ (Mahapatra and Zide no date, II.3)

The last of the common postpositional cases is the adessive or ad-lative in *-[n]nia/-rnia*. It means ‘be near’, ‘(move) up to near’ in contexts requiring these locational/directional semantics. It also is one of the variant means of encoding possession in Gta?; see below.

- (32) (a) *mæ ŋku-nnia we-ce sarlo+uhwě-ce ko-ke*
 He tiger-ADESS go=SS greet+ECHO/TAG-SS sit-T/A
 ‘He went up to the tiger, greeted him and sat down.’ (Mahapatra and Zide no date, H.4)

- (b) *ɖukri* *p-n-aʔ=cweʔ=ɽiaʔ* *ɖokra-ke* *pno-mæɡ-nnia*
 Old.woman millet-gruel oldman-CASE anus-near
aʔ-ug-ce *aʔ-ro+aʔ-beʔ-ce* *osmar* *hnor*
 CAUS-drink-SS CAUS-compose+ECHO-SS song CONCOMITANT
aʔ-ro-ce *hoʔ-ke*
 CAUS-compose-SS cry-T/A
 ‘The old woman smearing millet gruel near the oldman’s anus composing songs, cried while singing.’ (Mahapatra and Zide no date, D.14)

The most straightforward of the grammatical or structural cases is the genitive or possessive case in *-ne*. This same element also functions as a definite marker (sometimes with vaguely assertive semantics), as an adnominal attributive formant, as well as one of the component elements (along with the tense/aspect or participle element) in the means to form relative clause structures in the language as well (see section 4.2.1).

The case suffix *-ne* appears on the first of two nouns that stand in a possessor–possessed relationship to one another, appearing in that order. Both pronouns (including demonstratives) and nouns may be marked by the genitive/possessive suffix in *-ne* in Gtaʔ (33). Thus, the possessor is marked by *-ne* and is followed by the possessum, which if the latter belongs to the class of inalienably possessed nouns (see section 3.1.3) may appear in the possessed form as well (with third person possessors) (34).

- (33) (a) *te-la* *tæn* *kitonj* *basonj-ke* *næŋ-ne* *paiŋi* *kæn-haʔ*, *na*
 that-DS that god say-T/A I-GEN work this-EMPH you
ɖi-rokom *kia* *na-big-ke* *næŋ* *ɖiŋ* *se-rokom*
 what-manner paddy 2-SOW-T/A I also that-manner
ceʔmwa *bbig-ɖij-ke*
 grass REDPL:SOW-AUX-T/A
 ‘Then that god said, “my task is this: whatever manner you sow in, I will sow in that manner too.”’ (Mahapatra and Zide no date, I.17)
- (b) *te-la* *a-tæn* *hni-ne* *pujari* *at go=re*
 This-DS that.very village-GEN priest that hunt=person
bitre mmwij-ja *pujari* *læʔ-ge* *tæn* *pujari* *mæʔ pag=ra*
 among one-person priest remain-T/A that priest what offering
biʔ-ke *tæn* *konɖa-ne* *bri-huŋ-ke* *pag=ra biʔ-ke*
 give-T/A that mountain-GEN mountain-child-CASE offering give-T/A
 ‘Then the priest of that village who was also among the hunters; that priest made some offering to the mountain-god of that mountain.’ (Mahapatra and Zide no date, A.52)
- (c) *na* *sapa* *konɖa-ne* *jantu-ke* *wa-raʔ-sij-ce*
 you all mountain-GEN animal-CASE call-CAUS-gather-SS
ɖwa *aʔ-gga*
 home CAUS-REDPL:enter
 ‘You call and gather all the animals of the mountain and get them to enter the house.’ (Mahapatra and Zide no date, D.10)

- (34) *mæ-ne knweʔ-ɾæ nɖiaʔ kuma-ge*
 3PR-GEN wife-3 water bath-T/A
 ‘His wife gave birth.’ (Mahapatra and Zide no date, B.2)

Conjoined possessor nouns appear without a conjunction as is typical of Gta? syntax, with both nouns marked for genitive.

- (35) *ɖetwaʔ laʔɾig laʔɾig laʔɾig-ce mal-bhaʔ aʔ-ɾij-ce*
 like.that night night night-SS five-head:CLSSFR CAUS-do-SS
tur-bhaʔ sgwa bir=tia-ne sre-ne ɖuŋ+tæ-ce ɖwa
 six-head:CLSSFR like gayil-GEN/DEF deer-GEN/DEF carry-SS home
raŋ-wig-har-ke
 bring-come.home-PL-T/A
 ‘Spending night after night like that they brought home five or six head of gayil and deer.’ (Mahapatra and Zide no date, A.59)

Complex or recursive/embedded genitive phrases are also found in Gta?, each noun appearing in a genitive-marked form. Note that these do not formally differ from the asyndetic or conjunctionless coordinative structure just mentioned, distinguishable only by context.

- (36) *mōjæ hni we-har-la mæ-hij basoŋ-ke næ-ne hni-ne*
 Middle village GO-PL-SUB 3PR-PL say-T/A WE-GEN village-GEN
slaʔ kaʔɾu pe paŋ-ce-ka bari sapa
 tree branch? you.PL come-SS-only again entirely
gwe+hæʔ-biʔ-pe-ŋɖij-ke
 cut+ECHO-AUX-2PL-AUX-T/A
 ‘Because they settled in the middle of the village, those others said “you are coming here and cutting down all the trees of our village and their branches.”’ (Mahapatra and Zide no date, F.14)

Note that one of the variant means of encoding ‘have’ constructions in Gta? uses the genitive as well, explained later.

Another grammatical case in Gta? is *-ke*. The functions of this case are quite varied, but many cluster around a dative or a dative-accusative familiar from a range of languages. However, its functions are numerous and for this reason it is simply glossed as *-CASE*, as it is probably the default non-subject grammatical case. The system described below is relevant only to noun stems, and not demonstratives or pronouns, where other subsystems are attested (see discussion of *a-* below). It is likely but not certain, that *-ke* is a loan (from Desia?) in Gta?, as its functions partly overlap with other elements or forms in the system, and as just mentioned, pronouns also exhibit their own phenomena.

In its basic functions, the *-ke* case suffix can mark a recipient or ‘indirect object’ (i) or a patient or ‘direct object’ (i/ii). It, thus, functions as a kind of ‘primary object’ marker in the Dryer (1986) sense.

- (37) *wig-ce* ***bba-ɾæ-ke*** *basoŋ-ke* “*nɖu aba* *næ?* *ljo*
 go/come-SS father-3-CASE say-T/A well father:VOC we:INC land
ja-ŋdɛ-ha? *kia* *bbig-dij-ke* *dakce* ***bba-ɾæ-ke***
 who-3-EMPH paddy REDPL:SOW-AUX-T/A QUOT father-3-CASE
basoŋ-ke”
 say-T/A
 ‘She came home and told her father “well daddy, is someone (supposed to be) sowing our field?”’ (Mahapatra and Zide no date, I.10)
- (38) (a) *hni-[n]-[d]re* *basoŋ-ke* *næ ne-we-e* *tæn* ***bir=tia-ke***
 Village-GEN-person say-T/A we 1PL-go-FUT that gayil-CASE
ne-twiŋ-e
 1PL-shoot-FUT
 ‘The villagers said, “we will go, we will shoot down those wild cattle.”’ (Mahapatra and Zide no date, A.4)
- (b) *ljo* *habo?-bo-la* ***huŋ-dæ-ke*** *bæ-ke*
 land forget-AUX-DS child-3-CASE send-T/A
 ‘It was forgotten at the field so he sent his child (for it).’ (Mahapatra and Zide no date, I.5)

For other verb stems, it is not clear how to analyze the *-ke*-marked noun either syntactically or semantically (there is sufficient variation), but, nevertheless, it seems to mark the primary object of the clause.

- (39) (a) ***kitoy-ke*** *b-a?-to-ke* *dapɾe* *kitoy* *bto?-ke*
 god-CASE frighten\CAUS\T/A immediately god fear-T/A
 ‘He frightened the god and the god became fearful.’ (Mahapatra and Zide no date, I.25)
- (b) *he?-dig* *tæn* *kitoy* *big-ne* *ce?mwa*
 today-even that god SOW-PRTCPL/NF/GEN grass
dakce ***kitoy-ke*** *gta?=re-hij* *gge-mia?-ke*
 COMP god-CASE Gta?-people-PL REDPL:worship-CUST-T/A
 ‘Even today the Gta? people worship that god as he might [have] sow[n] grass.’ (Mahapatra and Zide no date, I.32)

Another, related function in Gta? of the case suffix *-ke* is to mark the causee in a causative construction.

- (40) (a) *tæn* ***goɾæ=o-ke*** *a?-ga-ce* *mæ* *kia*
 that boy=child-CASE CAUS-enter-SS he paddy
tto-dij-ge
 REDPL:pound-AUX-T/A
 ‘Then (she) made that boy come in and start pounding paddy.’ (Mahapatra and Zide no date, L.33)
- (b) *na* *sapa* *koŋda-ne* ***jantu-ke*** *wa-ra?-siŋ-ce*
 you all mountain-GEN animal-CASE call-CAUS-gather-SS

ɖwa aʔ-gga
 home CAUS-REDPL:enter
 ‘You call and gather all the animals of the mountain and get them to enter the house.’ (Mahapatra and Zide no date, D.10)

A superlative object, perhaps better conceived of as ‘primarily affected’ object in Gta? may be marked by *-ke* as well.

- (41) (a) *tæn kiton remwa-ke bweʔ-tur-la remwa-ke gæ-ke*
 that god person-CASE spit-DS man-CASE itch-T/A
 ‘The god spit on the man, and the man started to feel itchy.’ (Mahapatra and Zide no date, I.22)
- (b) *remwa ɖi-le kiton-ke bweʔ-tur-la kiton-ke poka-poka*
 man oh.boy! god-CASE spit-DS god-CASE blisters
tar-ke
 emerge-T/A
 ‘The man spat on the god and blisters broke out on the god.’ (Mahapatra and Zide no date, I.23)

The case element *-ke* may optionally (though most commonly does) appear within adverbs of various types in Gta? (cf. English ‘to-day’).

- (42) (a) *eʔ mæʔ-saʔ næʔ neʔ-we-e ɖakce ja-ɾig a-we-ke*
 today what-PURP we.INC 1PL.INC-go-FUT QUOT who-ADD NEG-go-T/A
 ‘“For what will we go today?” (they said) and no one came.’ (Mahapatra and Zide no date, D.29)
- (b) *eʔ-ke næʔ-nnia swa nɖu na wig-la-ce*
 Today we-ADESS fire NEG.COP you go.home-IMP-SS
swa toʔ-raŋ-la ɖak-ce griŋ-ke toʔ=so bæ-ke
 fire pull.out-bring-IMP state-SS cat-CASE pull.out=fire send-T/A
 ‘“Today we have no fire, you go home and bring the fire,” saying thus, the bear sent the cat to bring fire.’ (Mahapatra and Zide no date, H.13)

As in many South Asian languages, a noun functioning as an experiencer subject may be case-marked by *-ke* in Gta?.

- (43) (a) *tæn kiton remwa-ke bweʔ-tur-la remwa-ke gæ-ke*
 that god person-CASE spit-DS man-CASE itch-T/A
 ‘The god spit on the man, and the man started to feel itchy.’ (Mahapatra and Zide no date, I.22)
- (b) *remwa ɖi-le kiton-ke bweʔ-tur-la kiton-ke poka-poka*
 man INTERJ god-CASE spit-DS god-CASE blister:COPY
tar-ke
 emerge-T/A
 ‘The man spat on the god and blisters broke out on the god.’ (Mahapatra and Zide no date, I.23)

Finally a kind of possessive notion may be marked by a noun in *-ke* in Gta? in certain configurations as well.

- (44) *ɖukri* *p-n-aʔ=cweʔ=ɽiaʔ* ***ɖokra-ke*** *pnomæɡ-nnia*
 Old.woman millet-gruel oldman-CASE anus-ADESS
aʔ-ug-ce *aʔ-ro+aʔ-beʔ-ce* *osmar* *hnor*
 CAUS-drink-SS CAUS-compose+ECHO-SS song CONCOMITANT
aʔ-ro-ce *hoʔ-ke*
 CAUS-compose-SS cry-T/A
 ‘The old woman smearing millet gruel near the oldman’s anus composing songs, cried while singing.’ (Mahapatra and Zide no date, D.14)

Gtaʔ like some of the other South Munda languages, makes use of an objective case prefix in *a-*. As discussed in Anderson (2007), this is an archaic feature in Munda and in the Austroasiatic language family as a whole, with cognates in Mon-Khmer as well. It may appear with personal pronouns and demonstrative stems as well and often marks direct objects (or possessors).

- (45) *me a-nij bug-ke*
 he OBJ-1 beat-PAST
 ‘He beat me.’ (Bhattacharya 1975: 166)

The objective prefix in *a-* may combine with the case suffix in *-ke* in Gtaʔ as well. Here an indirect object (46) is the typical meaning found (or possession as well, see below) or objects of verbs of mental action (47).

- (46) (a) *dæk eʔ-baj næy ljo ceʔmwa na-big-la a-na-ke bol*
 look today-ABL I field grass 2-SOW-DS OBJ-YOU-CASE good
n-a-dij ɖæt ɖakce-ka bari kitoy-ke basoy-ke
 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god “look, you sow grass in my field (again), then I will not be nice to you.”’ (Mahapatra and Zide no date, I.28)
- (b) ***a-k-mæ-hij-ke*** *mmwiŋ coʔa-k[e]-ne ne-[m]biʔ ɖakce*
 OBJ-DEM-3PR-CASE one lame-T/A-DEP 1PL-give QUOT
bari basoy-ke
 again/also/then say-T/A
 ‘Let’s give them that lame one’. (Mahapatra and Zide no date, A.49)
- (47) (a) *coʔa=bir ne-biʔ-la k-mæ-hij a-næ-ke mæʔ mane-e be*
 Lame=gayil 1PL-give-DS DEM-3p-PLO BJ-we-CASE what obey-FUT Q
 ‘We will give them the lame gayil (but) will they obey us?’ (Mahapatra and Zide no date, A.35)
- (b) *k-mæ-hij ke-la olæŋ-læʔ-k[e]-ne remwa-hij*
 DEM-3PR-PL see-DS wander-AUX-T/A-DEP person-PL
a-næ-ke *k-mæ-hij a-miaʔ*
 OBJ-we-CASE DEM-3PR-PL NEG-know
 ‘They, (see you), are nomads, and they do not know us.’ (Mahapatra and Zide no date, A.36)

As aforementioned, there is a wide range of constructions to mark possessive clauses in Gtaʔ. Generally speaking, these consist of a copula verb or an element functioning as a copula with a case-marked noun phrase. In the negative, the negative

(present) copula is *ŋdu* (also the interjection meaning ‘No!’ or ‘so’/‘well’). Pronominal possessors in such constructions may be marked by the *a-* objective prefix (a), the *a-* prefix and the *-ke* suffix (b/c) or it may appear in the adessive form in *-nia* (d).

- (48) (a) *niã ni-læʔ-ke sina a-na ke-la knweʔ-huy*
 We.DL 1DL-remain-T/A though OBJ-you see-COND/DS wife-child
ŋdu a-næŋ knweʔ-huy ŋdu
 NEG.COP OBJ-I wife-child NEG.COP
 ‘Though we both live, you see, you don’t have a wife or child and I don’t have a wife or child.’ (Mahapatra and Zide no date, H.6)
- (b) *a-mæ-pa-ke go+gsiaʔ ŋdu*
 OBJ-3PR-DL-CASE child+monkey NEG.COP
 ‘They had no child[ren].’ (Mahapatra and Zide no date, M.2)
- (c) *a-mæ-hij-ke oron+paron ŋdu ki tmwaʔ d̥ãton ŋdu*
 OBJ-3PR-PL-CASE food+ECHO NEG.COP or mouth brushing NEG.COP
d̥æt d̥ij-ce k-mæ-hij olæŋ-læʔ-ke
 like.that do-SS DEM-3PR-PL wander-AUX-T/A
 ‘They have no food, no [way of] brushing teeth, they’ve been wandering around like that.’ (Mahapatra and Zide no date, A.33)
- (d) *eʔ-ke næʔ-nia swa ŋdu na wig-la-ce swa*
 Today we-ADESS fire NEG.COP you go.home-IMP-SS fire
toʔ-raŋ-la d̥ak-ce griŋ-ke toʔ=so bæ-ke
 pull.out-bring-IMP state-SS cat-CASE pull.out=fire send-T/A
 ‘Today we have no fire, you go home and bring the fire; saying thus, the bear sent the cat to bring fire.’ (Mahapatra and Zide no date, H.13)

Non-negative formations are also varied. In the past, the verb *læʔ-ge* may be found, in combination with the genitive suffix in *-ne*.

- (49) *mæ-hij-ne mbar-klig bũʔi læʔ-ge*
 3PR-PL-GEN two-CLSSFR buffalo COP-T/A
 ‘They had two buffaloes.’ (Mahapatra and Zide no date, F.2)

Circumlocutions are also possible, such as the borrowed verb *mile-* ‘be available’ which takes a dative/objective marked form of the logical possessor (*a-...-ke* with this pronominal stem).

- (50) *gte-la a-næʔ-ke gaʔi-si cili ccoŋ-ne mile-e*
 then OBJ-WE.INC-CASE more=day meat REDPL:eat-GEN available-FUT
 ‘Then we will have meat for eating for many days.’ (Mahapatra and Zide no date, D.12)

Keeping the distribution of these elements in mind, it appears that certain constructions in Gta? seem to have the formal properties of structures called, among various terms, ‘possessor raising’. One such construction aforementioned, is the use of *-ke* for nouns and *a-...-ke* for pronouns to mark possessors or primarily affected referents with certain body part formations.

- (51) *næŋ hweʔ-dʊlæŋ-n-læʔ-e na a-næŋ-ke pnomæŋ*
 I pretend-sleep-1-AUX-FUT you OBJ-1-CASE anus
p-n-aʔ-cweʔ-ʕiaʔ kig-biʔ
 millet.gruel pour.down-AUX
 ‘I’ll pretend to sleep, you smear some millet gruel around my anus.’
 (Mahapatra and Zide no date, D.8)

Such formations are also found when the body parts in question are incorporated into the verbal stem (see section 3.2.11) as in the following example.

- (52) *a-næʔ-ke kala-haʔ dɛkwa gæʔ-mu gæʔ=moaʔ-læʔ-e*
 OBJ-WE.INC-CASE daily like.this fry=nose fry=eye-AUX-FUT
 ‘They shall be frying our noses and eyes up like this every day.’ (Mahapatra and Zide no date, C.5)

Finally, note the following unusual structure in Gtaʔ which combines many of the features just mentioned. Here the ‘objective’ element *a-* seems to mark a pronominal possessor, or perhaps it is merely a pronominal ‘affected discourse referent’. Here the pronoun is marked *a-* alone (cf. Juang and North Munda *-a[ʔ]* genitive) but the noun it is associated with appears with the suffix *-ke*. It is as if the two elements were split across the two components of the ‘affected’ NP, with the pronoun taking the *a-* prefix as expected and the noun the *-ke* suffix. More analysis is required to determine how widespread this construction is in the language.

- (53) *ssæ dɛk dʒiŋ-ce a-næʔ remwa-hiŋ-ke sapa*
 Before like.this do-SS OBJ-WE.INC person-PL-CASE all
bagweʔ-coŋ-har-ke
 kill-eat-PL-T/A
 ‘Doing so earlier, they killed and ate up all of our people.’ (Mahapatra and Zide no date, D.28)

3.1.3 Person

A small number of stems in Gtaʔ, primarily kin-terms, (may) appear in an inalienably possessed form. For first and second person possessors, a personal prefix is used (corresponding to the subject prefixes/proclitics) while with third person possessors, a suffix is used instead which takes the form *-[ŋ]dɛ* after consonant-final stems and *-ɾæ* (for some speakers) after vowel-final ones. Thus, templatically we find

- (54) 1st/2nd person: 1/2-Noun[-PL/DL-CASE]

- (55) 3rd person: Noun-3-[PL/DL-CASE]

with number suffixes and case suffixes following in that order.

- (56) (a) *pe-ntaʔ* (b) *næ=dʒre-hiŋ*
 2_{PL}-grandfather 1_{PL}-person-PL
 ‘your grandfather’ ‘our people’
 (Mahapatra and Zide no date, L.41) (Mahapatra and Zide no date, F.11)
- (c) *lʒo haboʔ-bo-la huŋ-dæ-ke bæ-ke*
 land forget-AUX-DS child-3-CASE send-T/A
 ‘It was forgotten at the field so he sent his child (for it).’ (Mahapatra and Zide no date, I.5)

The possessed form of the noun may occur with a genitive-marked [pro]noun that refers to the possessor of the possessed noun.

- (57) *mæ-ne knweʔ-ɾæ nɕiaʔ kuma-ge*
 3PR-GEN wife-3 water bath-T/A
 ‘His wife gave birth.’ (Mahapatra and Zide no date, B.2)

Sometimes the inalienable possession form seems to correspond to a definite referent rather than one that is explicitly possessed or referred to as such in that context:

- (58) (a) *nɕia kuma-ce nji harke ɕij læʔ-la knweʔ-ɾæ gweʔ-we-ge*
 Water bath-SS three month do COP-DS wife-3 die-go/AUX-T/A
 ‘(after) she gave birth, three months passed and the wife died.’ (Mahapatra and Zide no date, B.4)
- (b) *meʔ=swa-ne-ɾak-ke gweʔ-k[ɛ]-ne knweʔ-ɾæ ɕwa wig-ce*
 One=day-DEF-state-T/A die-T/A-DEP wife-3 home come.home-SS
huŋ-ɕæ-ke aʔ-bbuʔ+aʔ-ccoŋ-ce
 child-3-CASE CAUS-REDPL:SUCK+CAUS-REDPL:eat-SS
bole+hmaʔ ɕwe+twɛʔ-ce huŋ-ɕæ-ke aʔcoŋ+aʔug bari
 rice+curry cook+serve-SS child-3-CASE feed+give.drink also
huŋ-ɕæ-ke basoŋ-ke
 child-3-CASE say-T/A
 ‘It is said that one day the dead wife came home to feed and suckle her/the children cook and serve up rice and curry, she fed her/the children and then said to the (older) child’ (Mahapatra and Zide no date, B.8)

With certain stems, there is an assimilation of [e/æ] --> [i] in the possessive prefix form; compare the forms for ‘your father’ and ‘your mother’ below.

- (59) (a) *pe-mba wig-ce a-pe-ke saliaʔ+ku-la*
 2PL-father come.home-SS OBJ-2PL-CASE ask+TAG-COND/DS
ni-iaŋ wig-læʔ-ge ɕakce a-basoŋ-ge-pa
 1PL-mother:VOC come.home-AUX-T/A QUOT NEG-say-PROHIB-2DL
 ‘When your father comes back home and asks you, do not say “our mother has come home.”’ (Mahapatra and Zide no date, B.9)
- (b) *ɕokri basoŋ-ke pe-mba bari poʔ-la caŋ*
 Old.woman say-T/A 2PL-father again stab-DS SOUND:SYMBOLIC
ɕak-ke pi-iaŋ
 state-T/A 2PL-mother
 ‘The old woman said “your father” and he stabbed again, “caŋ” it sounded and she said “your mother”.’ (Mahapatra and Zide no date, L.23)

Note that certain kin-terms appear in numerous forms in Gta?, for example, some have derived or suppletive vocative forms (e.g. *iaŋ* ‘mother’, *aba* ‘father’). Note that with second plural possessors ‘father’ appears in the stem-form *-mba* (e.g. *pe-mba*) with a syllabic nasal prefix that has derived the free form of the noun while third person possessors, on the other hand, take the variant *bba* (e.g. *bba-ɾæ*) with a reduplicated free form.²

- (60) *wig-ce* *bba-ræ-ke* *basoŋ-ke* “*nɖu aba* *næ?* *ljo*
 go/come-SS father-3-CASE say-T/A well father:VOC we:INC land
ja-ŋɖe-ha? *kia* *bbig-ɖij-ke* *ɖakce*
 who-3-EMPH paddy REDPL:SOW-AUX-T/A QUOT
bba-ræ-ke *basoŋ-ke*
 father-3-CASE say-T/A
 ‘She came home and told her father “well daddy, is someone (supposed to be) sowing our field?”’ (Mahapatra and Zide no date, I.10)

3.1.4 Definiteness

Definiteness as a morphologically expressed category is not *per se* a feature of the grammar of Gta?. There are certain uses of what appears to be the genitive case marker in a small number of constructions without any possessive interpretation possible. These seem to function as some kind of emphatic or definite marker. As briefly noted above, sometimes the possessive marking on nouns themselves seem to carry a connotation of definiteness rather than possession *per se* as well. The use of this genitive-definite marker in Gta? is found most typically with two kinds of expressions, one with demonstrative stems (61) and the other in certain kinds of (usually clause-initial) adverbial expressions found in, for example, formulaic expressions in tales (62).

- (61) (a) *ɖo kra* *a-tæŋ-ne* *gbug=ci* *e-liŋ-raŋ-ke* *nluŋ=bug* *mari*
 Old.man OBJ-that-DEF pig=meat go-lift-bring-T/A ear=pig again
gwa?=luŋ *raŋ-wig-ke*
 cut=ear bring-come.home-T/A
 ‘The old man went and brought that pig meat and cut and brought back home [pieces of] the pig’s ear.’ (Mahapatra and Zide no date, M.7)
- (b) *ɖæt* *ne?-ɖij* *sina* *pe* *pe-ŋ-ɖo-la* *nor+nara-hij*
 That.like 1PL.INC-do though you.PL 2PL-AUGM-run-DS human.being+ECHO-PL
ggwe?-we-har-e *ɖakce* *kæŋ-ne* *ɖira* *basoŋ-ke*
 REDPL:die-AUX-PL-FUT QUOT this-DEF/GEN millet.species say-T/A
 ‘“Let’s do it like this (as an acceptable alternative), you run away, the human beings will all die” this millet said.’ (Mahapatra and Zide no date, C.16)
- (c) *mæ-hij* *int-baŋ* *janwai hni* *hātar-ce* *tarpa-ɖæŋ*
 3PR-PL that.side-ABL janwai village abandon-SS ceremonial.drum
liŋɖi=kwi *ale-tæŋ-ne* *sorte* *sa?-ce* *at*
 ceremonial.cooking.pot all-that-DEF entirely hold-SS there
jar-ce *haɖa?-koŋɖa* *basa-har-ke*
 climb.down-SS vicinity-mountain reside-PL-T/A
 ‘From that side they abandoned the village of Janwai, took up all their ceremonial drums and pots, climbed down and settled in the vicinity of the mountain.’ (Mahapatra and Zide no date, F.6)
- (d) *koŋɖa-ne* *nla?* *koŋɖa-ne* *pa?ɖu* *koŋɖa-ne* *cucu*
 mountain-GEN tuber mountain-GEN leaf mountain-GEN fruit
ale-kæŋ-ne *coŋ-coŋ-ce-ka* *bari* *mæ-hij* *brwa-læ?-ge*
 all-this-DEF COPY-eat-SS-only again 3PR-PL live-AUX-T/A

‘They lived only eating tubers, leaves and fruits available in the mountains.’
(Mahapatra and Zide no date, G.3)

- (62) *meʔ-swa-ne kia-manda-hij rsiŋ-ke bari ho-baʔlir-ke*
One-day-DEF paddy-group-PL gather-T/A again RECIP-talk-T/A
‘One day all the paddies assembled and spoke with each other.’ (Mahapatra
and Zide no date, C.4)

Other words or forms may appear with this genitive/definite marking in Gta? as well, for example, adjectival or demonstrative stems (a/b) or non-finite verb forms (c), where it marks a kind of infinitival structure.

- (63) (a) *akæŋ koŋɖa-ne gsiaʔ=o nu-n[e] koŋɖa we-ke*
This mountain-GEN monkey=child FAR-DEF mountain go-T/A
geʔ-la o gsiaʔ=o-hij pe-ntaʔ gweʔ-we-ge
come-IMP VOC monkey-child-PL 2PL-grandfather die-AUX-T/A
o-hoŋ o-hoŋ nu-n[e] koŋɖa-ne ŋko=o atæŋ
VOC child VOC child FAR-DEF mountain-GEN peacock=VOC that
koŋɖa we-ke geʔ-la o ŋko-o-hij pe-ntaʔ
mountain go-T/A come-IMP VOC peacock-VOC-PL 1PL-grandfather
gweʔ-we-ge o-hoŋ o-hoŋ
die-go-T/A VOC child VOC child
‘Oh monkeys of that far-off mountain, come here, your grandfather is dead;
the peacocks of that far-off mountain went away to that mountain yonder,
come back oh peacocks, your grandfather is dead.’ (Mahapatra and Zide no
date, D.15)
- (b) *ljo we-la mmwiŋ mæʔ-bare baya ɖokra sgwa ke-ce*
field go-DS one what-INTERJ mad old.man like look-ss
giŋ=siaʔ giŋ=siaʔ nluŋ-ce clæ clæ ugboʔ-ce-ka
cheek-monkey cheek-monkey ear-ss long long hair-ss-only
bari moʔmnæŋ-ne buti bwaʔ=tar-ce mmwiŋ
again very.big-DEF basket carry=shoulder-ss one
unæŋ=ula+toʔ-ceʔ-ka mæ big=cog ɖæte bbig-ɖiŋ-ke ljo
cigarhold.in-lips-ss-only s/he sow=basket with REDPL:SOW-AUX-T/A land
‘She went to the field, an old man looking like a cheek monkey was there
with long hair, carrying a basket on his shoulder, with a cigar in his mouth
and was sowing the field with a basket. (Mahapatra and Zide no date, I.8)
- (c) *bari jibon=re llæʔ-ne ʔan a-mile-ke*
again life=man REDPL:remain-DEF place NEG-available-T/A
‘There was no place for the soul to stay’ (Mahapatra and Zide no date, J.18)

It may be in this latter function that this element has found its way into the make up of the relative clause or participial form discussed in sections 3.2.8 and 4.2.1.

A small number of other elements may be included in this vaguely discourse based notion of definiteness. For example, the *-te* in the following form may be a frozen (locational/directional) case form or it may be a variant of *-ne*. Further the suffix or clitic *-noʔ* in the form below may be another example of a vaguely emphatic definiteness marker in Gta?. Much remains to be worked out in this domain of Gta? grammar.

- (64) *qæt qij-ce mba-ya-no? at-te wij-ha?+rij-ha?-har-ke*
 like.that do-ss two-person-DEF there-DEF rebuke-EMPH+fight-EMPH-PL-T/A
 ‘Like that the two people quarreled there.’ (Mahapatra and Zide no date, I.21)

3.1.5 Class/gender

Noun classes or gender are not active categories in the morphosyntax of Gta? *per se*. That said, there are some means of encoding various kinds of class- or gender-like distinctions in the language. A small set of mainly loan words appear in masculine and feminine varieties, often contrasting masculine *-a* and feminine *-i*, directly reflecting the system found in local Indo-Aryan varieties, for example, *qokra* vs. *qokril/qukri*. If an ambiguous animal term needs to be specified for gender, then classifiers can be used to disambiguate these, for example, *bbo-* for masculine for example, *grij* ‘cat’ *bbo-rij* ‘tomcat’ *gsu?* ‘dog’ *bbo-su?* ‘male-dog’ or *dja[N]-* for feminine *gmi?* ‘goat’ *diami?* ‘nanny’ *birbo* ‘buffalo bull’ vs. *diambo* ‘buffalo cow’. As is evident from the alternations above, it is clear that if a noun-derivational prefix is found in the gender-neutral term, this is replaced by the gender classifier prefix in these gender-specific forms.

There is also a combining form or suffix *-/=o* that seems to have masculine reference as well. Its use and distribution remains a subject for future research.

- (65) *goɾæ-o ljo we-ge*
 child-MASC field go-T/A
 ‘The child went to the field.’ (Mahapatra and Zide no date, I.7)
- ngir-boe?-o we-ke-la mæ dij wig-ce*
 male-child-MASC go-see-COND s/he also come.home-ss
- bba-ɾæ-ke basoŋ-ke*
 father-3-CASE say-T/A
 ‘When the male child went and saw, he came home and told his father.’
 (Mahapatra and Zide no date, I.13)

In addition, a covert or morphosyntactic distinction between human/animate and inanimate is also found in certain people’s speech. Thus for some speakers, only human nouns may be marked for number (dual or plural) or with the ‘objective’ or ‘dative’ case suffix *-ke*, while for others any animate noun can be so marked (and trigger plural verb agreement). For most speakers, plural marking, *-ke* object marking and especially plural verb agreement is found with inanimates only if these have been anthropomorphized, for example, talking rice paddies and millets from traditional narratives.

3.1.6 Pronouns

Gta? has a ten-member pronominal system consisting of three persons (first, second, third) and three numbers (singular, dual and plural) and an inclusive form of the first person dual and plural. As aforementioned, third dual is marked by a combination of the third singular pronoun stem and the second dual marker *-pa*. The pronoun forms are given in (66) and examples of their use follow. Note that Bhattacharya (1975)

has *nij* for the first person, and Ghosh (2003) follows this (though his examples mainly show *ney* instead). This may be a variant of the form found in our materials *næj*, but may also reflect the influence of Remo on the speech of their consultant[s], where *nij* is the normal first person singular form (see Anderson and Harrison this volume). Note also the subtle distinction between the exclusive and inclusive first person plural pronouns [*næ* vs. *næʔ*]

(66)	SG	DL	PL	
1	<i>næj</i>	<i>niaʔ</i> [INCL]	<i>næ</i>	<i>næʔ</i> [INCL]
2	<i>na</i>	<i>pa</i>	<i>pe</i>	
3	<i>mæ</i>	<i>mæpa</i>	<i>mæhij</i>	

Examples of the use of pronouns in *Gtaʔ* in various functions are offered in the following sections.

(67) *næj*
te-la tæn kitoŋ basoŋ-ke næj-ne paiŋi kæn-haʔ, na di-rokom
 that-DS that god say-T/A I-GEN work this-EMPH you what-manner
kia na-big-ke næj diŋ se-rokom ceʔmwa bbig-diŋ-ke
 paddy 2-SOW-T/A I also that-manner grass REDPL:SOW-AUX-T/A
 ‘Then that god said “my task is this: whatever manner you sow in, I will sow in that manner too.”’ (Mahapatra and Zide no date, I.17)

(68) *na*
 (a) *na kala-haʔ andi we-na-miaʔ-ke iaŋ*
 You daily-EMPH where go-2-CUST-T/A mother:VOC
 ‘Mother, where do you go everyday?’ (Mahapatra and Zide no date, B.13)
 (b) *dæk eʔ-baŋ næj ljo ceʔmwa na-big-la a-na-ke*
 look today-ABL I field grass 2-SOW-DS OBJ-you-CASE
bol n-a-diŋ dæŋ ɖakce-ka bari kitoŋ-ke basoŋ-ke
 good 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god “look, you sow grass in my field (again), then I will not be nice to you.”’ (Mahapatra and Zide no date, I.28)

Note that the agreement prefixes of *Gtaʔ* are in fact agreement markers, as they are used even in the presence of an overt subject pronoun. Note also that the uninflected form of the pronoun may serve as a genitive or possessive form in a small number of instances (cf. the inalienably possessed forms mentioned in 3.1.3).

(69) *kitoŋ basoŋ-ke ʔna ljo eʔ-baŋ næj n-a-big ɖakce basoŋ-ke*
 god say-T/A you field today-ABL I 1-NEG-SOW QUOT say-T/A
 ‘The god says “from today on I will not sow your field any longer.”’
 (Mahapatra and Zide no date, I.29)

(70) *næ*
tela eʔ næ diŋi ne-diŋ kitoŋ-ke mæʔ-ɽig ne-pag=ra
 then now we how 1PL-do god-CASE what-ADD 1PL-break=twig
pag=ra-ce-ka bari tæn koŋɖa e-go ne-ke ɖakce basoŋ-ke
 break=twig-ss-only again that mountain go-hunt 1PL-AUX QUOT say-T/A
 ‘Then what are we to do now? Let’s make an offering of something to the god and then let’s try to go hunting on that mountain.’ (Mahapatra and Zide no date, A.51)

- (71) *næ?*
e? *mæ?-sa?* *næ?* *ne?-we-e* *ɖakce* *ja-ɽig* *a-we-ke*
 today what-PURP WE.INC 1PL.INC-GO-FUT QUOT who-ADD NEG-GO-T/A
 ‘“For what will we go today?” (they said) and no one came.’ (Mahapatra and Zide no date, D.29)
- (72) *niã = nia[?]*
niã *ni-læ?-ke* *sina* *a-na* *ke-la* *knwe?-huŋ* *nɖu*
 We.DL 1DL-remain-T/A though OBJ-YOU see-COND/DS wife-child NEG.COP
a-næŋ *knwe?-huŋ* *nɖu*
 OBJ-I wife-child NEG.COP
 ‘Though we both live, you see, you don’t have a wife or child and I don’t have a wife or child.’ (Mahapatra and Zide no date, H.6)
- (73) *pe*
 (a) *pe* *sapa* *remwa* *pe-ɖulæŋ-ke* *næŋ* *ko-n-læ?-ge*
 You.PL all person 2PL-sleep-T/A I sit-1-AUX-T/A
 ‘(while) you people all slept, I (alone) sat (awake all night).’ (Mahapatra and Zide no date, A.44)
 (b) *gte-la* *k-mæ* *basoŋ-ke* *pe* *a-ɖo-ge-pe* *ge?-la-pe*
 this-DS DEM-3PR say-T/A YOU.PL NEG-flee-PROHIB-2PL come-IMP-2PL
ɖag-la-ŋɖig *ja-ɽig* *a-we-ke*
 state-DS-ADD who-ADD NEG-GO-T/A
 ‘Then he said “don’t flee, please come,” but nobody came/went.’ (Mahapatra and Zide no date, J.23)
 (c) *ɖæt* *ne?-ɖiŋ* *sina* *pe* *pe-ŋ-ɖo-la* *nor+nara-hiŋ*
 that.like 1PL.INC-do though you.PL 2PL-AUGM-run-DS human.being+ECHO-PL
ggwe?-we-har-e *ɖakce* *kæn-ne* *ɖira* *basoŋ-ke*
 REDPL:die-AUX-PL-FUT QUOT this-DEF millet.species say-T/A
 ‘“Let’s do it like this though, you run away, the human beings will all die”
 this millet said.’ (Mahapatra and Zide no date, C.16)
- (74) *pa*
pa *coŋ-pa*
 you.DL eat-2DL
 ‘eat you 2!’ (Mahapatra *et al.* 1989)

While unit/unanalyzable elements are found with first and second person non-singular forms, these are clearly agglutinative and combinatorial in form for third person pronouns.

- (75) *mæ/mæ-pa/mæ-hiŋ*
 (a) *mæ* *we-ce* *basoŋ-ke* ‘*na* *mæ?* *bbig* *na-ɖiŋ-ke?*
 he go-SS say-T/A ‘you what REDPL:SOW 2-AUX-T/A’
ɖakce *salia?+ku-ke*
 QUOT ask+ECHO-T/A
 ‘He (=father) goes and says “what are you sowing?” he asked.’ (Mahapatra and Zide no date, I.16)
 (b) *mæ-ne* *knwe?-ɽæ* *nɖia?* *kuma-ge*
 3PR-GEN wife-3 water bath-T/A
 ‘His wife gave birth.’ (Mahapatra and Zide no date, B.2)

- (c) *gmi? basoŋ-ke d̪æt n̪ɖu kæn k̄ɔ̄ɖa-ke ja*
 goat say-T/A that.like no this mountain-CASE who
pwetur-æɡ-aʔ-caʔ-la mæ ssæ paŋ-coŋ-e
 surround-shit-CAUS-arrive-DS he before come-eat-FUT
 ‘Goat said “No, not like that, (I suggest) whoever can surround the mountain in shit, let him come and eat first.”’ (Mahapatra and Zide no date, II.11)
- (d) *mæ-pa mba-ya hara-hari d̪iŋ-ce k̄ɔ̄ɖa pwetur-æɡ-har-ke*
 he-DL two-person defeat:ECHO AUX-SS mountain surround-shit-PL-T/A
 ‘the two of them trying to beat each other, went around the mountain shitting.’ (Mahapatra and Zide no date, II.12)
- (e) *mæ-hiŋ int-baŋ janwai hmi h̄atar-ce tarpa-ɾæŋ*
 3PR-PL that.side-ABL janwai village abandon-SS ceremonial.drum
lind̪i=kwi ale-tæn-ne sorte saʔ-ce at
 ceremonial.cooking.pot all-that-DEF entirely hold-SS there
jar-ce haɾaʔ-kon̪ɖa basa-har-ke
 climb.down-SS vicinity-mountain reside-PL-T/A
 ‘From that side they abandoned the village of Janwai, took up all their ceremonial drums and pots, climbed down and settled in the vicinity of the mountain.’ (Mahapatra and Zide no date, F.6)

Beside personal pronouns, a range of other pronominal elements are found in Gta?. These include interrogative pronouns, indefinite pronouns, and demonstratives. The demonstratives are discussed in 3.1.7, but the other categories are briefly introduced here.

The interrogative or Wh-pronouns in Gta? form the basis for the indefinite, etc. pronouns so are introduced first. A basic inanimate/animate (or perhaps human/non-human) distinction is maintained between the referents of *mæ?* ‘what’ and *ja* ‘who’.

- (76) *mæ?*
ŋku mæ? lig-ge d̪ak-ce hur-hur a-ia-ke
 tiger what work-T/A state-SS wait-wait NEG-CAP-T/A
 ‘The tiger saying “what did he do? Could not wait any more”.’ (Mahapatra and Zide no date, H.17)

Combined with the case suffix *-saʔ/-sar* ‘for’, the interrogative ‘why, for what reason’ is created.

- (77) (a) *eʔ mæʔ-saʔ n̄eʔ neʔ-we-e d̪akce ja-ɾiɡ a-we-ke*
 today what-PURP we.INC 1PL.INC-go-FUT QUOT who-ADD NEG-go-T/A
 ‘“For what will we go today?” (they said) and no one came.’ (Mahapatra and Zide no date, D.29)
- (b) *mæʔ-sar n-a-big d̪akce mæ basoŋ-ke*
 what-PURP 1-NEG-SOW QUOT he say-T/A
 ‘“Why shouldn’t I sow?” He said.’ (Mahapatra and Zide no date, I.20)

The bare form of this pronoun may carry this kind of interrogative meaning as well in a limited number of instances.

- (78) *ɖæʔ basoŋ-ke dapre-haʔ 'mæʔ næŋ ljo na-big-ke'*
 like.that say-T/A instantly-EMPH what I land 2-sow-T/A
ɖakce kmæ basoŋ-ke
 QUOT this.one say-T/A
 'Right away he spoke thus "why do you sow my land?"' (Mahapatra and Zide no date, I.19)

The basic form of the interrogative pronoun *mæʔ* may also carry an indefinite sense when used in certain kinds of rhetorical questions (with the question particle *be*), that is, 'anything' or 'something'. Usually this is marked by the additive focus marker *-riɡ* etc.

- (79) (a) *coʔa=bir ne-biʔ-la k-mæ-hiŋ a-næ-ke mæʔ mane-e be*
 lame=gayil 1PL-give-DS DEM-3PR-PL OBJ-we-CASE what obey-FUT Q
 'We will give them the lame gayil (but) will they obey us about anything?'
 (Mahapatra and Zide no date, A.35)
- (b) *tela eʔ næ ɖiɖi ne-ɖiŋ kiŋoŋ-ke mæʔ-riɡ ne-paɡ=ra*
 then now we how 1PL-do god-CASE what-ADD 1PL-break=twig
paɡ=ra-ce-ka bari tæŋ koŋɖa e-go ne-ke
 break=twig-ss-only again that mountain go-hunt 1PL-AUX
ɖakce basoŋ-ke
 QUOT say-T/A
 'Then what are we to do now? Let's make an offering of something to the god and then let's try to go hunting on that mountain.' (Mahapatra and Zide no date, A.51)

In negative structures (i.e. with negative verb forms) in combination with the emphatic particle, adverb, or additive focus marker, *ɖiɡ/riɡ*, negative pronouns of the 'nothing' sort are produced. Note that while usually together, the pronoun and the particle may be separated as well.

- (80) *mmwiŋ-mæʔ plig+ɡɾaʔ riɡ a-mile-ke*
 one-what bird+mouse ADD NEG-available-T/A
 'Nothing, not even a bird or mouse was available (to be hunted)' (Mahapatra and Zide no date, A.10)

Such negative pronoun forms are common with the animate pronoun *ja* 'who' in our materials.

- (81) (a) *kæn ɖokri jiwa-manda-ke e-wa-la ja-riɡ a-we-ke*
 this old.woman animal-herd-CASE go-call-DS who-ADD NEG-go-T/A
 '(When) the old woman called the animals, [but] no one came.' (Mahapatra and Zide no date, D.27)
- (b) *eʔ mæʔ-saʔ næʔ neʔ-we-e ɖakce ja-riɡ a-we-ke*
 today what-PURP we.INC 1PL.INC-go-FUT QUOT who-ADD NEG-go-T/A
 "'For what will we go today?" (they said) and no one came.' (Mahapatra and Zide no date, D.29)

Indefinite pronouns of the 'whoever'-type are also formed with non-negative verbs in a complex structure, akin to relative-correlative structures familiar from languages

of the region. These consist of two parallel clauses, one headed by *ja* and one by the third person pronoun *mæ*.

- (82) (a) *tte-la hīsa nḍu-kuḥa ja par-le mæ paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come
ccoŋ dḥiŋ-le
 REDPL:eat AUX-OPT
 ‘Thus there will be no envy; whoever wins, let him come and eat.’
 (Mahapatra and Zide no date, II.7)
- (b) *gbe basoŋ-ke akæn konḍa ja olæŋ aʔ-caʔ-la*
 Bear say-T/A this mountain who walk CAUS-arrive.at.destination-DS
mæ ssæ uli-alo tʰwā-læʔ-leʔ
 he before mango-under stand-remain-OPT
 ‘Bear says, “whoever gets to this here mountain first, let him stand under the mango tree.”’ (Mahapatra and Zide no date, II.9)

Other interrogatives include *oʔna* ‘when’ *oʔri* ‘how many’/‘how much’ and *anḍi* ‘where’ -*rokom* ‘what manner’. Like *mæʔ* and *ja*, these may occur as indefinites with positive and negative structures involving the particle/clitic/additive focus element -*riŋ* (etc.). Some examples of their use follow in both interrogative structures (83) and indefinite and correlative structures (84), not in combination with personal pronouns, but with subtypes of demonstratives.

- (83) (a) *anḍi we-pe-ḥdḥiŋ-ke ḍakce basoŋ-ke*
 where go-2PL-do-T/A QUOT say-T/A
 ‘“Where are you going?” he says.’ (Mahapatra and Zide no date, C.11)
- (b) *na kala-haʔ anḍi we-na-miaʔ-ke iaŋ*
 You daily-EMPH where go-2-CUST-T/A mother:VOC
 ‘Mother, where do you go everyday?’ (Mahapatra and Zide no date, B.13)
- (c) *mmwiŋ goʔæ=o basoŋ-ke oʔna-baŋ aʔsi-ge*
 One boy say-T/A when-ABL fever-T/A
 ‘One boy asked “since when has he been sick?”’ (Mahapatra and Zide no date, M.12)
- (84) (a) *ḍukri gewa=mwaʔ-riaʔ-ce basoŋ-ke ḥḍu ḍokra mari*
 old.woman shy=eye-ECHO-SS say-T/A no old.man again
oʔna-riŋ ũ-we-e
 when-also 1-go-FUT
 ‘The old woman was embarrassed and said “No, old man, I will go there again sometime.”’ (Mahapatra and Zide no date, L.26)
- (b) *ḥriŋ oʔri=mwa sgwa we-la ḍokra gweʔ=we-ge*
 later.on how.much=year like go-DS old.man die=AUX-T/A
 ‘Later on, after like several years passed, the man died.’ (Mahapatra and Zide no date)
- (c) *bba-ʔæ basoŋ-ke nḍu anḍi-anḍi-ne si-si=re-ḥiŋ at gali*
 father-3 say-T/A no COPY-where-GEN plough=man-PL there way
wig-dḥiŋ-e ḍakce-ka bari basoŋ-la-ḥḍig a-mane-ke
 go.home-AUX-FUT QUOT-only again say-DS-ADD NEG-agree-T/A
 ‘Her father disagreed, “no, ploughmen might be passing through on their way home.”’ (Mahapatra and Zide no date, I.11)

- (d) *te-la tæŋ kitoŋ basoŋ-ke næŋ-ne paiŋi kæn-haʔ, na ɟi-rokom*
 then that god say-T/A I-GEN work this-EMPH you what-manner
kia na-big-ke næŋ ɟig se-rokom ceŋmwa bbig-ɟiŋ-ke
 paddy 2-sow-T/A I also that-manner grass REDPL:SOW-AUX-T/A
 ‘Then that god said “my task is this: whatever manner you sow in, I will
 sow in that manner too.”’ (Mahapatra and Zide no date, I.17)

In the last example, the form *ɟi-rokom* bears mention. Its first element is something that appears in an emphatic interjection-like interrogative in *ɟiɟi[-le]*. Some examples of the use of this latter element are offered below.

- (85) (a) *hɽiŋ ɟi-le mæ jetek brwa-læʔ-ge oʔ-ɽi=si sgwa*
 afterwards what-INTERJ he as.many live-AUX-T/A how.many=day like
læʔ-ge oʔ-tæn=si jako ceŋmwa mulke-se poʔ-ɽiaʔ nɟu
 remain-T/A that.many=days till grass totally-DEF sprout NEG.COP
 ‘Afterwards, however long he lived, no grass ever grew up there again.’
 (Mahapatra and Zide no date, I.30)
- (b) *wō-ɟæ wa=huy we-ge; we-ce basoŋ-la ŋgire*
 Mother-in-law-3 call=child go-T/A go-SS say-DS young.man
basoŋ-ke me=mwa-ne sle ɟiɟi meʔ=swa n-aʔ-lug-e
 say-T/A one=year-GEN work how one=day I-CAUS-finish-FUT
 ‘The mother-in-law went to call the son-in-law and speak (with him); the
 young man said “how can I do one year’s work in just one day?”’ (Mahapatra
 and Zide no date, K.8)
- (c) *bari meʔ-swa-ne ɟiɟi-le haŋɟa-ŋɟæ pag=liʔ we-ke*
 then one-day=DEF how-INTERJ husband-3 break=shoot go-T/A
 ‘Then one day what happened but the husband went for shoots.’ (Mahapatra
 and Zide no date, G.7)

3.1.7 Demonstratives

The demonstrative system of Gtaʔ, like that of many other Munda languages is complex, indeed among the more elaborated of such systems found cross-linguistically. For a fuller description the reader is referred to Zide (1968, 1972). The basic opposition is between *k[æŋ]* ‘this’ or proximal (86) and *t[æŋ]* (87) ‘that’ (unmarked distal) but many other forms are attested. These demonstratives appear as modifiers in pre-nominal position.

- (86) *tæŋ*
- (a) *ɟæk basoŋ-la tæŋ ŋgire bari wō-ɟæ*
 Like.this say-DS that young.man again mother-in-law-3
plaʔ wig-ke
 behind come.home-T/A
 ‘She spoke like this and that young man came back home behind (=following)
 his mother-in-law.’ (Mahapatra and Zide no date, K.11)
- (b) *hni-[ŋ]-[ɟ]re basoŋ-ke næ ne-we-e tæŋ bir=tia-ke ne-twiŋ-e*
 village-GEN-person say-T/A we IPL-go-FUT that gayil-CASE IPL-shoot-FUT
 ‘The villagers said, “we will go, we will shoot down that gayil.”’
 (Mahapatra and Zide no date, A.4)

- (c) *bari tæn miŋdʒik at-te kndæŋ+hni aʔ-bog-ce-ka bari*
 again that night there drum+ECHO CAUS-beat-ss-only again
para-miŋdʒik waʔ=ræŋ+waʔ=so waʔ-ce njo njir
 whole-night dance=drum-dance=ECHO dance-ss next.day morning
sra-ke bir=tia-ke tæ-raŋ-we-ce
 deer-CASE gayil-CASE carry.on.head-bring-go-ss
ggwaʔ+ttæg-ce klig hni baʔa+kuʔa coŋ-ke
 REDPL.cut+REDPL.make.piece-ss whole village share+ECHO eat-T/A
 ‘Then that night they danced beating the drum and then on the morning of
 the next day they slaughtered the gayil and deer and prepared the meat and
 the whole village ate together.’ (Mahapatra and Zide no date, A.61)

(87) *kæn*

- (a) *kæn dʒokri jiwa-manda-ke e-wa-la ja-ʒig a-we-ke*
 this old.woman animal-herd-CASE go-call-DS who-ADD NEG-go-T/A
 ‘(when) the old woman called the animals, [but] no one came.’ (Mahapatra
 and Zide no date, D.27)
- (b) *koŋda-ne nlaʔ koŋda-ne paʔdu koŋda-ne cucu*
 mountain-GEN tuber mountain-GEN leaf mountain-GEN fruit
ale-kæn-ne coŋ-coŋ-ce-ka bari mæ-hiŋ brwa-læʔ-ge
 all-this-DEF eat-eat-ss-only again 3PR-PL live-AUX-T/A
 ‘They lived only eating tubers, leaves and fruits available in the mountains.’
 (Mahapatra and Zide no date, G.3)

A number of other categories appear to be derived from these basic ones. One such form appears to be derived from a stem *dæ-k/dæ-t* which forms adverbials ‘like this’ ‘like that’ that often appear in clause-initial position in narratives.

(88) *dæk*

- (a) *dæk basoŋ-la tæŋ ŋgire bari wõ-dæ*
 Like.this say-DS that young.man again mother-in-law-3
plaʔ wig-ke
 behind come.home-T/A
 ‘She spoke like this and that young man came back home behind (=following)
 his mother-in-law.’ (Mahapatra and Zide no date, K.11)
- (b) *dæk dʒiŋ-ce iʔtuʔa bole aʔ-dʒiŋ-ce mbar pwaʔ-coŋ*
 Like.this do-ss little cooked.rice CAUS-do-ss two leaf.cup
nturiaʔ aʔ-dʒiŋ-ce næ gtaʔ=re-hiŋ ccoŋ+uug-miaʔ-ke
 gruel CAUS-do-ss we Gtaʔ-person-PL REDPL.eat-REDPL.drink-CUST-T/A
 ‘And thus we did it this way, we Gtaʔ people used to make two leaf-cups
 and eat [only] a little cooked rice with gruel.’ (Mahapatra and Zide no
 date, C.23)

(89) *dæt*

- dæt dʒiŋ-ce mba-ya-noʔ at-te wiŋ-haʔ+riŋ-haʔ-har-ke*
 like.that do-ss two-person-DEF? there rebuke-EMPH+fight-EMPH-PL-T/A
 ‘Like that the two people quarreled there.’ (Mahapatra and Zide no
 date, I.21)

When used referentially without an accompanying noun that it modifies, the proximal demonstrative element *k-* may be added to the third person pronominal stem *-mæ* (inflected in any number). This corresponds to the use of ‘this one’, etc. in English.

(90) *kmæ*

(a) *na kia aʔ-ble na-coŋ-la næŋ ceʔmwa n-a-raʔ-ble*
 you paddy CAUS-ripen 2-eat-COND I grass 1-NEG-CAUS-ripen

ɖakce kmæ basoŋ-ke
 QUOT this.one say-T/A

‘“You cultivate paddy and eat it; shouldn’t I cultivate grass” this one said.’
 (Mahapatra and Zide no date, I.18)

(b) *kmæ-hiŋ ɖhã kuru=paka we-har-ke*
 this:one-PL small pit=side go-PL-T/A

‘These people went to the small-pit side’ (Mahapatra and Zide no date, A.5)

A number of other demonstrative forms are attested in our materials. One of these is *atte* ‘there’ which may have an archaic retention of an older now lost or lexicalized local/directional case *-te* in Gtaʔ.

(91) *atte hur-k-ne-nia ɖig go-k-ne-nia ɖig bir=tia-ɖig*

there guard-T/A-DEP-ADESS also hunt-T/A-DEP-ADESS also gayil-ADD

a-tar-ke

NEG-come.out-T/A

‘The gayil did not appear there, neither near the guard[s] nor the hunter[s].’
 (Mahapatra and Zide no date, A.7)

Other demonstrative elements found in Gtaʔ include forms such as *oʔ-kæn* ‘this much’ (92) and *int* ‘that side’ (93).

(92) *k-mæ-hiŋ oʔ-kæn ɖonɖ+kost ɖiŋ-ce*

DEM-3PR-PL this.much punishment+trouble do-SS

bar=si+jir=si

ɖɖiŋ-ɖiŋ-ke

two=day+three=day REDPL:do-AUX-T/A

‘These ones were working with so much trouble the last two or three days.’
 (Mahapatra and Zide no date, A.32)

(93) *mæ-hiŋ int-baŋ janwai hni hãtar-ce tarpa-ɽæŋ*
 3PR-PL that.side-ABL janwai village abandon-SS ceremonial.drum

linɖi=kwi

ale-tæn-ne

sorte

saʔ-ce

at

ceremonial.cooking.pot all-that-DEF entirely hold-SS there

jar-ce

haraʔ-konɖa

basa-har-ke

climb.down-SS vicinity-mountain reside-PL-T/A

‘From that side they abandoned the village of Janwai, took up all their ceremonial drums and pots, climbed down and settled in the vicinity of the mountain.’ (Mahapatra and Zide no date, F.6)

This has only scratched the surface of the complexity embodied in the demonstrative system of Gtaʔ. For more on Munda demonstrative systems, see any of the relevant sections in the chapters in this volume.

3.1.8 Numerals

The numerals from ‘one’ to ‘twenty’ in Gta? are

(94)	<i>m-mwiŋ/mmuiŋ</i>	<i>gomiŋ</i>
	<i>mbar/mbaya</i>	<i>gombar</i>
	<i>nji</i>	<i>gonji</i>
	<i>õ</i>	<i>go-õ</i>
	<i>malwal/malue</i>	<i>gomal</i>
	<i>tur</i>	<i>gotur/gu?tur</i>
	<i>gu</i>	<i>gogulgugu</i>
	<i>tma</i>	<i>gotma</i>
	<i>sontiŋ</i>	<i>gosontiŋ</i>
	<i>gua</i>	<i>gosal gallkuŋi</i> (Mahapatra <i>et al.</i> 1989)

Numerals take singular forms and may appear with classifiers in incorporated and/or derivational compound forms. The forms in (95) have the structure Numeral Noun.

- (95) (a) *qetwa qiiŋ-ce mbar hni-ne remwa-hiiŋ at wiŋ-ha?-har-ke*
 Like.that do-SS two village-GEN person-PL these quarrel-PL-T/A
 ‘Doing like that, the peoples of the two villages quarreled.’ (Mahapatra and Zide no date, F.16)
- (b) *ttar-ce mmwiŋ sla? alo we-ce e-rsiŋ-har-ke*
 REDPL:come.out-SS one tree under go-SS go-assemble-PL-T/A
 ‘They come out, go under a tree and (start to) assemble.’ (Mahapatra and Zide no date, A.43)

Note the animate form *mbaya* ‘2 persons’ (as well as plural verb agreement, see sections 3.1.1 and 3.2.1), which may be used without a following noun.

- (96) *e?-ke mba-ya hara+jita ne?-qiiŋ*
 today-CASE two-person defeat+win 1PL.INC-do
 ‘Let’s today decide by stake.’ (Mahapatra and Zide no date, II.6)

Other nouns require classifiers. These may be free-standing in form (97) or may combine with the numeral in a derivational compound (98).

- (97) *mæ-hiiŋ-ne mbar klig bũti læ?-ge*
 3PR-PL-GEN two CLSSFR buffalo COP-T/A
 ‘They had two buffaloes.’ (Mahapatra and Zide no date, F.2)
- (98) *qetwa? la?ŋig la?ŋig la?ŋig-ce mal-bha? a?-ŋiiŋ-ce*
 like.that night night night-SS five-head:CLSSFR CAUS-do-SS
tur-bha? sgwa bir=tia-ne sre-ne qun+tæ-ce
 six-head:CLSSFR like gayil-GEN/DEF deer-GEN/DEF carry-SS
qwa raŋ-wig-har-ke
 home bring-come.home-PL-T/A
 ‘Spending night after night like that they brought home five or six head of gayil and deer.’ (Mahapatra and Zide no date, A.59)

Duplication of numeral stems in Gta? serves the function of distributive numerals ‘X by X’.

- (99) *gmi? mmwiŋ mmwiŋ a?-dij-ce konḍa pwetur-æg-a?-ca?-ke*
 goat one one CAUS-DO-SS mountain surround-shit-CAUS-arrive-T/A
 ‘Goat one by one went around the mountain shitting.’ (Mahapatra and Zide
 no date, II.13)

3.1.9 Adpositions

A number of postpositional elements are found in Gta?. They appear with an uninflected (non-case marked) form of the noun they govern. Some examples of common postpositions in sentences in Gta? include the following.

- (100) (a) *ḍæk basoŋ-la tæŋ ŋgire bari wō-ḍæ*
 Like.this say-DS that young.man again mother-in-law-3
pla? wig-ke
 behind come.home-T/A
 ‘She spoke like this and that young man came back home behind (=following)
 his mother-in-law’ (Mahapatra and Zide no date, K.11)
- (b) *mæ?-swa-ne mæ-hiŋ læŋ-ḍia?-kig hli?-koṅḍa we-ce*
 One-day-DEF 3PR-PL rainy-season bamboo.shoot-mountain go-SS
gaŋ-hā sla? alo we-ge
 bamboo tree under go-T/A
 ‘One day in the rainy season going to Shoot Mountain they came under the
 bamboo trees.’ (Mahapatra and Zide no date, G.4)
- (c) *tæŋ k-n-ala e? sgwa oron ccoŋ a-mia?-har-ke*
 That time today like cooked.food REDPL:eat NEG-CUST-PL-T/A
e? sgwa paiṭi ḍiŋ a-mia?-har-ke
 today like work do NEG-CUST-PL-T/A
 ‘At that time unlike today, they did not eat cooked food nor did they work
 (the fields) like today.’ (Mahapatra and Zide no date, G.2)
- (d) *tela ataen hni-ne pujari at go=re bitre*
 then that.very village-GEN priest that hunt=person among
mmwiŋ-ja pujari læ?-ge tæŋ pujari mæ? pag=ra
 one-person priest remain-T/A that priest what offering
bi?-ke tæŋ koṅḍa-ne bri-huŋ-ke pag=ra bi?-ke
 give-T/A that mountain-GEN mountain-child-CASE offering give-T/A
 ‘Then the priest of that village who was also among the hunters, that priest
 made some offering to the mountain-god of that mountain.’ (Mahapatra and
 Zide no date, A.52)
- (e) *mæ-hiŋ at oriŋgi-hū ḍætte basa-læ?-ge*
 3PR-PL there Oringi-child with reside-AUX-T/A
 ‘They lived there with the people of Oringi.’ (Mahapatra and Zide no
 date, F.7)

3.1.10 Derivation

There are several different means of deriving nominal forms in Gta?. This includes pseudo-derivation or juxtaposition compounding of two free forms of nouns, and

derivational compounding, which consists of combining forms of nominals. One of the hallmarks of Gta? nominal structure (and verbal structure for that matter due to the system of incorporation) is the obligatory bimoraic full form [FF] of nouns that contrasts with a monosyllabic combining form [CF] that combines with other nominal or verbal stems. The combining form and the full form are often transparently related to each other, the combining form most commonly appearing to lack an initial consonant, perhaps a lexicalized prefix, that is found with the bimoraic full form. The two most common appear to be a syllabic nasal and *g-*, but reduplication of *C*₁ is also used. This is apparent when comparing forms of Gta? that are cognate with Kherwarian forms, where prefixes other than reduplication (which as in many other languages is a full reduplication in the case of ‘turmeric’) has been lost, in nouns replaced by suffixes, compounding or vowel extensions of various sorts.

(101) Selected Nouns in Gta? with Kherwarian cognates (i.e. which are Proto-Munda words)

Gta?	gloss	Kherwarian
<i>tti/nti</i>	‘hand’	<i>ti ~ tii</i>
<i>nco</i>	‘foot’	<i>jan̄ga</i>
<i>gcæŋ</i>	‘porcupine’	<i>ʃ^hīk jiki (H)</i>
<i>ŋgir</i>	‘young man’	---
<i>mmwa?</i>	‘eye’	<i>metlq [-e-, -ε?]</i>
<i>ŋku</i>	‘tiger’	<i>kul, kula (M)</i>
<i>ũhũl-ũ</i>	‘child’	<i>hɔn (o), hɔpɔn</i>
<i>gsu?</i>	‘dog’	<i>seta</i>
<i>nɖia?</i>	‘water’	<i>dak gl? Ø</i>
<i>ssia</i>	‘turmeric’	<i>sasaj</i>
<i>gsæŋ</i>	‘fowl’	<i>sim</i>
<i>ŋkwilŋ</i>	‘f-i-l, w.e.b.’	<i>hɔŋhar Ho honyar, Mundari honyar ~ hoŋear</i>
<i>gbe</i>	‘bear’	<i>bana</i>
<i>ncu</i>	‘oil’	<i>sunum</i>
<i>n(d)rwe</i>	‘fly’	<i>rɔ, roko</i>

There is variation in the derivational means used to form the free form from the combining form in Gta? in certain instances. For example, the forms for ‘hand’ above *tti* and *nti* show a root/stem *ti*, also the combining form, and what appears to be reduplication in the former case and syllabic nasal prefixation in the latter case. Note that the same derivational variation is seen in the forms found for ‘father’ in Gta? *mba* and *bba*; note also the vocative form ‘Dad[dy]!’ *aba*, each show a different form derived from =*ba*.

Many combining forms are used in Gta? (Mahapatra and Zide 1972). Some common ones are exemplified below. Such concepts as animals, for example, ‘pig’, ‘chicken’, ‘body parts’, ‘nose’, ‘mouth’, ‘day’, ‘man/person’ are typical combining forms.

(102) (a) *se-gsæŋ* *tiloŋ=po* *dur-gbug* *te?mur=bug*
 CLSFR:COLL-chicken spur=leg? CLSFR:COLL-pig snout=pig
c-m-oŋ-ũhũ *rupug-rupug*
 eat/AGENT/eat-child tra la la
 ‘spur of the chicken’s leg, snout of the pig, (you are) the child-eater tra la la.’ (Mahapatra and Zide no date, L.50)

- (b) *dukri basoŋ-ke* ‘*amu?* *bobo tbo?* *ke-la æg=bug*
 Old.woman say-T/A don’t child earth see-COND feces=pig
æg=sæŋ læ?-*ke ar-a?*-*lo?*-*ge na wæ?*=*go*
 feces=chicken remain-T/A NEG-CAUS-fall-PROHIB you cloth=fold
raŋ-jar-la-ce bi?-*la*
 bring-climb.down-IMP-SS give-IMP
 ‘The old woman said, “look, don’t let them fall, there are pig and chicken
 feces remaining (on the ground), climb down and bring some up in the fold of
 your loin cloth and give (them) to me.”’ (Mahapatra and Zide no date, L.11)
- (c) *gte-la a-næ?*-*ke gaŋi-si cili ccoŋ-ne mile-e*
 then OBJ-WE.INC-CASE more=day meat REDPL:eat-GEN available-FUT
 ‘Then we will have meat for eating for many days.’ (Mahapatra and Zide
 no date, D.12)
- (d) *ljo we-la mmwiŋ mæ?*-*bare baya ɖokra sgwa ke-ce*
 field go-DS one what-INTERJ mad old.man like look-SS
giŋ=sia? giŋ=sia? nluŋ-ce clæ clæ ugbo?-*ce-ka*
 cheek-monkey cheek-monkey ear-SS long long hair-SS-only
bari mo?mnaeg-ne buti bwa?=*tar-ce mmwiŋ unæŋ=u*
 again very.big-DEF basket carry=shoulder-SS one cigar
la+to?-*ce?*-*ka mæ big=cog ɖæte bbig-ɖiŋ-ke ljo*
 hold.in-lips-ss-only s/he sow=basket with REDPL:SOW-AUX-T/A land
 ‘She went to the field, an old man looking like a cheek monkey was there
 with long hair, carrying a basket on his shoulder, with a cigar in his mouth
 was sowing the field with a basket.’ (Mahapatra and Zide no date, I.8)

In the case of =*re* ‘person, man’ from *remwa*, it can attach to primarily nominal or verbal stems. In the latter case, either the basic form or the reduplicated form of the verb stem is used.

- (103) (a) *o?*=*sæ næ gta?*=*re mmwiŋ-ja læ?*-*ke*
 much.before we Gta?=people one-person remain-T/A
 ‘In the old days, there was one of us Gta? people.’ (Mahapatra and Zide no
 date, I.1)
- (b) *bba-ɾæ basoŋ-ke nɖu andi-andi-ne si-si=re-hiŋ at gali*
 father-3 say-T/A no COPY-where-GEN plough=man-PL there way
wiŋ-ɖiŋ-e ɖak-ce-ka bari basoŋ-la-ŋɖiŋ a-mane-ke
 go.home-AUX-FUT QUOT-only again say-DS-also NEG-agree-T/A
 ‘Her father disagreed, “no, ploughmen might be passing through on their
 way home.”’ (Mahapatra and Zide no date, I.11)
- (c) *te-la a-tæn hni-ne pujari at go=re bitre*
 This-DS that.very village-GEN priest that hunt=person among
mmwiŋ-ja pujari læ?-*ge tæŋ pujari mæ?* *pag=ra*
 one-person priest remain-T/A that priest what offering
bi?-*ke tæŋ koŋɖa-ne bri-huŋ-ke pag=ra bi?*-*ke*
 give-T/A that mountain-GEN mountain-child-CASE offering give-T/A
 ‘Then the priest of that village who was also among the hunters, that priest
 made some offering to the mountain-god of that mountain.’ (Mahapatra and
 Zide no date, A.52)

- (d) *eʔ-ʔe* *ggo=re-hij* *koŋɖa-baŋ* *wig-ɖij-la*
 now-time REDPL:hunt=person-PL mountain-ABL come.home-AUX-DS
sela-mbweʔ-hij *kero* *kisaʔlo* *gag-la* ***ggo=re-hij***
 female-PL way loin.cloth tie-DS REDPL:hunt=person-PL
hni *gga* *a-ia-har-ke*
 village REDPL:enter NEG-CAP-PL-T/A
 ‘Now [if] the hunters return home, and the women have blocked their access
 by tying loincloths together, the hunters can’t enter the village.’ (Mahapatra
 and Zide no date, A.63)

The other type of nominal formation is the juxtaposition compound. This is a concatenation of two free forms of nouns together.

- (104) *kmæ* *hʔij* *gweʔ-we-la* *mæ* *ljo+bri* *ceʔmwa* *poʔ-ʔiaʔ-miaʔ-ke*
 that.one afterwards die-AUX-DS he field+forest grass sprout-CUST-T/A
 ‘Only after he died, did grass sprout again on his fields.’ (Mahapatra and
 Zide no date, I.31)

One semi-productive means of deriving nominal forms from verbal or some nominal stems is the process of nasal infixation. Here *n* is inserted into the position immediately following the word-initial consonant. It, thus, becomes either part of a complex onset cluster or a syllabic nucleus in the case of CC- stems (i.e. ones already reduplicated, from which *-n-* forms may also be derived).

- (105) *pnoʔ* ‘spear’ *poʔ* ‘stab, poke’
cnoʔ ‘broom’ *coʔ* ‘sweep’
bnasaʔ ‘residence’ *basaʔ* ‘reside’
ɖnɖa ‘lid’ *ɖɖa* ‘close lid, cover pot’ (Mahapatra *et al.* 1989)
tæn ***k-n-ala*** *eʔ-sgwa* *oron* *cɔŋ* *a-miaʔ-har-ke*
 That time today like cooked.food REDPL: eat NEG-CUST-PL-T/A
 ‘At that time unlike today, they did not eat cooked food.’ <-- *kala* (Mahapatra
 and Zide no date, G.2)

Note that *-n-* infixated derived nominals may be derived from verbs that have incorporated nouns in them in Gtaʔ.

- (106) *bari* *gag-bo-k[e]-ne* *goʔæ=o-ke* *toʔ-biʔ-la* *tar-ce*
 Again tie-AUX-T/A-DEP boy=child-CASE open-AUX-DS come.out-SS
ɖukri-ne *huŋ-ɖæ-ne* *c-n-og=mu*
 old.woman-GEN child-3-GEN ornament/NOUN/=nose
c-n-og=mwaʔ *ttoʔ-biʔ-ce* *tæn* *goʔæ=o-ke*
 ornament/NOUN/=eye REDPL:open-AUX-SS that boy=child-CASE
aʔ-ga-ce *mæ* *kia* *tto-ɖij-ge*
 CAUS-enter-SS he paddy REDPL:pound-AUX-T/A
 ‘again she uncovered the confined boy [who] came out, took the old woman’s
 child’s nose ring and eye ornament ... then that boy started pounding paddy
 himself.’ (Mahapatra and Zide no date, L.33)

3.1.11 Adjectives

It is not clear that adjectives as a word class exist *per se* in Gta?, but some stems may appear modifying a nominal stem in an underived form and appear in a position to the left of the noun they modify. Such a formation is not overly frequent in the textual materials.

- (107) (a) *kmæ-hij* *qhã* *kuru=paka* *we-har-ke*
 This:person-PL small pit=side go-PL-T/A
 ‘These people went to the small-pit side.’ (Mahapatra and Zide no date, A.5)
- (b) *ntig-dæ* *minɖik* *ɖulæg* *kala* *mna?* *remwa* *hetæ-ke*
 Afterwards night sleep time elder man remember/think-T/A
næŋ-dig *cili* *liŋ-n-coŋ-la* *pa-ŋɖij* *hma?-ɽig* *sa?-n-coŋ-la*
 I-also meat lift-1-eat-DS CONTNG-AUX curry-ADD hold-1-eat-DS
pa-ŋɖij-ɽag-ce *hetæ-hetæ-ce* *ɖulæg-ge*
 CONTNG-AUX-state-SS COPY-remember/think-SS sleep-T/A
 ‘Afterwards at the time of night sleeping the elder man thought “I could have brought meat, I could have held and eaten curry, so thinking he fell asleep.”’ (Mahapatra and Zide no date, J.10)
- (c) *pe* *sapa* *remwa* *pe-ɖulæg-ke* *næŋ* *ko-n-læ?-ge*
 You.PL all person 2PL-sleep-T/A I sit-1-AUX-T/A
 ‘(while) you people all slept, I (alone) sat (awake all night).’ (Mahapatra and Zide no date, A.44)

A more common structure is a derivational compound where the adjective stem fuses with the combining form of the noun, as in *hɽæ?-riŋ* below, not *hɽæ? griŋ* ‘black cat’ or *coŋa-bir* for *coŋa bir-tia*.

- (108) (a) *gmi?* *basoŋ-ke* *mæ?* *ɖij-ge* *gia?=mwa?* *hɽæ?=riŋ* *ɽag-ke*
 goat say-T/A what do-T/A crow=eye black=cat state-T/A
 ‘Goat said “so what, you crow-eyed black cat.”’ (Mahapatra and Zide no date, II.4)
- (b) *coŋa=bir* *ne-bi?-la* *k-mæ-hij* *a-næ-ke* *mæ?* *mane-e* *be*
 Lame=gayil 1PL-give-DS DEM-3PR-PL OBJ-WE-CASE what obey-FUT Q
 ‘We will give them the lame gayil (but) will they obey us?’ (Mahapatra and Zide no date, A.35)

Partial reduplication can encode an augmented meaning to the adjectival semantics

- (109) *e-go-har-la* *mmwiŋ* *mo?-mnaæg* *ɖij-k[e]-ne* *bir=tia* *mmwiŋ* *tar-ke*
 go-hunt-PL-DS one very.big AUX-T/A-DEP gayil one come.out-T/A
 ‘They started hunting and one very large gayil appeared.’ (Mahapatra and Zide no date, A.54)

Adjectives may be derived from adverbials with the ‘definite’ or ‘genitive’ suffix *-ne*.

- (110) *wõ-dæ* *wa=huy* *we-ge;* *we-ce* *basoŋ-la* *ŋgire*
 Mother-in-law-3 call=child go-T/A go-SS say-DS young.man
basoŋ-ke *me=mwa-ne* *sle* *ɖiɖi* *me?=swa* *n-a?-lug-e*
 say-T/A one=year-GEN work how one=day 1-CAUS-finish-FUT
 ‘The mother-in-law went to call the son-in-law and speak (with him); the young man said “how can I do one year’s work in just one day?”’ (Mahapatra and Zide no date, K.8)

3.1.12 Adverb(ial)s

Adverbials are another category with heterogeneous formal properties, and, thus, defined functionally in Gta?. Most typically, temporal and manner adverbs cluster near the left clause periphery, in either initial position or in second position but they may appear throughout a clause (except in final position).

- (111) (a) *bba-ɾæ-ke basoŋ-ke dapɾe bba-ɾæ tar-ke*
 father-3-CASE say-T/A instantly father-3 go.out-T/A
 ‘As soon as he said this to his father, his father immediately left.’ (Mahapatra and Zide no date, I.14)
- (b) *gbe basoŋ-ke ‘na to kala-ha? ssæ paŋ-ce*
 bear say-T/A you surely daily before come-SS
cçoŋ na-mia? -ke næŋ kala-ha? hɾiŋ ppaŋ m-mia? -ke
 REDPL:eat 2-CUST-T/A I daily afterwards REDPL:come 1-CUST-T/A
 ‘Bear says “every day you come before and eat and then I come.”’ (Mahapatra and Zide no date, II.5)
- (c) *kmæ hɾiŋ gwe? -we-la mæ ljo+bri ce?mwa po? -ɾia? -mia? -ke*
 that.one afterwards die-AUX-DS he field+forest grass sprout-CUST-T/A
 ‘Only after he died, did grass sprout again on his fields.’
- (d) *ssæ dæk djiŋ-ce a-næ? remwa-hiŋ-ke sapa bagwe?-çoŋ-har-ke*
 Before like.this do-SS OBJ-WE.INC person-PL-CASE all kill-eat-PL-T/A
 ‘Doing so earlier, they killed and ate up all of our people.’ (Mahapatra and Zide no date, D.28)

Some adverbials constitute partially grammaticalized elements in Gta? such as *hnor*, an adverbial expressing concomitant action.

- (112) *ɖukri p-n-a? =cwe? =ɾia? ɖokra-ke pno-mæg-nnia a? -ug-ce*
 Old.woman millet-gruel oldman-CASE anus-near CAUS-drink-SS
a? -ro+a? -be? -ce osmar hnor a? -ro-ce ho? -ke
 CAUS-compose+CAUS-ECHO-SS song CONCOMITANT CAUS-compose-SS cry-T/A
 ‘The old woman smearing millet gruel near the oldman’s anus composing songs, cried while singing.’ (Mahapatra and Zide no date, D.14)

3.2 Verbal morphology

3.2.1 Subject

Gta? is like Remo, Gutob, and Kharia, (and increasingly Juang) in generally showing only subject agreement with the verb. A sample paradigm is given in (113).

- (113)
- | | | |
|---------|--------------------------|-------------------------|
| 1SG | <i>n-çoŋ-ke</i> | ‘I ate’ |
| 2SG | <i>na-çoŋ-ke</i> | ‘you ate’ |
| 1PL.INC | <i>ni-çoŋ-ke</i> | ‘we (include. you) ate’ |
| 1PL.EX | <i>næ-çoŋ-ke (~ nɛ-)</i> | ‘we (not you) ate’ |
| 2PL | <i>pa-çoŋ-ke</i> | ‘y’all ate’ |
| 2DL | <i>pe-çoŋ-ke</i> | ‘you two ate’ |
| 3SG | <i>çoŋ-ke</i> | ‘she/he ate’ |
| 3PL | <i>çoŋ-ke</i> | ‘they ate’ |
| | <i>çoŋ-har-ke</i> | |

Note *ni?*- first dual imperative

As is evident, unlike Remo, Gutob, and Kharia that have subject enclitics, Gta?, like Juang (and Gorum) has subject prefixes. This is an archaic feature shared by archaic (South) Munda languages (Anderson 2007). The only marker from a morphophonological perspective that merits particular attention is that for the first singular subject prefix, which is a syllabic nasal that is often homo-organic to the initial sound of the verb stem (e.g. *m* before labials, *ŋ* before velars; note also its realization as *ũ* before *we-* (also *n-*) but *ŋ* before *æ* (also *n-*)).

- (114) (a) *dukri gewa=mwaʔ-riaʔ-ce basoŋ-ke ŋɖu ɖokra mari*
 old.woman shy=eye-echo-ss say-T/A no old.man again
oʔna-ŋig ũ-we-e
 when-also 1-go-FUT
 ‘The old woman was embarrassed and said “No, old man, I will go there again sometime.”’ (Mahapatra and Zide no date, L.26)
- (b) *mæ we-ce basoŋ-ke ʔna mæʔ bbig na-ɖiŋ-ke ɖakce*
 he go-SS say-T/A ‘you what REDPL:SOW 2-AUX-T/A QUOT
saliaʔ+ku-ke
 ask+ask-T/A
 ‘He (=father) goes and says “what are you sowing?” he asked.’ (Mahapatra and Zide no date, I.16)
- (c) *ɖæt neʔ-ɖiŋ sina pe pe-ŋ-ɖo-la nor+nara-hiŋ*
 That.like 1PL.INC-do though you.PL 2PL-AUGM-run-DS human.being+ECHO-PL
ggweʔ-we-har-e ɖakce kæn-ne ɖira basoŋ-ke
 REDPL:die-AUX-PL-FUT QUOT this-DEF/GEN millet.species say-T/A
 ‘“Let’s do it like this (as an acceptable alternative), you run away, the human beings will all die,” this millet said.’ (Mahapatra and Zide no date, C.16)

In complex verb stem and serial verb constructions, in Gta? the referent property marker is usually affixed to the last member in the series/complex.

- (115) *wiŋ-haʔ-har-ke ho-m-m-og ho-t-m-u*
 Quarrel-PL-T/A RECIP-beat/RECIP/ RECIP-throw.stone/RECIP/
ho-s-m-iʔ+ho-s-m-aʔ-har-ke
 RECIP-cut/RECIP/+RECIP-catch/RECIP-PL-T/A
 ‘They beat each other, threw stones at each other, caught and butchered each other.’ (Mahapatra and Zide no date, F.17)

Compare the distribution of subject marking of the following two sentences which seems to be determined by the choice of lexical verb in the construction (two very important deictically active verb stems), not the auxiliary which remains functionally identical across the two.

- (116) (a) *na kala-haʔ andʒi we-na-miaʔ-ke iaŋ*
 you daily-EMPH where go-2-CUST-T/A mother.VOC
 ‘Mother, where do you go everyday?’ (Mahapatra and Zide no date, B.13)
- (b) *ɖia-ŋɖæ basoŋ-ke næŋ ɖwa ũ-wig-miaʔ-ke*
 mother-3 say-T/A I home 1-go.home-CUST-T/A
 ‘the mother said “I go home.”’ (Mahapatra and Zide no date, B.14)

Under, as yet, unclear circumstances, agreement may be suppressed in certain uses of particular verbs, for example, the final verbs in (117) below, both auxiliary verb structures that usually have subject inflection but here lack them (and both have reduplicated lexical verb components).

- (117) (a) *te-la tæn kitoŋ basoŋ-ke næŋ-ne paiŋi kæn-haʔ, na*
 then that god say-T/A I-GEN work this-EMPH you
dji-rokom kia na-big-ke næŋ dʒig se-rokom ceʔmwa
 what-manner paddy 2-SOW-T/A I also that-manner grass
bbig-dij-ke
 REDPL:SOW-AUX-T/A
 ‘Then that god said “my task is this: whatever manner you sow in, I will sow in that manner too.”’ (Mahapatra and Zide no date, I.17)
- (b) *dæk dij-ce iʔtuŋa bole aʔ-dij-ce mbar pwaʔ-coŋ*
 like.this do-ss little cooked.rice CAUS-do-ss two leaf.cup
nturiaʔ aʔ-dij-ce næ gtaʔ=re-hij ccoŋ+uug-miaʔ-ke
 gruel CAUS-do-ss we Gtaʔ-person-PL REDPL:eat-REDPL:drink-CUST-T/A
 ‘And thus we did it this way, we Gtaʔ people used to make two leaf-cups and eat [only] a little cooked rice with gruel.’ (Mahapatra and Zide no date, C.23)

Subject marking is the same for intransitive verb stems (-*we*-) and transitive ones (-*twij*-), that is, Gtaʔ exhibits nominative/accusative alignment.

- (118) *hni-[n]-[d]re basoŋ-ke næ ne-we-e tæn bir=tia-ke ne-twij-e*
 village-GEN-person say-T/A we 1PL-go-FUT that gayil-CASE 1PL-shoot-FUT
 ‘The villagers said, “we will go, we will shoot down those wild cattle.”’
 Mahapatra and Zide no date, A.4)

Incorporated stems (-*pag=ra*) inflect the same as unincorporated ones (-*dij*) or functional elements (-*ke*).

- (119) *tela eʔ næ dʒi dʒi ne-dij kitoŋ-ke mæʔ-ŋig ne-pag=ra*
 then now we how 1PL-do god-CASE what-ADD 1PL-break=twig
pag=ra-ce-ka bari tæn koŋda e-go ne-ke dʒakce basoŋ-ke
 break=twig-ss-only again that mountain go-hunt 1PL-AUX QUOT say-T/A
 ‘Then what are we to do now? Let’s make an offering of something to the god and then let’s try to go hunting on that mountain.’ (Mahapatra and Zide no date, A.51)

Usually only finite verbs mark subject but different subject constructions (see section 4.2.3) with personal subjects may take subject marking as well.

- (120) (a) *coʔa=bir ne-biʔ-la k-mæ-hij a-næ-ke mæʔ mane-e be*
 lame=gayil 1PL-give-DS DEM-3PR-PL OBJ-we-CASE what obey-FUT Q
 ‘We will give them the lame gayil (but) will they obey us about anything?’
 (Mahapatra and Zide no date, A.35)
- (b) *na kia aʔ-ble na-coŋ-la næŋ ceʔmwa n-a-raʔ-ble*
 you paddy CAUS-ripen 2-eat-COND I grass 1-NEG-CAUS-ripen

ɖakce kmæ basoŋ-ke
 QUOT this.one say-T/A
 ‘“You cultivate paddy and eat it; shouldn’t I cultivate grass,” this one said.’
 (Mahapatra and Zide no date, I.18)

As aforementioned, in serial verb constructions in Gta?, the referent property marker is usually prefixed to the last member in the series/complex. Note that in asyndetic paratactic complex predicate formations, doubled subject marking is rather found.

- (121) (a) *næŋ bagweʔ+joŋ n-taliŋ-e*
 I kill-pick.up 1-ECHO-FUT
 ‘I will kill (you) and throw (you) away.’ (Mahapatra and Zide, n.d.)
- (b) *aʔkuŋeʔ n-dŋŋ-ke*
 CAUS:meet 1-PROG-T/A.I
 ‘I am causing to meet’ = ‘fixing you up with a girl.’ (Mahapatra and Zide, n.d.)
- (c) *næŋ ɖuŋɖi e-ko n-læʔ-e*
 I garden go-sit-1-stay-FUT
 ‘I will go and sit and rest in the garden.’ (N. Zide 1997:332)
- (d) *næŋ hweʔ-ɖulæŋ-n-læʔ-e na a-næŋ-ke pnomæŋ*
 I pretend-sleep-1-AUX-FUT you OBJ-I-CASE anus
p-n-aʔ-cweʔ-ɽiaʔ kiŋ-biʔ
 millet.gruel pour.down-AUX
 ‘I’ll pretend to sleep, you smear some millet gruel around my anus.’
 (Mahapatra and Zide no date, D.8)

A small number of stems are discontinuous in Gta? and function like old serial structures of the above mentioned type, with agreement on the rightmost host. One common stem is ‘run away’ *lag -ɖo*.

- (122) *ã=say-heʔ bæ-ke ã=say+ã=ɽiaʔ-ce bini*
 thatch=house-EMPH send-T/A thatch=house+ECHO/TAG-SS separate
aʔ-nsiŋ+aʔ-lwe coŋ-pe ɖɖag ɖŋŋ-ge atæn-ɖiŋ
 CAUS-boil+CAUS-ECHO eat-2PL REDPL: state AUX-T/A that-also
na lag-na-ɖo-ke ɖba
 you run.away 2-run.away-T/A come/go.IMP
 ‘He asked you to thatch the house; he was telling you “make a thatch house ready (for you) to prepare meals and eat separately and for that reason you ran away,” now let’s go.’ (Mahapatra and Zide no date, K.10)

In a few instances, however, there are exceptions to this person marking pattern, and there are forms with the prefix at the beginning of the verb complex in Gta?.

- (123) *n-wiŋ-miaʔ-ke n-weʔ-gag-ce*
 1-go[.back]-CUST-T/A 1-swing-tie-ss
 ‘I habitually went back.’ ‘after I swung and tied’ (Zide 1997:333)

In some instances, verbal agreement prefixes may even be found on both members of a juxtaposed coordinate construction, or a mixed construction with double subject marking but tense/aspect only on the rightmost conjunct, reminiscent of so-called split/doubled inflectional structures discussed by Anderson (2006).

- (124) *qokra, na bole-nturia? na-coŋ na-ug-k(e)-ŋe*
 old.man you rice-millet.gruel 2-eat 2-drink-T/A-Q:RHET
 ‘Old man, didn’t you drink/eat your rice millet and gruel?’ (Mahapatra and Zide no date)

Apart from *-har-*, there is one other instance of suffixal (or enclitic really) subject marking in the *Gta?* verbal system and that is in the formation of the imperative and prohibitive (see section 3.2.9), where dual and plural (second person) markers are found at the end of the verbal complex, as it is so found in other Munda languages (Anderson 2007).

- (125) (a) *pe-mba wig-ce a-pe-ke salia?+ku-la*
 2PL-father come.home-SS OBJ-2PL-CASE ask+TAG-COND/DS
ni-iaŋ wig-læ?-ge dakce a-basoŋ-ge-pa
 1PL-mother:VOC come.home-AUX-T/A QUOT NEG-say-PROHIB-2DL
 ‘when your father comes back home and asks you, do not say “our mother has come home.”’ (Mahapatra and Zide no date, B.9)
- (b) *gte-la k-mæ basoŋ-ke pe a-ŋo-ge-pe ge?-la-pe*
 this-DS DEM-3PR say-T/A YOU.PL NEG-flee-PROHIB-2PL come-IMP-2PL
qag-la ŋdŋig ja-ŋig a-we-ke
 state-DS also who-ADD NEG-go-T/A
 ‘Then he said “don’t flee, please come,” but nobody came/went.’ (Mahapatra and Zide no date, J.23)

Although overtly or semantically dual, nominals can have only plural or singular agreement on the verb, not the dual marker *-pa* which in verbal forms is, therefore, restricted to its original second dual meaning.

- (126) [Noun]_{DL}.....[Verb]_{PL/SG}
- (a) *mæ-pa mba-ya hara-hari dŋiŋ-ce kōda pwetur-æg-har-ke*
 he-DL two-person defeat:ECHO AUX-SS mountain surround-shit-PL-T/A
 ‘the two of them trying to beat each other, went around the mountain shitting.’ (Mahapatra and Zide no date, II.12)
- (b) *qokra-qokri-pa sapa jiwa-ke bagwe?-har-ke*
 Old.man-old.woman-DL all animal-CASE kill-PL-T/A
 ‘the old man and old woman killed all the animals.’ (Mahapatra and Zide no date, D.24)
- (c) *griŋ-ŋku-pa ho-ba?lir-ke*
 cat-tiger-DL RECIP-talk-T/A
 ‘The cat and tiger talked to each other.’ (Mahapatra and Zide no date, H.5)

3.2.2 Object types

Object marking is, generally, not part of the *Gta?* verbal structure with one exception: The verbal plural marker *-har-* can refer to objects in certain of its uses. These are the only instances of bipersonal verb structure in *Gta?* but it follows the older Munda order of S-V-O in the make up of the verb-word.

- (127) *gte-la næŋ mria?-ce a-mæ-hiŋ-ke m-bagwe?-har-e*
 then I rise-SS OBJ-3PR-PL-CASE 1-kill-PL:OBJ-FUT
 ‘Then I will get up and kill them all.’ (Mahapatra and Zide no date, D.11)

3.2.3 *Tense*

Plains Gta? has two tense/aspect markers, *-ge* and *-ke*, both with several complicated uses, sometimes quite restricted. *-ge* may be cognate with Proto-Gutob Remo **-gi* (and Kharia *-ke*) while *-ke* is an imperfective with a complicated tense/aspect range of meanings. The Hill Gta? form corresponding to *-ke* is *-ti*. Gta? *-kel-ti* probably reflects not just something of the South Munda N_{PS}T but also the ancestor of the more obscure *-ki*, the only other reflex of which is in Hill Remo. Mahapatra *et al.* (1989) lists a large number of differences between the use of *-ke* and *-ge* in Gta?. These were summarized by Zide (1999) and are in (128).

(128)	<i>-ke</i>	<i>-ge</i>
(a)	indefinite/aorist	definite
(b)	cislocative	translocative
(c)	assertive	evidential
(d)	state/present	change/past
(e)	w/auxiliary=present/perfect	w/auxiliary=past/pluperfect
(f)	habitual/general present	definite past (w/Intransitives)
(g)	interrogative	Answer =assertive?
(h)	negative	prohibitive
(i)	participial	----
(j)	remote past	recent past
(k)	actant focus	action focus
	neutral	marked

These are mostly consistent within a single column, but there are some anomalies, for example, the use of *-ke* as a supposed assertive in declarative sentences, corresponding to a quasi-evidential use of *-ge* in contrast with the interrogative use of *-ke* and the assertive answer in *-ge*. However, such cross-categorical inconsistencies are found in the use of particular forms/constructions in a wide range of languages; for example, in the Turkic language Xakas the auxiliary verb *odir* 'sit' can be used as either a marker of imperfective or perfective action (Anderson 2004a). The tense/aspect element *-ke* is the default past marker, and the one possible in negative past forms (except in certain auxiliary constructions).

Examples of the use of *-ke* and *-ge* in Gta? include those listed in (129), (130) and (131)

- (129) (a) *mæ cili coŋ-ke* vs. (b) *mæ cili coŋ-ge*
 she/he meat eat-T/A she/he meat eat-T/A
 'She/he eats/ ate meat.' 'She/he's just eaten meat.'
- (c) *sari-ge*
 flower-T/A
 'flowered completely'

- (130) *-ke*
- (a) *oʔ=sæ næ gtaʔ=re mmwiŋ-ja læʔ-ke*
 much.before we Gtaʔ=people one-person remain-T/A
 'In the old days, there was one of us Gta? people.' (Mahapatra and Zide no date, I.1)
- (b) *gbe æg-aʔ-caʔ a-ia-ke*
 bear shit-CAUS-arrive NEG-CAP-T/A
 'Bear could not shit (around all of it).' (Mahapatra and Zide no date, II.14)

- (c) *ḍia-ṇḍæ basoṅ-ke næṅ ḍwa ũ-wig-mia?-ke*
 Mother-3 say-T/A I home 1-go.home-CUST-T/A
 ‘The mother said “I go home.”’ (Mahapatra and Zide no date, B.14)

-ge is used in perfect constructions, as the translocative with motion verbs (see section 3.2.6), as the default past with several verb stems (*a?lug-* ‘finish’, *ṇḍia? kuma-* ‘give birth’), and in prohibitive forms.

(131) *-ge*

- (a) *mmwiṅ goṛæ=o basoṅ-ke o?na-baṅ a?si-ge*
 One boy say-T/A when-ABL fever-T/A
 ‘One boy asked “since when has he been sick?”’ (Mahapatra and Zide no date, M.12)
- (b) *ggæ?=se?+bog-pa?=se? a?-lug-ge*
 REDPL:worship=spirit+ECHO/TAG=spirit CAUS-finish-T/A
 ‘He finished the death rites’ (Mahapatra and Zide no date, B.6)
- (c) *sapa jiwa-hiṅ ke-re ḍag-ce gga-har-ge*
 all animal-PL see-EXCLAM state-SS REDPL:enter-PL-T/A
 ‘“Let’s see” they said and all the animals entered.’ (Mahapatra and Zide no date, D.20)
- (d) *mæ-hiṅ at oriṅgi-hũ ḍætte basa-læ?-ge*
 3PR-PL there oringi-child with reside-AUX-T/A
 ‘They lived there along with the people of Oringi.’ (Mahapatra and Zide no date, F.7)
- (e) *mæ-ne knwe?-ṛæ ṇḍia? kuma-ge*
 3PR-GEN wife-3 water bath-T/A
 ‘His wife gave birth.’ (Mahapatra and Zide no date, B.2)

The Gta? future in *-e* may be cognate with the Gutob present participle *+el*, or possibly other anomalous futures in *-e* found in various Munda languages (Anderson 2007). In Proto-Gta?, as in Proto-Sora-Gorum, the transitive/intransitive inflectional dichotomy is not active in the tense/aspect system. Only the later South Munda ‘intransitive’ past (Anderson 2001, 2004b, 2007) was preserved as the past in **-ki* (cf. pre-Kharia *(*)-ki* and Proto-Gutob-Remo **-gi*). The subject proclitics were preserved in Proto-Gta?, however (and modern Gta? as well).

- (132) (a) *gmi? basoṅ-ke ḍæt ṇḍu kæn kḍḍa-ke ja*
 goat say-T/A that.like no this mountain-CASE who
pwetur-æḡ-a?-ca?-la mæ ssæ paṅ-coṅ-e
 surround-shit-CAUS-arrive-DS he before come-eat-FUT
 ‘Goat said “No, not like that, (I suggest) whoever can surround the mountain in shit, let him come and eat first.”’ (Mahapatra and Zide no date, II.11)
- (b) *hni-[n]-[ḍ]re basoṅ-ke næ ne-we-e tæṅ bir=tia-ke*
 village-GEN-person say-T/A we 1PL-go-FUT that gayil-CASE
ne-twiṅ-e
 1PL-shoot-FUT
 ‘The villagers said, “we will go, we will shoot down that gayil.”’ (Mahapatra and Zide no date, A.4)
- a-næ?-ke kala-ha? ḍekwa gæ?-mu gæ?=moa?-læ?-e*
 OBJ-WE.INC-CASE daily like.this fry=nose fry=eye-AUX-FUT
 ‘They shall be frying our noses and eyes up like this every day.’ (Mahapatra and Zide no date, C.5)

Note that in Gta?, the future also has a zero allomorph in the negative. There is also no tense/aspect marker in negative present formations (see also section 3.2.9).

- (133) (a) *mæʔ-sar n-a-biḡ ɖakce mæ basoŋ-ke*
 what-PURP 1-NEG-SOW QUOT he say-T/A
 ‘‘Why shouldn’t I sow?’’ He said.’ (Mahapatra and Zide no date, I.20)
- (b) *k-mæ-hij ke-la olæŋ-læʔ-k[e]-ne remwa-hij a-næ-ke*
 DEM-3PR-PL see-DS wander-AUX-T/A-DEP person-PL OBJ-WE-CASE
k-mæ-hij a-miaʔ
 DEM-3PR-PL NEG-know
 ‘They, (you see), are nomads, and they do not know us.’ (Mahapatra and Zide no date, A.36)

3.2.4 Aspect

A number of aspectual categories are found in Gta? grammar, mainly expressed by (fused) auxiliary verb constructions. These include perfect and progressive among others.

Perfect in Gta? is mostly marked by the auxiliary Gta? *ɖij-* which generally takes a reduplicated stem allomorph. In the negative past, no reduplication of the verb is permitted, even when otherwise dictated as such by *-ɖij-*.

- (134) (a) *c-coŋ-(n)-ɖij-ge* (b) *c-coŋ-ɖij-ge*
 REDPL-eat-1-PRF-T/A REDPL-eat-PRF-T/A
 ‘I have eaten.’ ‘She/he has eaten.’
- (c) *c-coŋ-n-ɖij-e* (d) *ccoŋ-ɖij-e*
 REDPL-eat-1- PRF-FUT REDPL-eat-PRF-FUT
 ‘I will eat.’ ‘She/he will eat.’
- (e) *c-coŋ-n-a-ɖij* (f) *c-coŋ-a-ɖij*
 REDPL-eat-1-NEG-PRF REDPL-eat-NEG-PRF
 ‘I won’t eat.’ ‘She/he won’t eat.’
- (g) *coŋ-n-a-ɖij-ge* (h) *coŋ-a-ɖij-ge*
 eat-1-NEG-PRF-T/A eat-NEG-PRF-T/A
 ‘I haven’t eaten.’ ‘She/he hasn’t eaten.’ (Mahapatra *et al.* 1989)

Note the use of *-ge* with *ɖij*, rather than the expected *-ke* in the negative conjugation (see section 3.2.9).

The Gta? perfect auxiliary has cognates in Gutob–Remo as well, viz. Remo *ɖen/m*, Gutob *ɖeŋ*. In these languages, the semantics of this auxiliary are not the same as in Gta?; In Remo, this is progressive (Anderson and Harrison this volume) while in Gutob this can function as a ‘light verb’ or a predicate meaning ‘become’.

- (135) Gutob
may-nen rone-bone ɖeŋ-gu buron-gu-nen-aʔso-gu-nen
 3-PL happy-ECHO AUX-PST.II live-PST.II-PL ECHO-PST.II-PL
 ‘They became happy and lived (on that way).’ (Zide 1997)

The Gta? progressive has *læʔ* for intransitive and *bo-* for transitive stems, including all causative marked verbs. This is one of the only domains in which the older conjugational distinction (whatever the original semantics distinguishing the two

classes may have been) that characterizes most modern Munda languages has been preserved in Gta? grammar. Note also the contrastive use of the tense markers *-ke* (with *bo*) and *-ge* (with *læ?*) as well.

- (136) (a) *coŋ-n-læ?-ge*
eat-1-PROG.I-T/A
'I was eating.'
- (b) *coŋ-læ?-ge*
eat-PROG.I-T/A
'She/he was eating.'
- (c) *a? -coŋ-m-bo-ke*
CAUS-eat-1-PROG.II-T/A
'I was feeding.'
- (d) *a?-coŋ-bo-ke*
CAUS-eat-PROG.II-T/A
'She/he was feeding.'
- (e) *coŋ-n-læ?-e*
eat-1-PROG.I-FUT
'I will be eating.'
- (f) *coŋ-læ?-e*
eat-PROG.I-FUT
'She/he will be eating.'
- (g) *a?-coŋ-m-bo-e*
CAUS-eat-1-PROG.II-FUT
'I will be feeding.'
- (h) *a?-coŋ-bo-e*
CAUS-eat-PROG.II-FUT
'She/he will be feeding.'
- (i) *at[e] gmi? copa=u ccoŋ dŋiŋ-g[e]=dŋak-ke*
there goat skin=mango REDPL:eat AUX-T/A=state-T/A
'it is said that Goat was there eating mango skins.' (Mahapatra and Zide no date, II.2)
- (j) *tte-la hīsa nŋu-kuŋa ja par+le mæ paŋ*
that-DS envy NEG.COP-NEG.COND who can+OPT he come
ccoŋ dŋiŋ-le
REDPL:eat AUX-OPT
'Thus there will be no envy; whoever wins, let him come and eat.' (Mahapatra and Zide no date, II.7)

3.2.5 Mood

There are a range of morphologically marked modal formations in Gta?. One set of these is found in the imperative/hortative paradigm and prohibitive forms (marked by *a[r]-. . .-ge*).

- (137) (a) *pa coŋ-pa*
you two eat-2DL
'Eat you-2!'
- (b) *pa a-coŋ-ge-pa*
you.two NEG-eat-PROHIB-2DL
'Don't eat you-2!'
- (c) *pa a?coŋ-pa*
you.two feed-2DL
'Feed you all!'
- (d) *pa ar-a?coŋ-ge-pa*
you.two NEG-feed-PROHIB-2DL
'Don't feed you all!'
- (e) *coŋ-pe*
you.PL-eat-2PL
'eat (you please)!'
- (f) *pe a-coŋ-ge-pe*
you.PL NEG-eat-PROHIB-2PL
'Don't eat (you please)!'
- (g) *pe a?coŋ-pa*
you.PL feed-2PL
'Feed (you two)!'
- (h) *pe ar-a?coŋ-ge-pe*
you.PL NEG-feed-PROHIB-2PL
'Don't feed (you 2)!'

Some more examples of these modal forms are offered below.

- (138) (a) *mbaya akwa ni[?] -læ? na dŋwa a-wiŋ-ge*
Both here 1DL[IMP]-remain you home NEG-go.home-PROHIB
'Let's both stay here, don't go home.' (Mahapatra and Zide no date, H.7)

- (b) *gte-la k-mæ basoŋ-ke pe a-dŋo-ge-pe geʔ-la-pe*
 then DEM-3PR say-T/A you.PL NEG-flee-PROHIB-2PL come-IMP-2PL
ɖag-la ŋɖig ja-rig a-we-ke
 state-DS also who-ADD NEG-go-T/A
 ‘Then he said “don’t flee, please come,” but nobody came/went.’ (Mahapatra and Zide no date, J.23)

Rhetorical forms can be marked by *-re/-re*:

- (139) *næŋ slweʔ-nsa konta ũ-we-re cili+haʔro-nsa ũ-we-e*
 I belly-PURP mountain 1-go-Q:RHET meat+fish-PURP 1-AUX/go-FUT
 ‘Am I to go to the mountain work in order to fill our bellies or in order to eat fish and meat?!’ (Mahapatra and Zide no date, M.17)

Conditional constructions come in several forms in Gtaʔ. The conditional subordinate clause formant (and different subject marker) *-la* is one such element found in formations of this type.

- (140) (a) *na kia aʔ-ble na-coŋ-la næŋ ceʔmwa n-ar-aʔ-ble*
 you paddy CAUS-ripen 2-eat-COND I grass 1-NEG-CAUS-ripen
ɖakce kmæ basoŋ-ke
 QUOT this.one say-T/A
 ‘“You cultivate paddy and eat it; shouldn’t I cultivate grass” this one said.’ (Mahapatra and Zide no date, I.18)
- (b) *dæk eʔ-baŋ næŋ ljo ceʔmwa na-big-la a-na-ke bol*
 look today-ABL I field grass 2-NEG-SOW OBJ-YOU-CASE good
n-a-dij ɖæt ɖakce-ka bari kitoŋ-ke basoŋ-ke
 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god “look, you sow grass in my field (again), then I will not be nice to you.”’ (Mahapatra and Zide no date, I.28)
- (c) *pe-mba wig-ce a-pe-ke saliaʔ+ku-la*
 2PL-father come.home-SS OBJ-2PL-CASE ask+TAG-COND/DS
ni-iaŋ wig-læʔ-ge ɖakce a-basoŋ-ge-pa
 1PL-mother:VOC come.home-AUX-T/A QUOT NEG-say-PROHIB-2DL
 ‘When your father comes back home and asks you, do not say “our mother has come home.”’ (Mahapatra and Zide no date, B.9)

The imperative occurs with second person, the subjunctive with the third and, rarely, first persons. The imperative marker, also *-la*, is used

- (141) (a) *ɖæt neʔ-dij sina pe pe-ŋ-dŋo-la nor+nara-hij*
 That.like 1PL.INC-do though you.PL 2PL/IMP-run-DS-AUGM human.
 being+ECHO-PL
ggweʔ-we-har-e ɖakce kæn-ne ɖira basoŋ-ke
 REDPL:die-AUX-PL-FUT QUOT this-DEF millet.species say-T/A
 ‘“let’s do it like this (as an acceptable alternative), you run away, the human beings will all die,” this millet said.’ (Mahapatra and Zide no date, C.16)
- (b) *ɖukri basoŋ-ke “amuʔ bobo tboʔ ke-la æg=bug*
 Old.woman say-T/A don’t child earth look-IMP feces=pig
æg=sæŋ læʔ-ke ar-aʔ-loʔ-ge na wæʔ=go
 feces=chicken remain-T/A NEG-CAUS-fall-PROHIB you cloth=fold

raŋ-jar-la-ce *bi?-la*
bring-climb.down-IMP-SS **give-IMP**

‘the old woman said, “look, don’t let them fall, there are pig and chicken feces remaining (on the ground), climb down and bring some up in the fold of your loin cloth and give (them) to me.”’ (Mahapatra and Zide no date, L.11)

- (c) *knwe?-ɾæ-hiŋ* *ɖiŋ* *basoŋ-ke* *we-la-pe* *bir=tia* *tar-læ?-ke*
 wife-3-PL also say-T/A **go-IMP-2PL** gayil come.out-AUX-T/A
 ‘The[ir] wives also said “go, the gayil has come out.”’ (Mahapatra and Zide no date, A.3)

Gta? also makes use of a small number of suppletive imperative verb stems. For example, *ɖba* ‘let’s go!’

The hortative/optative in Gta? is *-le[ʔ]*

- (142) (a) *tte-la* *hīsa* *nɖu-kuɕa* *ja* *par-le* *mæ* *paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come
cɔŋ *ɖiŋ-le*
 REDPL:eat AUX-OPT
 ‘Thus, there will be no envy; whoever wins, let him come and eat.’
 (Mahapatra and Zide no date, II.7)
- (b) *gbe* *basoŋ-ke* *akæŋ* *konɖa* *ja* *olæŋ* *a?-ca?-la*
 Bear say-T/A this mountain who walk CAUS-arrive.at.destination-DS
mæ *ssæ* *uli-alo* *tʰwã-læ?-le?*
 he before mango-under stand-remain-OPT
 ‘Bear says, “whoever gets to this here mountain first, let him stand under the mango tree.”’ (Mahapatra and Zide no date, II.9)
- (c) *nɖu* *mæ-hiŋ* *mia?-le?* *a-mia?-le?*
 Well 3PR-PL know-IMP NEG-know-IMP
 ‘Well, whether they know or not’ (Mahapatra and Zide no date, A.37)

Gta? also has a contingent mood prefix in *pa-*. There do not appear to be any obvious cognate elements in any other Munda language, but the fact that it is a prefix, suggests that it may well be old. One possibility would be the dubitative particle in Sora (Anderson and Harrison, this volume-b).

- (143) *ntiŋ-ɖe* *miŋɖik* *ɖulæŋ* *kala* *mna?* *remwa* *hetæ-ke*
 Afterwards night sleep time elder man remember/think-T/A
næŋ-ɖiŋ *cili* *liŋ-n-coŋ-la* *pa-ŋɖiŋ* *hma?-ɾiŋ* *sa?-n-coŋ-la*
 I-also meat lift-1-eat-DS CONTNG-AUX curry-ADD hold-1-eat-DS
pa-ŋɖiŋ-ɾag-ce *hetæ-hetæ-ce* *ɖulæŋ-ge*
 CONTNG-AUX-state-SS COPY-remember/think-SS sleep-T/A
 ‘Afterwards at the time of night sleeping the elder man thought “I could have brought meat, I could have held and eaten curry, so thinking he fell asleep.”’
 (Mahapatra and Zide no date, J.10)

3.2.6 Orientation/directionality

As described in section 3.2.3, the only basic means of encoding these categories in Gta? with our current state of understanding of its grammar is through the use of *-ke* vs. *-ge* with the stem *we-* to mark cislocative and translocative, respectively.

- (144) (a) *meʔ-swa-ne mæ-hij læŋ-ɖiaʔ-kig hliʔ-konɖa*
 One-day-DEF 3PR-PL rainy-season bamboo.shoot-mountain
we-ce gaŋ-hã slaʔ alo we-ge
 go-ss bamboo tree under go-T/A
 ‘One day in the rainy season going to Shoot Mountain they came under
 the bamboo trees.’ (Mahapatra and Zide no date, G.4)

vs.

- bari meʔ-swa-ne ɖiɖi-le haŋɖa-ŋɖæ pag=liʔ we-ke*
 then one-day=DEF how-INTERJ husband-3 break=shoot go-T/A
 ‘Then one day what happened but the husband went for shoots.’
 (Mahapatra and Zide no date, G.7)
- (b) *mæ qua wig-ke* vs. *mæ qua wig-ge*
 he home go-T/A he home go-T/A
 ‘He came back home.’ ‘He went home.’

Similar quasi-covert cislocative/translocative systems encoded in a small number of motion verb stems differentiated by two classes of tense/aspect markers, albeit with only partially overlapping formal markers, is found in Remo (Anderson and Harrison, this volume-a) and Gutob (Griffiths, this volume) as well.

3.2.7 Voicelversion

As in other South Munda languages, a causative may be marked by a prefix. In Gtaʔ this is realized as *aʔ-*. It has an infixed allomorph as well in *-aʔ-*. This distribution is subject to the same conditioning factors as in Kharia, Juang, and Sora, that is, the infix is used with polysyllabic (verb) stems, or functions as a prefix to the rightmost stem syllable. It adds a (usually animate) argument to a construction, which may be marked by the case suffix *-ke*, or serves as a mere transitivizer.

- (145) (a) *na sapa konɖa-ne jantu-ke wa-raʔ-sij-ce*
 you all mountain-GEN animal-CASE call-CAUS-gather-ss
ɖwa aʔ-gga
 home CAUS-REDPL:enter
 ‘You call and gather all the animals of the mountain and get them to
 enter the house.’ (Mahapatra and Zide no date, D.10)
- (b) *ggæʔ=seʔ+bog-paʔ=seʔ aʔ-lug-ge*
 REDPL:worship=speech+ECHO/TAG=speech CAUS-finish-T/A
 ‘He finished the death rites.’ (Mahapatra and Zide no date, B.6)
- mæ ljo big-ce siʔ+hæʔ-ce-ka bari taen kala*
 he land sow-ss cut-splash.water-ss-only again that day
poel aʔ-loʔ-cmu ɖij-ke
 first CAUS-fall-seed AUX-T/A
 ‘that day he prepared the soil by cutting weeds and splashing water and
 then sowed seeds of the first (crop).’ (Mahapatra and Zide no date, I.2)
- kiton-ke b-aʔ-to-ke dapre kiton btoʔ-ke*
 god-CASE frighten\CAUS-T/A immediately god fear-T/A
 ‘He frightened the god and the god became fearful.’ (Mahapatra and
 Zide no date, I.25)

Reciprocal is also marked morphologically in the Gta? verbal complex. In fact, it is likely that the most archaic means of realizing this category is to be found in Gta? where there is an infix *-m-* corresponding to North Munda *-p-* in conjunction with the prefix *ho-* cognate with the reciprocal found in other South Munda languages. In other words, Gta? may have preserved a Proto-Munda combination of prefix and infix to index reciprocal.

- (146) *wiŋ-haʔ-har-ke ho-m-m-og ho-t-m-u*
 Quarrel-PL-T/A RECIP-beat/RECIP/ RECIP-throw.stone/RECIP/
ho-s-m-iʔ+ho-s-m-aʔ-har-ke
 RECIP-cut/RECIP/+RECIP-catch/RECIP-PL-T/A
 ‘They beat each other, threw stones at each other, caught and butchered each other.’ (Mahapatra and Zide no date, F.17)

The *-m-* may have been preserved as *-n-* in a few Remo forms that have reciprocal semantics as well, but no cognate prefix. The *ho-* prefix may also appear alone with reciprocal semantics in Gta?. This is probably the more common of the two reciprocal patterns.

- (147) (a) *tur-la-ŋdʒig griŋ-ke a-ho-ba-ke*
 Search-COND-ADD cat-CASE NEG-RECIP-get-T/A
 ‘Although he searched, he did not meet the cat.’ (Mahapatra and Zide no date, H.19)
- (b) *meʔ-swa-ne kia-manda-hiŋ rsiŋ-ke bari ho-baʔlir-ke*
 One-day-DEF paddy-group-PL gather-T/A again RECIP-talk-T/A
 ‘One day all the paddies assembled and spoke with each other.’ (Mahapatra and Zide no date, C.4)

3.2.8 Finiteness

A small number of non-finite verb forms may be found in Gta?. One of these is the infinitive or gerund form. This is created from monosyllabic verb stems by a copy of the initial consonant as a prefix, yielding a geminate consonant. A certain number of complement taking or auxiliary verbs require their complement verb to appear in this reduplicated infinitive form, for example, *aʔlug*, *dʒiŋ*, or [*a-*]*miaʔ* (though with the latter it does not appear to be obligatory).

- (148) (a) *k-mæ tto+hhuʔ aʔlug-ce ko-læʔ-ge*
 DEM-3PR REDPL:pound+REDPL:ECHO/TAG CAUS-finish-SS sit-remain-T/A
 ‘This one finished pounding and remained sitting there.’ (Mahapatra and Zide no date, L.36)
- (b) *heʔ-dig tæŋ kitoŋ big-ne ceʔmwa dʒakce*
 today-even that god SOW-PRTCPL/NF/GEN grass COMP
kitoŋ-ke gtaʔ=re-hiŋ gge-miaʔ-ke
 god-CASE Gta?-people-PL REDPL:worship-CUST-T/A
 ‘Even today the Gta? people worship that god as he might [have] sow[n] grass.’ (Mahapatra and Zide no date, I.32)
- (c) *tte-la hiisa ndu-kuʔa ja par-le mæ paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come

ccoŋ *dij-le*
 REDPL:eat AUX-OPT
 ‘Thus there will be no envy; whoever wins, let him come and eat.’
 (Mahapatra and Zide no date, II.7)

A C₁-reduplicated verb form can appear with the definite (or genitive) suffix as well, for example, as the complement of *mile*. As this is a loan stem, this construction is not likely to be an old one in the language.

- (149) (a) *gte-la a-næʔ-ke gaʔi-si cili ccoŋ-ne mile-e*
 then OBJ-WE.INC-CASE more=day meat REDPL:eat-GEN available-FUT
 ‘Then we will have meat for eating for many days.’ (Mahapatra and Zide no date, D.12)
- (b) *bari jibon=re llæʔ-ne ʔan a-mile-ke*
 again life=man REDPL:remain-DEF place NEG-available-T/A
 ‘There was no place for the soul to stay.’ (Mahapatra and Zide no date, J.18)

Another non-finite form is the different subject or conditional suffix in *-la* (see sections 3.2.5 and 4.2.3). Corresponding to this is the same subject or coordinative/conjunctive suffix in *-ce* (section 4.2.3).

The *-ce* form of the verb *ɖak* ‘say, state’ has been grammaticalized as both a quotative and a complementizer with verbs of mental action.

- (150) (a) *bʔo-k[e]-ne dapre heʔ-baŋ næŋ ljo ceʔmwa*
 fear-T/A-NF immediately today-ABL I field grass
na-big-ɖe n[a]-á-big ɖakce basoŋ-ke
 2-SOW-OR 2-NEG-SOW QUOT say-T/A
 ‘Fearing, he said “from today will you or will you not sow my field?”’
 (Mahapatra and Zide no date, I.26)
- (b) *ɖæt neʔ-dij sina pe pe-ŋ-ɖo-la nor+nara-hij*
 That.like 1PL.INC-do though you.PL 2PL-AUGM-run-DS human.
being+ECHO-PL
ggweʔ-we-har-e ɖakce kæn-ne ɖira basoŋ-ke
 REDPL: die-AUX-PL-FUT QUOT this-DEF/GEN millet.species say-T/A
 ‘“Let’s do it like this (as an acceptable alternative), you run away, the human beings will all die,” this millet said.’ (Mahapatra and Zide no date, C.16)

Gtaʔ also makes use of a participle form that can function nominally or as an adjectival modifier. It is formed with the complex suffix *-k-ne*, probably historically, derived from the definite/genitive form of the verb marked in the *-ke* tense/aspect form. Some examples of its use follow below.

- (151) (a) *atte hur-k-ne-nia ɖig go-k-ne-nia ɖig*
 There guard-T/A-GEN-ADESS also hunt-T/A-GEN-ADESS also
bir=tia-ɖig a-tar-ke
 gayil-ADD NEG-appear-T/A
 ‘The gayil did not appear there, neither near the guard[s] nor the hunter[s].’ (Mahapatra and Zide no date, A.7)
- (b) *k-mæ-hij ke-la olæŋ-læʔ-k[e]-ne remwa-hij a-næ-ke*
 DEM-3PR-PL see-DS wander-AUX-T/A-DEP person-PL OBJ-WE-CASE

- k-mæ-hij a-mia?*
DEM-3PR-PL NEG-know
'They, (you see), are nomads, and they do not know us.' (Mahapatra and Zide no date, A.36)
- (c) *oriŋgi basa-k-ne mlæ-hū k-mæ-hij gsæŋ*
Oringi reside-T/A-DEP mlæ-son DEM-3PR-PL chicken
ŋko-hū=ja-ne rko? dæt bi?-e
peacock-child=paddy-GEN rice like.that give-FUT
'The original settlers of Oringi, those of the Mlæ clan, offer things like chicken and fine rice.' (Mahapatra and Zide no date, F.10)
- (d) *hɽij santa we-k-ne samwa sapa basoŋ-ke*
Afterwards market go-T/A-DEP story whole say-T/A
'Later he told the whole story about coming to the market.' (Mahapatra and Zide no date, J.28)

Duplication or full reduplication of a verb stem may also be found. Note the following two uses. In the first example, the duplication is iconic of an iterative action in an augmented same subject form (see section 4.2.3). The second form, on the other hand, shows a partially grammaticalized use of full reduplication in this capability construction.

- (152) (a) *koŋða-ne nla? koŋða-ne pa?dʉ koŋða-ne cucu*
mountain-GEN tuber mountain-GEN leaf mountain-GEN fruit
ale-kæn-ne coŋ-coŋ-ce-ka bari mæ-hij brwa-læ?-ge
all-this-DEF COPY-eat-ss-only again 3PR-PL live-AUX-T/A
'They lived only eating tubers, leaves, and fruits available in the mountains.' (Mahapatra and Zide no date, G.3)
- (b) *ŋku mæ? lig-ge dæk-ce hur-hur a-ia-ke*
tiger what work-T/A state-ss wait-wait NEG-CAP-T/A
'The tiger saying "what did he do" could not wait any more.' (Mahapatra and Zide no date, H.17)

3.2.9 Negation

The negative prefix in Gta? is *a[r]-*. In combination with a zero tense ending (or lack of a tense marker), this creates a negative present (sometimes future) form.

- (153) *k-mæ-hij ke-la olæŋ-læ?-k[e]-ne remwa-hij a-næ-ke*
DEM-3PR-PL see-DS wander-AUX-T/A-DEP person-PL OBJ-WE-CASE
k-mæ-hij a-mia?
DEM-3PR-PL NEG-know
'They, (you see), are nomads, and they do not know us.' (Mahapatra and Zide no date, A.36)

Third person modal forms are also formed via a straightforward combination of the negative prefix *-a* and the optative suffix *-le?*.

- (154) *ndu mæ-hij mia?-le? a-mia?-le?*
Well 3PR-PL know-IMP NEG-know-IMP
'Well, whether they know or not' (Mahapatra and Zide no date, A.37)

Only *-ke* is permitted in negative past contexts in Gta? (i.e. *a- -ke*), neutralizing the *-gel-ke* dichotomies at work in positive conjugations. Note the use of epenthetic (and possibly archaic) *-r-* preceding the causative prefix in the first example below. Note that some auxiliaries inflected in the negative past do, however, allow (or, in fact, require) *-ge*.

- (155) (a) *tæn k-n-ala e? sgwa oron ccoŋ*
 That time today like cooked.food REDPL:eat
a-mia?-har-ke e? sgwa paiŋi dŋij a-mia?-har-ke
 NEG-CUST-PL-T/A today like work do NEG-CUST-PL-T/A
 'At that time unlike today, they did not eat cooked food nor did they work (the fields) like today.' (Mahapatra and Zide no date, G.2)
- (b) *kæŋ dɔkri jiwa-manda-ke e-wa-la ja-rig a-we-ke*
 this old.woman animal-herd-CASE go-call-DS who-ADD NEG-go-T/A
 '(when) the old woman called the animals, [but] no one came.' (Mahapatra and Zide no date, D.27)

Subject prefixes come outside the negative prefix in Gta?. Note that this is also true in Juang and Gorum (Anderson 2007).

- (156) *na kia a?-ble na-coŋ-la næŋ ce?mwa n-ar-a?-ble*
 you paddy CAUS-ripen 2-eat-COND I grass 1-NEG-CAUS-ripen
ɖakce kmæ basoŋ-ke
 QUOT this.one say-T/A
 '“You cultivate paddy and eat it; shouldn't I cultivate grass” this one said.'
 (Mahapatra and Zide no date, I.18)

Second singular and first person non-singular subject prefixes often coalesce with the negative prefix and attract stress.

- (157) (a) *mæ?-sar n-a-big ɖakce mæ basoŋ-ke*
 what-PURP 1-NEG-SOW QUOT he say-T/A
 '“Why shouldn't I sow?” He said.' (Mahapatra and Zide no date, I.20)
- (b) *bʈo-k[e]-ne dapɽe he?-baŋ næŋ ljo ce?mwa*
 fear-T/A-NF immediately today-ABL I field grass
na-big-ɖe n[a]-á-big ɖakce basoŋ-ke
 2-SOW-OR 2-NEG-SOW QUOT say-T/A
 'Fearing, he said, “from today will you or will you not sow my field?”'
 (Mahapatra and Zide no date, I.26)

The negative prefix may occur in non-finite forms as well.

- (158) *a-mane-la bar-mmwiŋ ŋgir=boe? huŋ-dæ-ke bæ-ke*
 NEG-agree-DS another male child-3-CASE send-T/A
 'He disagreed and then sent another child, a male one.' (Mahapatra and Zide no date, I.12)

Prohibitives are formed by the combination *a-...-ge*. Dual and plural subject markers occur on the outside of the *-ge*.

- (159) (a) *a-poʔ-ge*
NEG-pierce-PROHIB
'Don't stab (me).' (Mahapatra and Zide no date, L.39)
- (b) *gte-la k-mæ basoŋ-ke pe a-ɔo-ge-pe*
this-DS DEM-3PR say-T/A you.PL NEG-flee-PROHIB-2PL
geʔ-la-pe ɔag-la ŋɔig ja-ɾig a-we-ke
come-IMP-2PL state-DS also who-ADD NEG-go-T/A
'Then he said "don't flee, please come," but nobody came/went.'
(Mahapatra and Zide no date, J.23)
- (c) *mbaya akwa ni[ʔ]-læʔ na ɔwa a-wig-ge*
Both here 1DL[:IMP]-remain you home NEG-go.home-PROHIB
'Let's both stay here, don't go home.' (Mahapatra and Zide no date, H.7)

In addition to the *a[r]-ge* prohibitive construction, prohibitives may also be augmented or softened by the particles *amuʔ* or *namuʔ*, which occur in clause-initial position often before a vocative.

- (160) (a) *ɔukri basoŋ-ke "amuʔ bobo tboʔ ke-la æg=bug*
Old.woman say-T/A don't child earth see-COND feces=pig
æg=sæŋ læʔ-ke ar-aʔ-loʔ-ge na wæʔ=go
feces=chicken remain-T/A NEG-CAUS-fall-PROHIB you cloth=fold
raŋ-jar-la-ce biʔ-la"
bring-climb.down-IMP-SS give-IMP
'The old woman said, "look, don't let them fall, there are pig and chicken feces remaining (on the ground), climb down and bring some up in the fold of your loin cloth and give (them) to me."' (Mahapatra and Zide no date, L.11)
- (b) *bari poʔ-la 'namuʔ aba' ɔak-ke*
Again stab-DS "please.don't father:voc" state-T/A
'Again he stabbed "please don't father" she said.' (Mahapatra and Zide no date, L.40)

Negative copula constructions are usually formed with *ŋɔdu*, also the interjection 'No!'. As discussed in section 3.1.2 above, the logical possessor may occur in a range of different case forms, dative, adessive, objective.

- (161) (a) *a-mæ-pa-ke go+gsiaʔ ŋɔdu*
OBJ-3PR-DL-CASE child+monkey NEG.COP
'They had no child[ren].' (Mahapatra and Zide no date, M.2)
- (b) *eʔ-ke næʔ-nnia swa ŋɔdu na wig-la-ce swa*
Today we-ADESS fire NEG.COP you go.home-IMP-SS fire
toʔ-raŋ-la ɔak-ce griŋ-ke toʔ=so bæ-ke
pull.out-bring-IMP state-SS cat-CASE pull.out=fire send-T/A
'"Today we have no fire, you go home and bring the fire," saying thus, the bear sent the cat to bring fire.' (Mahapatra and Zide no date, H.13)
- (c) *niã ni-læʔ-ke sina a-na ke-la knweʔ-huŋ*
We.DL 1DL-remain-T/A though OBJ-you see-COND/DS wife-child

nḍu a-næŋ knweʔ-huŋ nḍu
 NEG.COP OBJ-I wife-child NEG.COP

‘Though we both live, you see, you don’t have a wife or child and I don’t have a wife or child.’ (Mahapatra and Zide no date, H.6)

3.2.10 Derivation

Verb derivation is mainly unproductive in Gtaʔ. Causative and reciprocal stem formation were presented briefly in section 3.2.7. Stem compounding, akin to juxtaposition compounding in nominal derivation, and also showing distinct formal overlap with certain kinds of serial verb constructions or echo and tag constructions (see section 3.2.12) is also found to a certain extent in Gtaʔ as well. One such form is offered below.

- (162) *eʔ-ke mba-ya hara+jita neʔ-dij*
 today-CASE two-person defeat+win 1DL.INCL-do
 ‘Let’s today decide by stake.’ (Mahapatra and Zide no date, II.6)

3.2.11 Noun incorporation and combining forms

Gtaʔ, like many of its South Munda languages, but more so than any other except than Sora, makes extensive use of verb–noun stem compounding or noun incorporation. The forms of the nouns are the combining forms already alluded to and exemplified in section 3.1.10 above.

Combining forms often look like the root of the noun with a prefix or infix (or suffix) removed.

- | (163) Gtaʔ full forms | Gtaʔ combining forms | gloss |
|-----------------------|----------------------|-------------------|
| <i>ncu</i> | = <i>cu</i> | ‘oil’ |
| <i>titi</i> | = <i>ti</i> | ‘hand’ |
| <i>gsi</i> | = <i>si</i> | ‘louse’ |
| <i>gbe</i> | = <i>be</i> | ‘bear’ |
| <i>gnar</i> | = <i>gar</i> | ‘strip of bamboo’ |
| <i>remwa</i> | = <i>re</i> | ‘person’ |
| <i>kumḍa</i> | = <i>koŋ</i> | ‘pumpkin’ |

- (a) *ḍokra atæn-ne gbug=ci e-liŋ-raŋ-ke nlug=bug mari*
 Old.man that-DEF pig=meat go-lift-bring-T/A ear=pig again

gwaʔ=lug raŋ-wig-ke
 cut=ear bring-come.home-T/A

‘The old man went and brought that pig meat and cut and brought back home [pieces of] the pig’s ear.’ (Mahapatra and Zide no date, M.7)

- (b) *ḍukri gewa=mwaʔ-riaʔ-ce bason-ke ḥḍu ḍokra mari*
 Old.woman shy=eye-ECHO-SS say-T/A no old.man again

oʔna-riŋ ũ-we-e
 when-also 1-go-FUT

‘The old woman was embarrassed and said “No, old man, I will go there again sometime.”’ (Mahapatra and Zide no date, L.26)

- (c) *a-næʔ-ke kala-haʔ ḍekwa gæʔ-mu gæʔ=moaʔ-læʔ-e*
 OBJ-we.INC-CASE daily like.this fry=nose fry=eye-AUX-FUT
 ‘They shall be frying our noses and eyes up like this every day.’ (Mahapatra and Zide no date, C.5)

- (d) *ljo we-la mmwiŋ mæʔ-bare baya ɖokra sgwa ke-ce*
 field go-DS one what-INTERJ mad old.man like look-ss
giŋ=sia? giŋ=sia? nluŋ-ce clæ clæ ugboʔ-ce-ka
 cheek-monkey cheek-monkey ear-ss long long hair-ss-only
bari moʔmnæg-ne buti bwaʔ=tar-ce mmwiŋ unæŋ=u
 again very.big-DEF basket carry=shoulder-ss one cigar
la+toʔ-ceʔ-ka mæ big=cog ɖæte bbig-ɖiŋ-ke ljo
 hold.in-lips-ss-only s/he sow=basket with REDPL:SOW-AUX-T/A land
 ‘She went to the field, an old man looking like a cheek monkey was there
 with long hair, carrying a basket on his shoulder, with a cigar in his mouth
 was sowing the field with a basket.’ (Mahapatra and Zide no date, I.8)

In Gta?, as in all the other South Munda languages, combining forms are generally monosyllabic ‘root’ elements. Free-standing nouns in Gta?, on the other hand, are minimally bi-moraic: they must have at least some (derivational) element not found in the corresponding combining form. In Gta?, this latter constraint is rigid: there are no (indigenous) nouns not having at least two morae, while the vast majority of combining forms are monosyllabic and monomoraic (Mahapatra *et al.* 1989, Mahapatra and Zide 1972). This correspondence, between a monomoraic/monosyllabic combining form and a sesqui-syllabic/bisyllabic free form of nouns has parallels in Sora, in Proto-South Munda, and in both Khasi and Nicobarese (Anderson 2007, Anderson and Zide 2002).

Note that one of the noteworthy aspects of Gta? nominal incorporation is that this language permits an external modifier to modify or qualify a nominal element that has been incorporated into the verb.

- (164) Gta?
 [næŋ] *mbæʔsia? gweʔ-ti-ke*
 I left.hand wash-hand-T/A
 ‘Washed my left hand.’ (Sadock 1991: 97)

Gta? with its large degree of syntactic transparency between its incorporated noun and the external syntax of the clause is, thus, like Greenlandic and Ket in this respect (Anderson 2007), allowing for the syntactic transparency of incorporated nouns on a number of levels (incorporated nouns may ‘refer’ and be syntactically modified). Many other languages do not have such a transparent access between incorporation and syntax.

3.2.12 Auxiliary verb constructions, serial verb constructions, and other complex predicate types

Gta? makes use of a wide range of complex predicate types including two part co-lexicalized stems, serial verb constructions, light verb constructions, echo-formations, and auxiliary verb constructions of various types. The formal and functional properties of these numerous formations in the language could form a monograph in their own right (see Hook 1991 for a brief introduction).

Auxiliary verb constructions in Gta? express mainly aspectual and modal categories. The order is Lexical Verb–Auxiliary Verb as is expected in Eurasian SOV languages. Inflectionally speaking, auxiliary verb constructions (and really [almost]

all complex predicate structures) in Gta? are AUX-headed in the Anderson (2006) sense. The lexical verb in these structures either appears in a reduplicated form or in the unmarked/basic stem form. Whether or not speakers run together AVCs as single intonation units or not varies considerably.

Compound predicates, most (nuclear) serial structures and AUX-headed auxiliary verb constructions have the structure X_{α} subj-Y where X is one or more predicates serialized or otherwise concatenated followed by a fully inflected verb (subject, polarity, tense/aspect/mood).

An example of a compound predicate in Gta? is *lag -dɔ* ‘run away’ the uninflecting element precedes the inflecting element, which in turn is inflected for subject, and may be reduplicated if embedded in a larger (e.g. auxiliary) structure.

- (165) (a) *nɖu næʔ des lag-neʔ-dɔ dæk dækce-ka bari*
 No we.INC country away-1PL-run like.this QUOT-only again
kia-hiŋ ho-baʔlir-ke
 paddy-PL RECIP-talk-T/A
 ‘“No let us run away from this country” the paddies conversed’
 (Mahapatra and Zide no date, C.7)
- (b) *eʔ-ke-dɔg gbe gmiʔ ke-la lag dɔdɔ-miaʔ-ke*
 today-CASE-also bear goat see-COND run.:REDPL:away-CUST-T/A
 ‘Even today bears run away from goats when they see them.’ (Mahapatra and Zide no date, II.16)

Noun + verb predicates are found as well in a limited set of instances in Gta?.

- (166) *nɖia kuma-ce nji harke dɔŋ læʔ-la knweʔ-ræ gweʔ-we-ge*
 Water bath-SS three month do COP-DS wife-3 die-AUX-T/A
 ‘(after) she gave birth, three months passed and the wife died.’ (Mahapatra and Zide no date, B.4)

Auxiliary constructions often have the same formal structures as the complex serialized predicate, with the last auxiliary appearing in a fully inflected form (subject- and tense-marked).

- (167) *gbe basoŋ-ke ʔna to kala-haʔ ssæ paŋ-ce ccoŋ*
 bear say-T/A you surely daily before come-SS REDPL:eat
na-miaʔ-ke næŋ kala-haʔ hɔŋ ppaŋ m-miaʔ-ke
 2-CUST-T/A I daily afterwards REDPL:come 1-CUST-T/A
 ‘Bear says “every day you come before and eat and then I come.”’ (Mahapatra and Zide no date, II.5)

The use of or lack of reduplication (with serialized predicates when embedded) within larger auxiliary structures that at least in certain instances require or prefer reduplicated or ‘infinitive’ complement forms is complex in Gta? and the details remain to be worked out.

- (168) (a) *tte-la hɔsa nɖu-kuʔa ja par-le mæ paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come
ccoŋ dɔŋ-le
 REDPL:eat AUX-OPT
 ‘Thus, there will be no envy; whoever wins, let him come and eat.’
 (Mahapatra and Zide no date, II.7)

- (b) *na kala-ha? anđi we-na-mia?-ke iaŋ*
 You daily-EMPH where go-2-CUST-T/A mother:VOC
 ‘Mother, where do you go everyday?’ (Mahapatra and Zide no date, B.13)

Some common auxiliary stems in Gta? include *mia?* marking customary action, *ia* the capabilitive mood form, *lae?*, *bo*, and *điŋ* which are [im]perfective markers, *bi?* which can mark an intensive or a benefactive action (< ‘give’) and *ke* which creates an attemptive mood marker < ‘see’.

- (169) (a) *đæk điŋ-ce i?tuŋa bole a?-điŋ-ce mbar pwa?-coŋ*
 Like.this do-ss little cooked.rice CAUS-do-ss two leaf.cup
nturia? a?-điŋ-ce næ gta?=re-hiŋ ccoŋ+uug- mia?-ke
 gruel CAUS-do-ss we Gta?-person-PL REDPL:eat-REDPL:drink-CUST-T/A
 ‘and thus we did it this way, we Gta? people used to make two leaf-cups and eat [only] a little cooked rice with gruel.’ (Mahapatra and Zide no date, C.23)
- (b) *wig-la đwa wig-ce ke-la đwa muta*
 come-DS home come-ss see-DS home waterjug
a-lae?-ke ljo habo? bo-ke
 NEG-remain-T/A field forget AUX-T/A
 ‘He came home and saw (that) the water jug was not there; he forgot it in the field.’ (Mahapatra and Zide no date, I.4)
- (c) *gbe æg-a?-ca? a-ia-ke*
 bear shit-CAUS-arrive NEG-CAP-T/A
 ‘Bear could not shit (around all of it).’ (Mahapatra and Zide no date, II.14)
- (d) *at-baŋ gbe gmi?-ke bto-lae?-ke*
 there/then-ABL bear goat-CASE fear-AUX-T/A
 ‘Since then Bear has been afraid of Goat.’ (Mahapatra and Zide no date, II.15)
- (e) *ã=sañ-he? bæ-ke ã=sañ+ã=ɰia?-ce bini*
 Thatch=house-EMPH send-T/A thatch=house+ECHO/TAG-SS separate
a?-nsiŋ+a?-lwe coŋ-pe đđag điŋ-ge atæn-điŋ
 CAUS-boil+CAUS-ECHO/TAG eat-2PL REDPL:state AUX-T/A that-also
na lag-na-đo-ke đba
 you run.away 2-run.away-T/A come/go.IMP
 ‘He asked you to thatch the house; he was telling you “make a thatch house ready (for you) to prepare meals and eat separately” and for that reason you ran away; now let’s go.’ (Mahapatra and Zide no date, K.10)
- (f) *tela e? næ đidi ne-điŋ kitoŋ-ke mæ?-ɰig ne-pag=ra*
 then now we how 1PL-do god-CASE what-ADD 1PL-break=twig
pag=ra-ce-ka bari tæŋ koŋða e-go ne-ke
 break=twig-ss-only again that mountain go-hunt 1PL-AUX
đakce basoŋ-ke
 QUOT say-T/A
 ‘Then what are we to do now? Let’s make an offering of something to the god and then let’s try to go hunting on that mountain. (Mahapatra and Zide no date, A.51)

While it is not the place to go into a detailed argumentation in favor of this position, in addition to ‘auxiliary’ or ‘light’ verb constructions, there are also a range of other complex predicate types evidenced in Gta?. These include serial verb constructions in which there is a componential, combinatorial semantics associated with the individual elements and the larger construction, as well as so-called echo formations in which one element in the complex predicate structure is semantically opaque and co-lexicalized. Both structures are found in Gta? and both are briefly presented below.

In Gta? (and in a variety of other South Munda languages) there are a range of complex verb forms that are probably best analyzed as fused (or in some instances still free-standing) serial verb forms. Not infrequently, some element in a complex predicate in Munda languages is completely opaque in its meaning/function and origin and lexically determined as to which elements it may appear with. These are known as ‘echo’ forms (K. Mahapatra 1976, N. Zide 1976). A few examples of these will suffice to demonstrate the range of constructions found in Gta? of this type.

Most nuclear serial structures show a similar behavior in Gta? as well. Two or three stems may commonly be serialized (usually following the iconic order of events described in the complex predicate) followed by the last, appropriately tense- and subject-inflected final verb.

- (170) (a) *at-bay bari dia-dæ wig-bia-ke*
 That-ABL again mother-3 come.home-decline-T/A
 ‘From that time on, the mother declined to come home.’ (Mahapatra and Zide no date, B.24)
- (b) *e-tur n-ke-e*
 go-look.for 1-see-FUT
 ‘I will go and look for and find (girl for you).’ (Mahapatra and Zide no date)
- (c) *gmi? mmwiŋ mmwiŋ a?-diŋ-ce konða*
 goat one one CAUS-do-SS mountain
pwetur-æg-a?-ca?-ke
 surround-shit-CAUS-arrive-T/A
 ‘Goat one by one went around the mountain shitting.’ (Mahapatra and Zide no date, II.13)

Various serialized and auxiliated structures can combine to yield fairly large sequences of verbs.

- (171) (a) *næŋ hwe?-dʉlæŋ-n-læ?-e na a-næŋ-ke pnomæŋ*
 I pretend-sleep-1-AUX-FUT you OBJ-I-CASE anus
p-n-a?-cwe?-ɽia? kig-bi?
 millet.gruel pour.down-AUX
 ‘I’ll pretend to sleep, you smear some millet gruel around my anus.’ (Mahapatra and Zide no date, D.8)
- (b) *tæn ŋgire atwa?-ha? sle-gnar-bi?-læ?-ge*
 That young.man there-EMPH work-??-AUX:BEN-AUX-T/A
 ‘That young man worked there for them.’ (Mahapatra and Zide no date, K.5)

Auxiliaries and serialized forms may appear in head-to-tail linkage structures and in non-finite formations with this same inflectional pattern (with subject (and non-finite)

marking) found on the rightmost verb in the structure. Sometimes variation is seen between the relative order of elements in nuclear serialized forms in *Gta?*.

- (172) *bandʒi saliaʔ-we-ke* vs. *bandʒi e-saliaʔ-ke*
 Wage beg-go-T/A wage go-beg-T/A
 ‘came to beg her wages’ ‘came to beg her wages’
 (Mahapatra and Zide no date, A.24) (Mahapatra and Zide no date, A.26)

Notice the inflectional and morphophonological structure of echo-formations in *Gta?*. In most such formations, it is the final element that serves as the inflectional head (i.e. it mimics an AUX-headed inflectional structure in auxiliary verb constructions or serial verb constructions). Thus, in the following form the subject prefix and the tense clitic/affix are found on the final element, which in this particular example happens to belong to the class of opaque ‘echo’ formants. Such a formation may have derived, historically, from a (nuclear) serialized formation (cf. Anderson 2006, Bril 2004, Crowley 2002, Senft 2004).

- (173) *næŋ bagweʔ-jog n-talig-e*
 I kill-pick.up 1-ECHO-FUT
 ‘I will kill (you) and throw (you) away.’ (Mahapatra and Zide no date)

Other forms use less opaque second elements and are, thus, akin to the ‘semi-lexical’ nature of light verb formations as conceived of by Butt and Geuder (2001).

- (174) *bobo, a-na-ke mmwiŋ sela bihæ m-biʔ-e*
 child OBJ-YOU-CASE one girl marry 1-AUX-FUT
 ‘Child, I will marry you off to a girl.’ (Mahapatra and Zide no date)

In some forms one encounters a construction reminiscent of the split-doubled inflectional type (Anderson 2006), with subject marking on both elements in a construct but tense marking only on the final one in *Gta?*.

- (175) *ɖokra, na bole-nturiaʔ na-coŋ na-ug-k(e)=ɾe*
 old.man you rice-millet.gruel 2-eat 2-drink-T/A=Q
 ‘Old man, didn’t you drink/eat your rice millet and gruel. (Mahapatra and Zide no date)

On occasion, one finds subject marking on the first element and not the last one in *Gta?*. That is, the two elements form a tighter, compound-like bond, and, thus, functions as a unitary (historically complex) stem in this sense. These are also reminiscent of ‘nuclear’ serialized forms as described in the literature (Bril 2004, Crowley 2002, Foley and Olson 1985, Schiller 1990, Senft 2004), but appear to be left-headed, not right-headed.

- (176) *Gta?*
n-weʔ-gag-ce
 1-swing-tie-ss
 ‘After I swung and tied’ (Mahapatra and Zide, n.d.)

For more on *Gta?* complex predicate structures, see Anderson (2007) chapter 8.

3.3 Expressives

In *Gta?*, extensive use is made of forms that only occur as the second member of a co-lexicalized pair which are partially based phonologically on the basic stem. These

are known in the literature on Austroasiatic under a broad heading called expressives, that also include onomatopoeic and other sound symbolic sets of related stems. Some examples of these copy forms are illustrated in (177)

- (177) (a) *lag-ḍo-ce basoŋ-ke at hni-ne remwa ḍætte pe*
 ‘run.away-SS say-T/A there village-GEN person with you.PL
bari mæ-hiŋ ḍætte wiy-ha?+riy-ha? a-we-g[e]-pe
 then 3PR-PL with quarrel-ECHO NEG-GO-PROHIB-2PL
 ‘the head ran away and said “you don’t go fight with them, these villagers.”’ (Mahapatra and Zide no date, F.23)
- (b) *ḍæt ne?-ḍiŋ sina pe pe-ŋ-ḍo-la*
 That.like 1PL.INC-do though you.PL 2PL-AUGM-run-DS
nor+nara-hiŋ ggwe?-we-har-e ḍakce
 human.being+ECHO-PL REDPL:die-AUX-PL-FUT QUOT
kæn-ne ḍira basoŋ-ke
 this-DEF millet.species say-T/A
 ‘“let’s do it like this (as an acceptable alternative), you run away, the human beings will all die” this millet said.’ (Mahapatra and Zide no date, C.16)
- (c) *ḍukri gewa=mwa?-ria?-ce basoŋ-ke ṅḍu ḍokra mari*
 old.woman shy=eye-ECHO-SS say-T/A no old.man again
o?na-riŋ ũ-we-e
 when-also 1-go-FUT
 ‘The old woman was embarrassed and said “No, old man, I will go there again sometime.”’ (Mahapatra and Zide no date, L.26)
- (d) *mæ-pa mba-ya hara-hari ḍiŋ-ce kōḍa*
 he-DL two-person defeat:ECHO AUX-SS mountain
pwetur-æg-har-ke
 surround-shit-PL-T/A
 ‘The two of them trying to beat each other, went around the mountain shitting. (Mahapatra and Zide no date, II.12)
- (e) *a-mæ-hiŋ-ke oron+paron ṅḍu ki tmwa? ḍāton*
 OBJ-3PR-PL-CASE food+ECHO NEG.COP or mouth brushing
ṅḍu ḍæt ḍiŋ-ce k-mæ-hiŋ olæŋ-læ?-ke
 NEG.COP like.that do-SS DEM-3PR-PL wander-AUX-T/A
 ‘They have no food, no [way of] brushing teeth, they’ve been wandering around like that. (Mahapatra and Zide no date, A.33)

These are related to other co-lexicalized predicate pairings where the second member is not or differently related to the basic stem form (which may include incorporated nominals). Such formations include the echo-formations mentioned above (Mahapatra 1976, Zide 1976).

- (178) (a) *ggæ?=se?+bog-pa?=se? a?-lug-ge*
 REDPL:worship=spirit+ECHO/TAG=spirit CAUS-finish-T/A
 ‘He finished the death rites. (Mahapatra and Zide no date, B.6)
- (b) *mæ ŋku-nnia we-ce sarlo+uhwē-ce ko-ke*
 He tiger-ADESS GO=SS greet+ECHO/TAG-SS sit-T/A

‘He went up to the tiger, greeted him and sat down. (Mahapatra and Zide no date, H.4)

- (c) *ã=saj-he?* *bæ-ke* *ã=saj+ã=ʈia?-ce* *bini*
 Thatch=house-EMPH send-T/A thatch=house+ECHO/TAG-SS separate
a?-nsij+a?-lwe *coŋ-pe* *qɔqag* *qij-ge* *ataen-qig*
 CAUS-boil+CAUS-ECHO/TAG eat-2PL REDPL:state AUX-T/A that-also
na *lag-na-qo-ke* *qba*
 you run.away 2-run.away-T/A come/go.IMP
 ‘He asked you to thatch the house; he was telling you “make a thatch house ready (for you) to prepare meals and eat separately” and for that reason you ran away; now let’s go.’ (Mahapatra and Zide no date, K.10)

4 SYNTAX

4.1 Syntax of the simple sentence

The basic sentence order of Gta? is SOV. Of course, examples of this type with S and O represented by full NPs are relatively rare in texts and speech, but OV features and general Eurasian verb-final characteristics dominate Gta? syntax. Textually speaking, verb-final structure is nearly universal and speakers reject such formations with something other than the finite verb outside of their textual context.

- (179) (a) *oriŋgi* *basa-k-ne* *mlæ-hũ* *k-mæ-hij* *gsæŋ*
 Oringi reside-T/A-DEP Mlæ-son DEM-3PR-PL chicken
ŋko-hũ=ja-ne *rko?* *qæt* *bi?-e*
 peacock-child=paddy-GEN rice like.that give-FUT
 ‘The original settlers of Oringi, those of the Mlæ clan, offer things like chicken and fine rice.’ (Mahapatra and Zide no date, F.10)
- (b) *qæt* *basoŋ-ke* *dapre-ha?* *‘mæ? næŋ* *ljo na-big-ke’*
 like.that say-T/A instantly-EMPH not what I land 2-sow-T/A
qakce *kmæ basoŋ-ke*
 QUOT this.one say-T/A
 ‘Right away he spoke thus “why do you sow my land?”’ (Mahapatra and Zide no date, I.19)
- (c) *qokra* *mria?-ce* *qukri-ke* *basoŋ-ke* *na* *tojkae?*
 Old.man rise-SS old.woman-CASE say-T/A you pestle
ssa? *næŋ* *bha?-lo* *n-sa?-e*
 REDPL:hold I club 1-hold-FUT
 ‘The old man woke up and said to the old woman “you hold the pestle, I’ll hold the club.”’ (Mahapatra and Zide no date, D.23)

However, inverted syntax of various sorts may be found as well in Gta?. The nuances of these variants have yet to be investigated.

- (180) *ljo* *we-la* *mmwiŋ* *mæ?-bare* *baya* *qokra*
 field go-DS one what-INTERJ mad old.man
sgwa *ke-ce* *giŋ=sia?* *giŋ=sia?* *nlug-ce*
 like look-ss cheek-monkey cheek-monkey ear-ss

clæ clæ ugboʔ-ce-ka bari moʔmnæg-ne buti
 long long hair-ss-only again very.big-DEF basket
bwaʔ=tar-ce mmwiŋ unæŋ=u la+toʔ-ceʔ-ka
 carry=shoulder-ss one cigar hold.in-lips-ss-only
mæ big=cog dæte bbig-diy-ke ljo
 s/he sow=basket with REDPL:SOW-AUX-T/A land

'She went to the field, an old man looking like a cheek monkey was there with long hair, carrying a basket on his shoulder, with a cigar in his mouth was sowing the field with a basket.' (Mahapatra and Zide no date, I.8)

Positive copula sentences in the past use *læʔ-ge* and a case-marked possessor (see section 3.1.2 above). Negative copular forms are marked by *nɖu* and a case-marked possessor as well.

- (181) (a) *mæ-hiy-ne mbar-klig būŋi læʔ-ge*
 3PR-PL-GEN two-CLSSFR buffalo COP-T/A
 'They had two buffaloes. (Mahapatra and Zide no date, F.2)
- (b) *mmwiŋ sakar læʔ-ge a-mæ-ke mmwiŋ sela-[m]boeʔ*
 One rich.man COP-T/A OBJ-3PR-CASE one girl
huŋ-dæ læʔ-ge
 child-3 COP-T/A
 'There was a rich man, he had one daughter.' (Mahapatra and Zide no date, K.1)
- (c) *a-mæ-pa-ke go+gsiaʔ nɖu*
 OBJ-3PR-DL-CASE child+monkey NEG.COP
 'They had no child[ren]. (Mahapatra and Zide no date, M.2)

Quotations of direct speech are followed by the quotative *ɖakce*, a same subject/conjunctive non-finite form of the verb *ɖak* 'say, state'. Such a formation is common not only across Munda and the other languages of South Asia but of Eurasian SOV languages more generally (e.g. Tuvan (Anderson and Harrison 1999)).

- (182) (a) *andɪ we-pe-nɖiy-ke ɖakce basoŋ-ke*
 where go-2PL-DO-T/A QUOT say-T/A
 "'Where are you going?" he says.' (Mahapatra and Zide no date, C.11)
- (b) *mæ we-ce basoŋ-ke 'na mæʔ bbig na-diy-ke'*
 he go-ss say-T/A 'you what REDPL:SOW 2-AUX-T/A
ɖakce saliaʔ+ku-ke
 QUOT' ask+ECHO-T/A
 'he (=father) goes and says "what are you sowing?" he asked.' (Mahapatra and Zide no date, I.16)

Possessor raising may occur in Gtaʔ with pronominal possessors through the combination of case marking on the pronoun and body-part incorporation in the verb.

- (183) *a-næʔ-ke kala-haʔ ɖekwa gæʔ-mu gæʔ=moaʔ-læʔ-e*
 OBJ-WE.INC-CASE daily like.this fry=nose fry=eye-AUX-FUT
 'They shall be frying our noses and eyes up like this every day.' (Mahapatra and Zide no date, C.5)

An interrogative sentence may have a final interrogative particle especially if it lacks a Wh-word (which although present in the following sentence is probably a rhetorical or emphatic use of this pronoun).

- (184) *coŋa=bir ne-bi?-la k-mæ-hij a-næ-ke mæ? mane-e be*
 Lame=gayil 1PL-give-DS DEM-3PR-PL OBJ-we-CASE what obey-FUT Q
 ‘We will give them the lame gayil (but) will they obey us about anything?’
 (Mahapatra and Zide no date, A.35)

4.1.1 Typological features

Gta? exhibits most of the phrasal syntactic features that characterize other SOV language characteristics. For noun phrases, this means Noun-final phrases and such features as Numeral Noun, with the Numeral followed by a classifier for non-human nouns and some inanimates (185), Demonstrative Noun (186), Adjective Noun (187), Genitive Noun (188), and Relative Clause Noun (189).

- (185) NUM [CLSSFR] N
 (a) *ŋetwa ŋij-ce mbar hni-ne remwa-hij at*
 Like.that do-ss two village-GEN person-PL these
wij-ha?-har-ke
 quarrel-PL-T/A
 ‘Doing like that, the peoples of the two villages quarreled.’ (Mahapatra and Zide no date, F.16)
 (b) *mæ-hij-ne mbar-klig būji læ?-ge*
 3PR-PL-GEN two-CLSSFR buffalo COP-T/A
 ‘They had two buffaloes.’ (Mahapatra and Zide no date, F.2)
- (186) DEM N
kæn ŋokri jiwa-manda-ke e-wa-la ja-ŋig a-we-ke
 This old.woman animal-herd-CASE go-call-DS who-ADD NEG-go-T/A
 ‘(when) the old woman called the animals, [but] no one came.’ (Mahapatra and Zide no date, D.27)
- (187) GEN N
næ-ne hin-ne sla?
 we-GEN village-GEN tree
 ‘the tree of our village’ (Mahapatra and Zide no date, F.14)
- (188) ADJ N (also true in derivational compounds)
 (a) *kmæ-hij ŋhā kuru=paka we-har-ke*
 This:person-PL small pit=side go-PL-T/A
 ‘These people went to the small-pit side.’ (Mahapatra and Zide no date, A.5)
 (b) *gmi? basoŋ-ke mæ? ŋij-ge giá?=mwa? hŋæ?=?rij ŋag-ke*
 goat say-T/A what do-T/A crow=eye black=cat state-T/A
 ‘Goat said “so what, you crow-eyed black cat.”’ (Mahapatra and Zide no date, II.4)
- (189) REL N
 (a) *hŋij santa we-k[e]-ne samwa sapa basoŋ-ke*
 Afterwards market go-T/A-DEP story whole say-T/A
 ‘Later he told the whole story about going/coming to the market.’
 (Mahapatra and Zide no date, J.28)

- (b) *gweʔ-k-ne knweʔ-ræ*
die-T/A-DEP wife-3
'the dead wife' (Mahapatra and Zide no date, B.8)
- (c) *orīgi basa-k-ne mlæ-hū k-mæ-hij gsæŋ*
Oringi reside-T/A-DEP Mlæ-son DEM-3PR-PL chicken
ŋko-hū=ja-ne rkoʔ dæʔ biʔ-e
peacock-child=paddy-GEN rice like.that give-FUT
'The original settlers of Oringi, those of the Mlæ clan, offer things like chicken and fine rice.' (Mahapatra and Zide no date, F.10)

Features of verb phrase syntax characteristic of Gtaʔ include OV (190) order and V AUX (191) order.

(190) OV

- (a) *na sapa koŋða-ne jantu-ke wa-raʔ-sij-ce dʒwa*
you all mountain-GEN animal-CASE call-CAUS-gather-SS home
aʔ-gga
CAUS-REDPL:enter
'You call and gather all the animals of the mountain and get them to enter the house.' (Mahapatra and Zide no date, D.10)
- (b) *hni-[n]-[d]re basoŋ-ke næ ne-we-e tæŋ bir=tia-ke*
Village-GEN-person say-T/A we 1PL-go-FUT that gayil-CASE
ne-twiŋ-e
1PL-shoot-FUT
'The villagers said, "we will go, we will shoot down those wild cattle."' (Mahapatra and Zide no date, A.4)

(191) V AUX

- (a) *tte-la hīsa nðu-kuʔa ja par-le mæ paŋ*
that-DS envy NEG.COP-NEG:COND who can-OPT he come
ccoŋ dīŋ-le
REDPL:eat AUX-OPT
'Thus, there will be no envy; whoever wins, let him come and eat.' (Mahapatra and Zide no date, II.7)
- (b) *mæ-ne b-n-asaʔ mæʔ jontu-dīŋ we-la*
He-GEN residence/NOUN/residence what animal-ADD go-DS
abaʔ nðu mæ tmwaʔ-haʔ ga-uwe-dīŋ-ke
survival NEG.COP he mouth-EMPH enter-REDPL:AUX:TLOC-AUX-T/A
'Whatever animal went by his residence does not survive, but rather ends up in his mouth.' (Mahapatra and Zide no date, H.2)

4.2 Complex sentence structure

For the most part a detailed analysis of Gtaʔ syntax whether on the level of phrase or simple sentence (section 4.1) or on the level of complex sentence structure as briefly outlined below, remains an object for future investigation. Only a few cursory comments are offered here.

4.2.1 Relative-type clauses

Relative clauses in Gta? are left-branching and right-headed. The verb is marked by the suffix *-k-ne* keeping its own internal phrasal syntax. This appears pre-nominally.

- (192) (a) *tela mæ trwe? gwa we-k[e]-ne remwa-ke*
 Then he in company go-T/A-DEP man-CASE
e-wa-raŋ-ce salia?-ku-har-ke
 go-call-bring-SS ask-PL-T/A
 ‘Then he brought the man who had come (the market) together with him and they asked him.’ (Mahapatra and Zide no date, J.29)
- (b) *bari gag-bo-k[e]-ne goɾæ=o-ke to?-bi?-la tar-ce*
 Again tie-AUX-T/A-DEP boy=child-CASE open-AUX-DS come.out-SS
ɖukri-ne huŋ-ɖæ-ne c-n-og=mu
 old.woman-GEN child-3-GEN ornament/NOUN/=nose
c-n-og=mwa? tto?-bi?-ce tæŋ goɾæ=o-ke
 ornament/NOUN/=eye REDPL:open-AUX-SS that boy=child-CASE
a?-ga-ce mæ kia tto-ɖiŋ-ge
 CAUS-enter-SS he paddy REDPL:pound-AUX-T/A
 ‘Again she uncovered the confined boy [who] came out, took the old woman’s child’s nose ring and eye ornament ... then that boy started pounding paddy himself.’ (Mahapatra and Zide no date, L.33)

These *-k-ne*-marked verbal forms can sometimes be used nominally at least notionally, and take case suffixes, for example,

- (193) *atte hur-k-ne-nia ɖiŋ go-k-ne-nia ɖiŋ*
 There guard-T/A-DEP-ADESS also hunt-T/A-DEP-ADESS also
bir=tia-ɖiŋ a-tar-ke
 gayil-ADD NEG-come.out-T/A
 ‘The gayil did not appear there, neither near the guard[s] nor the hunter[s].’ (Mahapatra and Zide no date, A.7)

Note *go-k-ne* ‘hunter’ may also be [*g-]**go=re* as well, with optional reduplication and an incorporated noun formant *=re* ‘man/person’.

Some relative/correlative type structures may be found with certain interrogative and indefinite pronouns/quantifiers.

- (194) (a) *te-la tæŋ kiton bason-ke næŋ-ne paiŋi kæn-ha?*
 then that god say-T/A I-GEN work this-EMPH
na ɖi-rokom kia na-big-ke næŋ ɖiŋ
 you what-manner paddy 2-sow-T/A I also
se-rokom ce?mwa bbig-ɖiŋ-ke
 that-manner grass REDPL:SOW-AUX-T/A
 ‘Then that god said “my task is this: whatever manner you sow in, I will sow in that manner too.”’ (Mahapatra and Zide no date, I.17)
- (b) *hɿiŋ ɖi-le mæ jetek brwa-læ?-ge*
 afterwards what-INTERJECT he as.many live-AUX-T/A
o?-ɿi=si sgwa læ?-ge o?-tæn=si jako ce?mwa
 how.many=day like AUX/COP-T/A that.many=days till grass

mulke-se po?-ria? nɖu
 totally-DEF sprout NEG.COP
 ‘Afterwards, however long he lived, no grass ever grew up there again.’
 (Mahapatra and Zide no date, I.30)

Indefinite relative/correlative constructions of the ‘whoever ... he’ type are formed by the complex structure *ja X-la mæ* in Gta?.

- (195) (a) *gbe basoŋ-ke akæŋ koŋɖa ja olaŋ*
 Bear say-T/A this mountain who walk
a?-ca?-la mæ ssæ uli-alo
 CAUS-arrive.at.destination-DS he before mango-under
t^hwā-læ?-le?
 stand-remain-OPT
 ‘Bear says, “whoever gets to this here mountain first, let him stand under the mango tree.”’ (Mahapatra and Zide no date, II.9)
- (b) *gmi? basoŋ-ke dæt nɖu kæn kəɖa-ke ja*
 goat say-T/A that.like no this mountain-CASE who
pwetur-æg-a?-ca?-la mæ ssæ paŋ-coŋ-e
 surround-shit-CAUS-arrive-DS he before come-eat-FUT
 ‘Goat said “No, not like that, (I suggest) whoever can surround the mountain in shit, let him come and eat first.”’ (Mahapatra and Zide no date, II.11)

4.2.2 Other subordinate clauses (time, manner, cause, purpose, etc.)

A common structure in Gta? complex sentences is the causally dependent clause form, which is created by the polyfunctional non-finite marker or subordinator *-la*.

- (196) (a) *næ ne-læ?-la d^hā remwa ne-dij-e*
 We 1PL-remain-SUB small people 1PL-do-FUT
 ‘Because/if we remain[ed], we shall become a small people.’ (Mahapatra and Zide no date, C.18)
- (b) *a-maŋe-la bar-mmwiŋ ŋgir=boe? huŋ-dæ-ke bæ-ke*
 NEG-agree-DS another male child-3-CASE send-T/A
 ‘He disagreed and then sent another child, a male one’ (Mahapatra and Zide no date, I.12)

In other instances, the function of *-la* is as a different subject marker (see section 4.2.3). A kind of conditional is clearly intended in some of its uses, as in, at least, the second of its two uses below in (197(a)), the other being a formant of a kind of temporal/conditional clause.

- (197) (a) *e?-re ggo=re-hij koŋɖa-baŋ*
 now-time REDPL:hunt=person-PL mountain-ABL
wig-dij-la sela-mbwe?-hij kero kisa?lo gag-la
 come.home-AUX-DS female-PL way loin.cloth tie-DS
ggo=re-hij hni gga a-ia-har-ke
 REDPL:hunt=person-PL village REDPL:enter NEG-CAP-PL-T/A
 ‘Now if/when the hunters return home, and the women have blocked their access by tying loincloths together, the hunters can’t enter the village.’ (Mahapatra and Zide no date, A.63)

- (b) *dæk eʔ-baŋ næŋ ljo ceʔmwa na-big-la a-na-ke*
 look today-ABL I field grass 2-NEG-SOW OBJ-YOU-CASE
bol n-a-dij ɖæt ɖakce-ka bari kiton-ke bason-ke
 good 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god “look, you sow grass in my field (again), then I will not be nice to you.”’ (Mahapatra and Zide no date, I.28)

Other temporal subordinate clause functions of *-la* may be seen in such examples as the following:

- (198) (a) *pe-mba wig-ce a-pe-ke saliaʔ+ku-la*
 2PL-father come.home-SS OBJ-2PL-CASE ask+TAG-COND/DS
ni-iaŋ wig-læʔ-ge ɖakce a-bason-ge-pa
 1PL-mother:VOC come.home-AUX-T/A QUOT NEG-SAY-PROHIB-2DL
 ‘When your father comes back home and asks you, do not say “our mother has come home.”’ (Mahapatra and Zide no date, B.9)
- (b) *eʔ-ke-dij gbe gmiʔ ke-la lag ɖɖo-miaʔ-ke*
 today-CASE-also bear goat see-COND run.:REDPL:away-CUST-T/A
 ‘Even today bears run away from goats when they see them.’ (Mahapatra and Zide no date, II.16)

The combination of the non-finite marker in *-la* plus the emphatic or additive focus clitic *-[ŋ]ɖig* forms concessive clauses (‘although’). A similar process is seen in unrelated languages like Tuvan as well (Anderson and Harrison 1999).

- (199) *tur-la-ŋɖig griŋ-ke a-ho-ba-ke*
 Search-COND-ADD cat-CASE NEG-RECIP-get-T/A
 ‘Although he searched, he did not meet the cat.’ (Mahapatra and Zide no date, H.19)

The quotative *ɖakce* is becoming a type of complementizer used in certain complex sentences with complement (clause) taking predicates in Gta?. This may have the connotation of a causal complement in some instances.

- (200) *heʔ-dig tæŋ kiton big-ne ceʔmwa ɖakce kiton-ke*
 today-even that god sow-PRTCPL/NF/GEN grass COMP god-CASE
gtaʔ=re-hij gge-miaʔ-ke
 Gta?-people-PL REDPL:worship-CUST-T/A
 ‘Even today the Gta? people worship that god as he might [have] sow[n] grass.’ (Mahapatra and Zide no date, I.32)

The participial form in *-k-ne* can be used with *dapɖe* to mark either a concomitant action or temporal subordinate clause of prior or simultaneous action.

- (201) *hʔo-k[e]-ne dapɖe heʔ-baŋ næŋ ljo ceʔmwa na-big-ɖe*
 fear-T/A-DEP immediately today-ABL I field grass 2-sow-or
n[a]-á-big ɖak-ce bason-ke
 2-NEG-SOW QUOT say-T/A
 ‘Fearing, he said “from today will you or will you not sow my field?”’ (Mahapatra and Zide no date, I.26)

The element *dapɽe* by itself with a finite verb encodes the notion ‘as soon as’, that is, a kind of notionally subordinate temporal clause of immediately prior action.

- (202) *bba-ɽæ-ke basoŋ-ke dapɽe bba-ɽæ tar-ke*
 father-3-CASE say-T/A instantly father-3 go.out-T/A
 ‘As soon as he said this to his father, his father immediately left.’ (Mahapatra and Zide no date, I.14)

The participle *-k-ne* can also appear as *-k-nia*, which appears to be an adessive case form attached to the participle element *-k-* (the dependent marker *-ne* may well be the genitive case, historically). As in many Eurasian SOV languages, for example, Burushaski (Anderson 2002) or Mundari (Osada this volume), such ‘case marked clausal subordination’ constructions serve to form temporally subordinate clauses of the ‘when’-type.

- (203) *hɽiŋ nswar-la bura-nɽia? ppaŋ-dɽiŋ-k-nia*
 later.on dry-DS flood-water REDPL:COME-AUX-T/A-ADESS
 ‘Later on, when (you) have dried (me) out and the floods come.’ (Mahapatra and Zide no date, I.14)

The combination *-ce+ka bari* forms a kind of complex same subject conjunctive, marking prior action.

- (204) *mōjæ hni we-har-la mæ-hiŋ basoŋ-ke næ-ne hni-ne*
 Middle village go-PL-SUB 3PR-PL say-T/A we-GEN village-GEN
sla? kaɽɽu pe paŋ-ce-ka bari sapa
 tree branch? you.PL come-ss-only again entirely
gwe+hæɽ-bi?-pe-ŋ-dɽiŋ-ke
 cut-AUX-2PL-AUX-T/A
 ‘Because they settled in the middle of the village, those others said “you are coming here and cutting down all the trees of our village and their branches.”’ (Mahapatra and Zide no date, F.14)

4.2.3 Coordination and switch reference

Gta? has been described by Anderson and Boyle (2002) as having a system of switch reference. Here verbs in complex sentences appear in a non-finite form that marks various kinds of inter-clausal dependencies and simultaneously serve to indicate subject [dis]continuity with the following clause (whether finite or non-finite) and ultimately with the (usually) one finite marked verb permitted in a sentence.

The same subject marker in Gta? is *-ce(?)*. Examples of its use can be found in (205).

- (205) Same subject
 (a) *ɽukri hoɽ-ru=hoɽ-ria-ce swa e-rro-raŋ-ce*
 old.woman weep=ECHO-SS fire go-REDPL.carry-bring-ss
hanɽa-nɽæ-ne moɽ-ke cwar-ce
 husband-3.REF-GEN corpse-CASE dry-ss
aɽ-nswar-bo-ke
 CAUS-dry-AUX-PAST
 ‘The old woman wept a lot and then made a fire, dried up her husband’s corpse and preserved it.’ (Mahapatra and Zide n.d.)

- (b) *ŋku gnag-hwa? to?-ce ga-ge*
 tiger door-rope open-ss enter-PAST
 ‘The tiger opened the door(-rope) and entered.’ (Mahapatra and Zide n.d.)

As is apparent from the examples above, all but the last of a series of verbs with the same subject take the same subject marker, the last verb appearing with finite inflection. This pattern is one of the defining characteristics of switch reference systems. More examples of this in Gta? include sentences such as the following:

- (206) (a) *mæ ŋku-nnia we-ce sarlo+uhwẽ-ce ko-ke*
 He tiger-ADESS go=SS greet+ECHO/TAG-SS sit-T/A
 ‘He went up to the tiger, greeted him and sat down.’ (Mahapatra and Zide no date, H.4)
- (b) *mæ-hij int-baj janwai hni hãtar-ce*
 3PR-PL that.side-ABL Janwai village abandon-ss
tarpa-řæŋ lindži=kwi ale-tæn-ne sorte
 ceremonial.drum ceremonial.cooking.pot all-that-DEF entirely
sa?-ce at jar-ce hařa?-koŋða basa-har-ke
 hold-ss there climb.down-ss vicinity-mountain reside-PL-T/A
 ‘From that side they abandoned the village of Janwai, took up all their ceremonial drums and pots, climbed down and settled in the vicinity of the mountain.’ (Mahapatra and Zide no date, F.6)
- (c) *mria?-ce ttar-har-ke*
 Wake-ss REDPL:COME.OUT-PL-T/A
 ‘They woke up and came out.’ (Mahapatra and Zide no date, A.42)
- (d) *ttar-ce mmwiŋ sla? alo we-ce e-rsiŋ-har-ke*
 REDPL:COME.OUT-SS one tree under go-ss go-assemble-PL-T/A
 ‘They come out, go under a tree and (start to) assemble.’ (Mahapatra and Zide no date, A.43)
- (e) *me?=swa-ne-řak-ke gwe?-k[e]-ne knwe?-řæ qwa*
 One=day-DEF-state-T/A die-T/A-DEP wife-3 home
wig-ce huŋ-dæ-ke a?-bbu?+a?-ccoŋ-ce
 come.home-ss child-3-CASE CAUS-REDPL:SUCK+CAUS-REDPL:eat-ss
bole+hma? qwe+twe?-ce huŋ-dæ-ke a?coŋ+a?ug bari
 rice+curry cook+serve-ss child-3-CASE feed+give.drink also
huŋ-dæ-ke basoŋ-ke
 child-3-CASE say-T/A
 ‘It is said that one day the dead wife came home to feed and suckle her/ the children cook and serve up rice and curry, she fed her/the children and then said to the (older) child.’ (Mahapatra and Zide no date, B.8)

The same subject or conjunctive form can appear in an expanded form as *-ce+ka bari*

- (207) *mæ ljo big-ce si?+hæ?-ce-ka bari tæn kala poel*
 he land sow-ss cut-splash.water-ss-only again that day first
a?-lo? cmu qiŋ-ke
 CAUS-fall seed AUX-T/A
 ‘That day he prepared the soil by cutting weeds and splashing water and then sowed seeds of the first (crop).’ (Mahapatra and Zide no date, I.2)

The Gta? different subject marker is *-la*. It is cognate in form and function with the Proto-Gutob–Remo conditional and different subject marker **-na*.

(208) Different subject

- (a) *hɽij oʔɽi=mwa sgwa we-la ɖokra gweʔ=we-ge*
 later.on how.much=year like go-DS old.man die=AUX-T/A
 ‘Later on, after like several years passed, the man died.’ (Mahapatra and Zide no date)
- (b) *ljo haboʔ-bo-la huŋ-ɖæ-ke bæ-ke*
 field forget-AUX-DS child-3.RFLXV-CASE send-PAST
 ‘It (the jug) was forgotten on the field, (so) he sent his child.’ (Mahapatra and Zide no date: 32)
- (c) *a-mane-la barmmwij ŋgir=boeʔ huŋ-ɖæ-ke bæ-ke*
 NEG-agree-DS another male child-3.RFLXV-CASE send-PAST
 ‘She did not agree and he sent another male child.’ (Mahapatra and Zide no date: 33)

Other examples of the different subject marker in *-la* used conjunctively in Gta? include the following:

- (209) (a) *e-go-har-la mmwij moʔ-mnæŋ ɖij-k[e]-ne*
 go-hunt-PL-DS one very.big AUX-T/A-DEP
bir=tia mmwij tar-ke
 gayil one come.out-T/A
 ‘They started hunting and one very large gayil appeared.’ (Mahapatra and Zide no date, A.54)
- (b) *ljo haboʔ-bo-la huŋ-dæ-ke bæ-ke*
 land forget-AUX-DS child-3-CASE send-T/A
 ‘It was forgotten at the field so he sent his child (for it).’ (Mahapatra and Zide no date, I.5)
- (c) *tæn kitoŋ remwa-ke bweʔ-tur-la remwa-ke gæ-ke*
 that god person-CASE spit-DS man-CASE itch-T/A
 ‘The god spit on the man, and the man started to feel itchy.’ (Mahapatra and Zide no date, I.22)

The different subject and same subject elements can be used together, in appropriate contexts, to encode and track subject [dis]continuity across clauses in complex sentences in Gta?.

- (210) (a) *karwali li-ce cu-la goʔæ=o atwaʔ-haʔ*
 Bitter-gourd creep-SS bear.fruit-DS boy there-EMPH
llaʔ-miaʔ-k[e]-ɽak-ke
 REDPL:remain-CUST-T/A-state-T/A
 ‘It is said that the bitter gourd sent out creepers and bore fruit, and the child stayed there (as always).’ (Mahapatra and Zide no date, L.3)
- (b) *tar-ce ke-la mæ sote-haʔ bbig-ɖij-ke*
 go.out-SS see-DS he true=EMPH REDPL:SOW-AUX-T/A
 ‘He (=boy) went out and saw, he (=old man) really was sowing (the field).’ (Mahapatra and Zide no date, I.15)

In (211) through (215) five consecutive lines from a single text are offered. They demonstrate a number of interesting points germane to the present discussion. The first sentence (211) exhibits a straightforward use of the same subject marker.

- (211) *qæt qak-ce knwe? hanḍa wiŋ=haʔ-har-ke*
 like.that say-ss wife husband quarrel-PL-T/A
 ‘Speaking like that, the wife and husband quarreled.’ (Mahapatra and Zide, n.d.: 47)

The second sentence (212) exhibits head-to-tail linkage, that is, it starts with a resumption of the preceding sentence’s finite verb – a common narrative device in the South Asian linguistic area. However, this is not a rote repetition of the preceding verb in a particular, non-finite morphological form, as is found, for example, in certain ‘Aryanized’ Juang texts, but rather, this resumptive connective element is marked for whether it has the same or a different subject as the following clause. In this instance, the subjects are different, the husband and wife together vs. the wife alone (that is, there is only partial subject co-reference) and the form is accordingly found with the different subject marker. Similarly, the initial verb form in (214) is a resumptive form marked for different subject as well.

- (212) *wiŋ=haʔ-la meʔ-swa-ne knweʔ-ɾæ diḍile hli? sambo?*
 quarrel-DS one-day-DEF wife-RFLXV PRTCL shoot plant
we-ke
 go-T/A
 ‘They quarreled and one day the wife went to the bamboo-plant place.’ (Mahapatra and Zide, n.d.: 47)

In the beginning of the sentence (213) this ‘resumptive’ or head-to-tail linked use of the same subject marker is found. This sentence also has a fascinating example of an unconscious speaker correction with regards to both the use of the appropriate case inflection and switch reference marker. At first the speaker uses a verb in the different subject form and a nominal complement in the adessive case, realizes he misspoke and ‘corrects’ himself to use the oblique object case marker on the noun and the same subject marker on the verb. This sentence constitutes a clear and strong confirmation of the switch reference system of Gta?.

- (213) *hli? sambo? we-ce poga diḍile hli?-nnia*
 shoot plant go-ss tobacco PRTCL shoot-near
cu-biʔ-la tan hliʔ-ke cu-biʔ-ce wiŋ-ke
 smear-AUX-DS that shoot-CASE smear-AUX-ss go-T/A
 ‘She went to the bamboo-plant place, smeared tobacco on the shoots and returned home.’ (Mahapatra and Zide, n.d.: 47)
- (214) *wiŋ-la hɕiŋ hanḍa-ŋḍe pag=liʔ we-ke*
 go-DS afterwards husband-RFLXV break=shoots go-T/A
 ‘She went and afterwards, the husband went for bamboo shoots.’ (Mahapatra and Zide, n.d.: 47)

The final sentence (215) begins with a same subject marker on the first of two consecutive actions by the husband, the second of which is marked different subject, as the final predicate of the sentence refers to the shoots tasting bitter, not the husband.

- (215) *hli?* *pag-ce* *coŋke-la* *poga* *sgwa* *bsæ?* *læ?-ke*
 shoot break-SS taste-DS tobacco like bitter AUX-T/A
 ‘He broke the shoots and tasted them, they were bitter like tobacco.’
 (Mahapatra and Zide, n.d.: 47)

Sometimes sequences of juxtaposed finite clauses may be found as well for coordinative constructions.

- (216) *ɖukri* *haŋɖæ-ŋɖæ-ke* *hāwe* *pno?* *bæ-ke*
 Old.woman husband-3-CASE sharpen spear send-T/A
huŋ-ɖæ-ke *hā* *beŋɖa=ha?* *bæ-ke*
 child-3-CASE pluck okra=vegetable send-T/A
 ‘The old woman sent her husband to sharpen a spear and sent her child to pluck some okra.’ (Mahapatra and Zide no date, L.18)

Of course, nuclear serialization is possible in which the first verb simply appears in an unmarked form altogether.

- (217) (a) *tte-la* *hīsa* *ŋɖu-kuɖa* *ja* *par-le* *mæ* *paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come
ccoŋ *ɖiŋ-le*
 REDPL:eat AUX-OPT
 ‘Thus, there will be no envy; whoever wins, let him come and eat’
 (Mahapatra and Zide no date, II.7)

5 SEMANTICS/DISOURSE

Neither semantics of any domain really nor discourse on any meaningful level has been investigated *per se* in Gta?

5.1 Semantics

Some verbs in Gta? described very specific actions and concepts. A fair bit can be inferred about Gta? culture through a careful inspection of its lexicon (as is true of any language). Some examples of interesting Gta? words include

- (218) *a?-wæ?-bo?* ‘nod head horizontally’ *a?-ɖar-bo?* ‘nod head side to side’
a?-bæg ‘split open a crab’ *oo?* ‘carry baby in sling’
urig ‘sky to become clear of clouds’ *cog* ‘put on ornaments’
war ‘eat boiled bamboo’ *bno?* ‘ladder made of single bamboo’
 shoots in special way’
curia? ‘to flow out of nasal mucus’ *gæiŋ* ‘twist long fibers to make rope’
gotæ? ‘bring something from inaccessible place with help of long stick’
hmua?-bug ‘sliding door of pig-sty’ *koræg* ‘dried out from fire too long’
kuig ‘make hairknot’ *lbwe?* ‘broken rice particles’
nosor ‘to free from tiger’ *po* ‘kill louse by pressing under nail’
rgæiŋ ‘rub body against wall (like animal)’ *rri?* ‘eat flesh from bones’
snɖæg ‘symbolic ritual stool’ *sko?* ‘gourd-ladle’

5.2 Discourse

The study of discourse particles and discourse structures in Gta? remains embryonic. Certain interjections, clitics and particles appear in the text materials, but their use has not been explored at all. Such forms include *qi[qi][le]*, *-ha?*, etc.

- (219) (a) *bari me?-swa-ne qidi-le handa-ndæ pag=li? we-ke*
 then one-day=DEF HOW-INTERJ husband-3 break=shoot go-T/A
 ‘Then one day what happened but the husband went for shoots.’
 (Mahapatra and Zide no date, G.7)
- (b) *remwa qi-le kiton-ke bwe?-tur-la kiton-ke poka-poka*
 man oh.boy! god-CASE spit-DS god-CASE blisters
tar-ke
 emerge-T/A
 ‘The man spat on the god and blisters broke out on the god.’ (Mahapatra and Zide no date, I.23)
- (c) *wig-ce bba-ɾæ-ke bason-ke “ndu aba næ?*
 go/come-SS father-3-CASE say-T/A well father:VOC we:INC
ljo ja-ndæ-ha? kia bbig-dij-ke” dak-ce
 land who-3-EMPH paddy REDPL:SOW-AUX-T/A QUOT<state-SS
bba-ɾæ-ke bason-ke
 father-3-CASE say-T/A
 ‘She came home and told her father “well daddy, is someone (supposed to be) sowing our field?”’ (Mahapatra and Zide no date, I.10)
- (d) *tar-ce ke-la mæ sote-ha? bbig-dij-ke*
 go.out-SS see-DS he true=EMPH REDPL:SOW-AUX-T/A
 ‘He went out and saw, he (=old man) really was sowing (the field).’
 (Mahapatra and Zide no date, I.15)

Duplication or full reduplication may be used for emphasis in certain cases.

- (220) (a) *bba-ɾæ bason-ke ndu andi-andi-ne si-si=re-hij at*
 father-3 say-T/A no COPY-where-GEN plough=man-PL there
gali wig-dij-e dakce-ka bari bason-la-ndig
 way go.home-AUX-FUT QUOT-only again say-DS-also
a-mane-ke
 NEG-agree-T/A
 ‘Her father disagreed, “no, ploughmen might be passing through on their way home.”’ (Mahapatra and Zide no date, I.11)
- (b) *ŋku mæ? lig-ge dak-ce hur-hur a-ia-ke*
 tiger what work-T/A state-SS wait-wait NEG-CAP-T/A
 ‘The tiger saying “what did he do” could not wait any more.’ (Mahapatra and Zide no date, H.17)

Up to five iterations of a single verb stem can be found in such formations.

- (221) *dæt dij-ce riŋ-riŋ riŋ-riŋ riŋ-har-ce hriŋ*
 like.that do-SS REDPL:fight REDPL:fight fight-PL afterwards
kiton-ke b-a?-to-ke
 god-CASE frighten\CAUS-T/A
 ‘They fought like that and frightened the god. (Mahapatra and Zide no date, I.24)

6 LEXICON

6.1 Austroasiatic/Munda components

Much of the Gta? lexicon is both old Munda and old Austroasiatic. Gta? is in some senses the most archaic of the Munda languages in terms of noun structure, although its syllable structure, seemingly similar to that of various Mon-Khmer subgroups of Southeast Asia, is as mentioned above, likely to be a pseudo-archaic feature of the language, that is, one that is accidentally (?) mimicking that feature. Note that the second part of some CC clusters in onset position Gta? result from C₂ originally in a coda-position, for example, *sni* ‘sun’, < **siŋ*. Words of Munda/Austroasiatic origin in Gta? include *tiłnti* ‘hand’ *gcaey* ‘porcupine’ *m-mwa?* ‘eye’, *nlia* ‘tongue’ *shwe?* ‘belly’ *gni* ‘tooth’ *nlug* ‘ear’ among many, many others.

6.2 Loan strata

A number of words have been borrowed in Gta?, mainly from Desia and other local and world languages, even English, for example, *mas mas* ‘early night’ *somara* ‘Monday’ *dukan* ‘shop’ *mesfer* ‘teacher’ *iskool* ‘school’.

7 BRIEF ANALYZED TEXTS

7.1 Text 1: *ce?mwa kitonj* God of Grasses

- (i) *o?*=*sæ* *næ* *gta?*=*re* *mmwiŋ-ja* *læ?*-*ke*
 much.before we Gta? people one.person remain-T/A
 ‘In the old days, there was one of us Gta? people.’
- (ii) *mæ* *ljo* *big-ce* *si?*+*hæ?*-*ce-ka* *bari* *tæn* *kala* *poel*
 he land sow-ss cut-splash.water-ss-only again that day first
a?-*lo?*-*cmu* *điŋ-ke*
 CAUS-fall-seed AUX-T/A
 ‘That day he prepared the soil by cutting weeds and splashing water and then sowed seeds of the first (crop).’
- (iii) *big+ta-ce?*-*ka* *bari* *kri*=*lo?*-*nđæ* *đi?*-*pia?*-*ce?*-*ka*
 sow+?/-ss-only again clod.of.earth-etc. beat-break-ss-only
bari *hriŋ* *toŋ-gwar-la* *đile* *đwa*
 again afterwards passing.of.time-DS what:INTERJECT home
ug-nturia? *wig-ke*
 drink-gruel come.home-T/A
 ‘He sowed and loosened soil, it became late, so he came home for a meal.’
- (iv) *wig-la* *đwa* *wig-ce* *ke-la* *đwa* *muta*
 come-DS home come-ss see-DS home water.jug
a-læ?-*ke* *ljo* *habo?* *bo-ke*
 NEG-remain-T/A field forget AUX-T/A
 ‘He came home and saw (that) the water jug was not there; he forgot it in the field.’

- (v) *ljo haboʔ-bo-la huŋ-dæ-ke bæ-ke*
land forget-AUX-DS child-3-CASE send-T/A
'It was forgotten at the field so he sent his child (for it).'
- (vi) *sela-mboeʔ huŋ-dæ-ke bæ-ke ɖe bobo muta*
girl child-3-CASE send-T/A VOC child:VOC waterjug
e-jog-raj-la ljo læʔ-ke ɖakce basoŋ-ke
go-pick.up-bring-DS field remain-T/A QUOT say-T/A
'He sent his daughter; he said "hey sweetie, go bring back the water jug, it was left on the field."
- (vii) *goʔæ-o ljo we-ge*
child-masc field go-T/A
'The child went to the field.'
- (viii) *ljo we-la mmwiŋ mæʔ-bare baya ɖokra sgwa ke-ce*
field go-DS one what-INTERJ mad old.man like look-ss
giŋ=siaʔ giŋ=siaʔ nluŋ-ce clæ clæ ugboʔ-ce-ka
cheek-monkey cheek-monkey ear-ss long long hair-ss-only
bari moʔmnæŋ-ne buti bwaʔ=tar-ce mmwiŋ unæŋ=u
again very.big-DEF basket carry=shoulder-ss one cigar
la+toʔ-ceʔ-ka mæ big=cog ɖæte bbig-ɖiŋ-ke ljo
hold.in-lips-ss-only s/he sow=basket with REDPL:SOW-AUX-T/A land
'She went to the field, an old man looking like a cheek monkey was there with long hair, carrying a basket on his shoulder, with a cigar in his mouth was sowing the field with a basket.'
- (ix) *bbig-ɖiŋ-la mæ ke-ke bari wig-ke*
sow-AUX-DS s/he see-T/A again come.home-T/A
'She saw the one that was sowing and she went home again.'
- (x) *wig-ce bba-ʔæ-ke basoŋ-ke "ndu aba næʔ ljo*
go/come-ss father-3-CASE say-T/A well father:VOC we:INC land
ja-ŋɖe-haʔ kia bbig-ɖiŋ-ke ɖakce bba-ʔæ-ke basoŋ-ke"
who-3-EMPH paddy REDPL:SOW-AUX-T/A QUOT father-3-CASE say-T/A
'She came home and told her father "well daddy, is someone (supposed to be) sowing our field?"'
- (xi) *bba-ʔæ basoŋ-ke ndu andi-andi-ne si-si=re-hiŋ at*
father-3 say-T/A no COPY:where-GEN plough=man-PL there
gali wig-ɖiŋ-e ɖakce-ka bari basoŋ-la-ŋɖig a-mane-ke
way go.home-AUX-FUT QUOT-only again say-DS-also NEG-agree-T/A
'Her father disagreed, "no, ploughmen might be passing through on their way home."
- (xii) *a-mane-la bar-mmwiŋ ŋgir=boeʔ huŋ-ɖæ-ke bæ-ke*
NEG-agree-DS another male=child 3-CASE send-T/A
'He disagreed and then sent another child, a male one.'
- (xiii) *ŋgir-boeʔ-o we-ke-la mæ ɖig wig-ce*
male-child-MASC go-see-COND s/he also come.home-ss

- bba-ɾæ-ke basoŋ-ke*
father-3-CASE say-T/A
'When the male child went and saw, he came home and told his father.'
- (xiv) *bba-ɾæ-ke basoŋ-ke dapɾe bba-ɾæ tar-ke*
father-3-CASE say-T/A instantly father-3 go.out-T/A
'As soon as he said this to his father, his father immediately left.'
- (xv) *tar-ce ke-la mæ sɔte-ha? bbig-dij-ke*
go.out-SS see-DS he true=EMPH REDPL:SOW-AUX-T/A
'He (=boy) went out and saw, he (=old man) really was sowing (the field).'
- (xvi) *mæ we-ce basoŋ-ke 'na mæ? bbig na-dij-ke' ɖakce salia?+ku-ke*
he go-SS say-T/A 'you what REDPL:SOW 2-AUX-T/A QUOT ask+ECHO-T/A
'He (=father) goes and says "what are you sowing" he asked.'
- (xvii) *te-la tæŋ kitoŋ basoŋ-ke næŋ-ne paiŋi kæn-ha?, na di-rokom*
that-DS that god say-T/A I-GEN work this-EMPH you what-manner
kia na-big-ke næŋ diŋ se-rokom ce?mwa bbig-dij-ke
paddy 2-SOW-T/A I also that-manner grass REDPL:SOW-AUX-T/A
'Then that god said "my task is this: whatever manner you sow in, I will sow in that manner too."
- (xviii) *na kia a?-ble na-coŋ-la næŋ ce?mwa n-ar-a?-ble*
you paddy CAUS-ripen 2-eat-COND I grass 1-NEG-CAUS-ripen
ɖakce kmæ basoŋ-ke
QUOT this.one say-T/A
'"You cultivate paddy and eat it; shouldn't I cultivate grass" this one said.'
- (xix) *ɖæt basoŋ-ke dapɾe-ha? 'mæ? næŋ ljo na-big-ke'*
like.that say-T/A instantly-EMPH what I land 2-SOW-T/A
ɖakce kmæ basoŋ-ke
QUOT this.one say-T/A
'Right away he spoke thus "why do you sow my land?"'
- (xx) *mæ?-sar n-a-big ɖakce mæ basoŋ-ke*
what-PURP 1-NEG-SOW QUOT he say-T/A
'"Why shouldn't I sow?" He said.'
- (xxi) *ɖæt dij-ce mba-ya-no? atte wiŋ-ha?+riŋ-ha?-har-ke*
like.that do-SS two-person-DEF there rebuke-EMPH+fight-EMPH-PL-T/A
'Like that the two people quarreled there.'
- (xxii) *tæŋ kitoŋ remwa-ke bwe?-tur-la remwa-ke gæ-ke*
that god person-CASE spit-DS man-CASE itch-T/A
'The god spit on the man, and the man started to feel itchy.'
- (xxiii) *remwa di-le kitoŋ-ke bwe?-tur-la kitoŋ-ke poka-poka tar-ke*
man oh.boy! god-CASE spit-DS god-CASE blisters emerge-T/A
'The man spat on the god and blisters broke out on the god.'

- (xxiv) *dɛt dɪŋ-ce riŋ-riŋ riŋ-riŋ riŋ-har-ce hɪŋ*
 like.that do-SS REDPL:fight REDPL:fight fight-PL afterwards
kitoŋ-ke b-aʔ-ʔo-ke
 god-CASE frighten\CAUS\T/A
 ‘They fought like that and frightened the god.’
- (xxv) *kitoŋ-ke b-aʔ-ʔo-ke dapɽe kitoŋ bʔoʔ-ke*
 god-CASE frighten\CAUS\T/A immediately god fear-T/A
 ‘He frightened the god and the god became fearful.’
- (xxvi) *bʔo-k[e]-ne dapɽe heʔ-baŋ næŋ ljo ceʔmwa na-big-dɛ*
 fear-T/A-NF immediately today-ABL I field grass 2-sow-or
n[a]-á-big dɛkce basoŋ-ke
 2-NEG-SOW QUOT say-T/A
 ‘Fearing, he said “from today will you or will you not sow my field?”’
- (xxvii) *kitoŋ bʔo-ce basoŋ-ke næŋ n-a-big dɛkce basoŋ-ke*
 god fear-SS say-T/A I 1-NEG-SOW QUOT say-T/A
 ‘The god was scared and said “I will not sow (any more).”’
- (xxviii) *dæk eʔ-baŋ næŋ ljo ceʔmwa na-big-la a-na-ke bol*
 look today-ABL I field grass 2-NEG-SOW OBJ-you-CASE good
n-a-dɪŋ dɛt dɛkce-ka bari kitoŋ-ke basoŋ-ke
 1-NEG-do like.that QUOT-only again god-CASE say-T/A
 ‘Then he said to the god “look, you sow grass in my field (again), then I will not be nice to you.”’
- (xxix) *kitoŋ basoŋ-ke ʔna ljo eʔ-baŋ næŋ n-a-big dɛkce basoŋ-ke*
 god say-T/A you field today-ABL I 1-NEG-SOW QUOT say-T/A
 ‘The god says “from today on I will not sow your field any longer.”’
- (xxx) *hɪŋ dɪ-le mæ jetek brwa-læʔ-ge oʔ-ɽi=si*
 afterwards what-INTERJECT he as.many live-AUX-T/A how.many=day
sgwa læʔ-ge oʔ-tæn=si jako ceʔmwa mulke-se
 like remain-T/A that.many=days till grass totally-DEF
poʔ-ɽiaʔ nɔɽu
 sprout NEG.COP
 ‘Afterwards, however long he lived, no grass ever grew up there again.’
- (xxxi) *kmæ hɪŋ gweʔ-we-la mæ ljo+bri ceʔmwa*
 that.one afterwards die-AUX-DS he field+forest grass
poʔ-ɽiaʔ-miaʔ-ke
 sprout-CUST-T/A
 ‘Only after he died, did grass sprout again on his fields.’
- (xxxii) *heʔ-dig tæŋ kitoŋ big-ne ceʔmwa dɛkce kitoŋ-ke*
 today-even that god sow-DEP grass COMP god-CASE
gtaʔ=re-hiŋ gge-miaʔ-ke
 Gtaʔ-people-PL REDPL:worship-CUST-T/A
 ‘Even today the Gtaʔ people worship that god as he might [have] sow[n] grass.’

7.2 Text 2: *gbe-gmi?* Bear and Goat

- (i) *me?*-*swa-ne* *gbe* *uli-alo* *we-g[e]=dqak-ke*
 one-day-DEF bear mango-under go-T/A=state-T/A
 'It is told that one day Bear went under the mango tree.'
- (ii) *at[e]* *gmi?* *copa=u* *ccoŋ* *dij-g[e]=dqak-ke*
 there goat skin=mango REDPL:eat AUX-T/A=state-T/A
 'It is said that Goat was there eating mango skins.'
- (iii) *gbe* *basoŋ-ke* *arei* *clæ+nlug* *kaŋi+djiraŋ* *na* *mæ?*-*nsa?* *kala-ha?*
 bear say-T/A ho long+ear stick+horn you what-PURP daily
ra?=*sæ* *ra?*=*sæ* *ppaŋ* *na-mia?*-*ke*
 go.ahead=obstruct REDPL:COME 2-CUST-T/A
 'Bear said "who there long-ear stick-horn why do come first every day and block (me)?"'
- (iv) *gmi?* *basoŋ-ke* *mæ?* *dij-ge* *gia?*=*mwa?* *hŋæ?*=*riŋ* *ŋag-ke*
 goat say-T/A what do-T/A crow=eye black=cat state-T/A
 'Goat said "so what, you crow-eyed black cat."'
- (v) *gbe* *basoŋ-ke* *'na* *to* *kala-ha?* *ssæ* *paŋ-ce* *ccoŋ*
 bear say-T/A you surely daily-EMPH before come-SS REDPL:eat
na-mia?-*ke* *næŋ* *kala-ha?* *hŋiŋ* *ppaŋ* *m-mia?*-*ke*
 2-CUST-T/A I daily afterwards REDPL:come 1-CUST-T/A
 'Bear says "every day you come before and eat and then I come."'
- (vi) *e?*-*ke* *mba-ya* *hara+jita* *ne?*-*dij*
 today-CASE two-person defeat+win 1PL.INC-do
 'Let's today decide by wager.'
- (vii) *tte-la* *hisa* *nɖu-kuŋa* *ja* *par-le* *mæ* *paŋ*
 that-DS envy NEG.COP-NEG:COND who can-OPT he come
ccoŋ *dij-le*
 REDPL:eat AUX-OPT
 'Thus, there will be no envy; whoever wins, let him come and eat.'
- (viii) *gmi?* *basoŋ-ke* *mæ?*-*dij-ka?* *tæn-dig* *bol-ha?*
 goat say-T/A what-do-EMPH/MOD that-also good-EMPH
 'Goat says "whatever, that's also good."'
- (ix) *gbe* *basoŋ-ke* *akæn* *konda* *ja* *olæŋ* *a?-ca?*-*la*
 Bear say-T/A this mountain who walk CAUS-arrive.at.destination-DS
mæ *ssæ* *uli-alo* *t^hwā-læ?*-*le?*
 he before mango-under stand-remain-OPT
 'Bear says, "whoever gets to this here mountain first, let him stand under the mango tree."'
- (x) *gte-la* *mæ* *jine-e* *ssæ* *paŋ-ce* *kala-ha?* *ssæ* *paŋ-coŋ-e*
 this-DS he win-FUT before come-SS daily-EMPH before come-eat-FUT
 'This being so, he will win and come first, he will come and eat first.'

- (xi) *gmi? basoŋ-ke d̪æt n̪ɖu kæn kãḍa-ke ja*
 goat say-T/A that.like no this mountain-CASE who
pwetur-æg-a?-ca?-la mæ ssæ paŋ-coŋ-e
 surround-shit-CAUS-arrive-DS he before come-eat-FUT
 ‘Goat said “No, not like that, (I suggest) whoever can surround the mountain
 in shit, let him come and eat first.”’
- (xii) *mæ-pa mba-ya hara-hari ḍiŋ-ce kōḍa pwetur-æg-har-ke*
 he-DL two-person defeat:ECHO AUX-SS mountain surround-shit-PL-T/A
 ‘The two of them trying to beat each other, went around the mountain shitting.’
- (xiii) *gmi? mmwiŋ mmwiŋ a?-ḍiŋ-ce koḍa pwetur-æg-a?-ca?-ke*
 goat one one CAUS-do-SS mountain surround-shit-CAUS-arrive-T/A
 ‘Goat one by one went around the mountain shitting.’
- (xiv) *gbe æg-a?-ca? a-ia-ke*
 bear shit-CAUS-arrive NEG-CAP-T/A
 ‘Bear could not shit (around all of it).’
- (xv) *at-baŋ gbe gmi?-ke b̪to-læ?-ke*
 there/then-ABL bear goat-CASE fear-AUX-T/A
 ‘Since then Bear has been afraid of Goat.’
- (xvi) *e?-ke-ḍig gbe gmi? ke-la lag-ḍḍo-mia?-ke*
 today-CASE-also bear goat see-COND run.:REDPL:away-CUST-T/A
 ‘Even today bears run away from goats when they see them.’

NOTES

- 1 Almost all of the data used in the preparation of this chapter comes from an unpublished collection of Gta? texts collected by K. Mahapatra and N. Zide. The format for citing these examples is (Mahapatra and Zide no date) followed by an alphanumeric code. The letter refers to the text order in the collection and the number the sentence number from that text. Thus, H.2 means the second sentence from the eighth text. Two of these texts are given at the end of this collection and used with Professor Zide’s permission. These are listed as text 1 and text 2 and are cited in the chapter body in the following manner I.5 and II.12, which mean the fifth line of the first text and the twelfth line from the second of these texts.
- 2 Note that some other nouns show similar variation in the realization of their free form, for example, *nti* vs. *tti* ‘hand’ also with a syllabic nasal prefix in opposition to a reduplicated form.

REFERENCES

- Anderson, Gregory D.S. (2001a) ‘A new classification of Munda: Evidence from comparative verb morphology’, *Indian Linguistics*, 62:27–42.
- (2001b) ‘Santali’, in Jane Garry and Carl Rubino (eds) *Facts About the World’s Languages*, Bronx: H. W. Wilson, 623–628.
- (2002) ‘Case marked clausal subordination in Burushaski complex sentence structure’, *Studies in Language*, 26(3):547–571.

- Anderson, Gregory D.S. (2004a) *Auxiliary Verb Constructions in Altai-Sayan Turkic*, Turcologica 51, Wiesbaden: Harrassowitz Verlag.
- (2004b) ‘Advances in Proto-Munda reconstruction’, *Mon-Khmer Studies*, 34: 159–184.
- (2006) *Auxiliary Verb Constructions*, Oxford: Oxford University Press.
- (2007) *The Munda Verb: Typological Perspectives*, Berlin: Mouton de Gruyter.
- Anderson, Gregory D.S. and John P. Boyle (2002) ‘Switch-reference in South Munda’, in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 39–54.
- Anderson, Gregory D.S. and K. David Harrison (1999) *Tyvan*, Languages of the World Materials 257, München: Lincom Europa.
- (this volume-a, Chapter 11) ‘Remo (Bonda)’.
- (this volume-b, Chapter 7) ‘Sora’.
- Anderson, Gregory D.S. and Norman H. Zide (2002) ‘Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint’, in M.A. Macken (ed.) *Papers from the 10th Annual Meeting of the Southeast Asian Linguistics Society*, Tempe: Arizona State University, South East Asian Studies Program, Monograph Series Press, 55–74.
- Bhattacharya, S. (1975) *Studies in Comparative Munda Linguistics*, Simla: IAS.
- Bril, I. (2004) ‘Complex nuclei in oceanic languages: Contribution to an areal typology’, in I. Brill and F. Ozanne-Rivierre (eds) *Complex Predicates in Oceanic Languages*, Berlin: Mouton de Gruyter, 1–48.
- Butt, Miriam, and Wilhelm Geuder (2001) ‘On the (Semi)lexical status of light verbs’, in Norbert Corver and Henk Van Riemsdijk (eds) *Semi-Lexical Categories: On the Content of Function Words and the Function of Content Words*, Berlin: Mouton de Gruyter, 323–370.
- Crowley, Terry (2002) *Serial Verbs in Oceanic*, Oxford: Oxford University Press.
- Dryer, Matthew (1986) ‘Primary objects, secondary objects, and antitativity’, *Language*, 62:808–845.
- Evans, N. and T. Osada (2005) ‘Mundari: The myth of a language without word classes’, *Linguistic Typology*, 9(3):351–390.
- Ghosh, Arun (2003) *An Ethnolinguistic Profile of Eastern India: A case of South Orissa*, Burdwan: University of Burdwan.
- Hook, P.E. (1991) ‘The compound verb in Munda: An areal overview’, *Language Sciences*, 13(2):181–195.
- Mahapatra, K. (1976) ‘Echo formation in Gta?’, in P.N. Jenner, L.C. Thompson, and S. Starosta (eds) *Austroasiatic Studies, Part II*, Oceanic Linguistics Special Publication (13), Honolulu: University of Hawaii Press, 815–832.
- Mahapatra, K. and N.H. Zide (1972) ‘Nominal Combining Forms in Gta?’, *Indian Linguistics*, 33:79–102.
- Mahapatra, K.P. and N.H. Zide (no date) *Gta? Texts*, Unpublished-ms.
- Osada, T. (this volume, Chapter 3) ‘Mundari’.
- Senft, Gunther (2004) ‘What do we really know about serial verb constructions in Austronesian and Papuan languages?’, in I. Brill and F. Ozanne-Rivierre (eds) *Complex Predicates in Oceanic Languages*, Berlin: Mouton de Gruyter, 49–64.
- Zide, Norman H. (1968) ‘Some comparative notes on Gata? pronouns’, in J.C. Heesterman, G.H. Schokker, V.I. Subramoniam (eds) *Pratidānam. Indian, Iranian and Indo-European Studies Presented to Franciscus Bernardus Kuiper on his Sixtieth Birthday*, The Hague / Paris: Mouton, 348–358.

-
- (1969) 'Munda and non-Munda Austroasiatic languages', in Thomas Sebeok (ed.) *Current Trends in Linguistics*, The Hague: Mouton, 5:411–430.
- (1972) 'A Munda demonstrative system: Santali,' in J.M.C.Thomas and L. Bernot (eds) *Langues et techniques. Nature et Societe. Tom I*, Paris: Klincksieck, 267–274.
- (1976) 'A note on Gta? echo forms', in P.N. Jenner, L.C. Thompson, and S. Starosta (eds) *Austroasiatic Studies, Part II*, Oceanic Linguistics Special Publication (13), Honolulu: University of Hawaii Press, 1335–1343.
- (1999) 'Gta? -ke and -ge', Presented at the 209th AOS meeting, Baltimore.

ON NIHALI*

Norman H. Zide

1 OVERVIEW OF NIHALI

The interest in Nihali, such as it is, in certain narrow academic quarters, lies in the fact – possible fact – that it is (in interesting ways) no language at all but a ‘so-called’ or seeming language, and/or that it is a mystery, a lost – possibly ‘paleolithic’ – language (something like the Tasaday of the Philippines, what Tasaday was purported to be but without the heavy public relations flak that surrounded it). It is, perhaps, the only remnant in India of an ancient – pre-Munda, pre-Dravidian, pre-Indo-Aryan language family, with no living relatives, but, perhaps, a sister language of the language the Bhils spoke before they lost their own language and it was supplanted by the various Indo-Aryan ‘Bhils’. Nihali has been noticed by historical linguists for the very high percentage of borrowed vocabulary, and the variety of (proposed) sources for that borrowing, and the ‘suspiciously simplified’ syntax of the language. What is a mystery academically and popularly can be an administrative headache. Early notices of the Nihals describe them as nuisances, hill marauders and plunderers, ‘caterans’ who were ‘incorrigible’, and needed to be exterminated, and almost were on a couple of occasions. (I use Mundlay’s spelling, Nihali, which represents the local pronunciation; Kuiper and others write ‘Nahali’. Berger’s paper goes into the history of the name. The name the Nihals use for themselves is Kalto or Kaltu.) It is due to the work of Professor F.B.J. Kuiper that Nihali has been brought to the attention of Indologists, and what we say here addresses matters that Kuiper has been the first to foreground, and to treat in impressive detail.

Nihali has been referred to several times as a ‘so-called (sogennante)’ language or something similar by Koppers, Konow, and Kuiper. Even Fuchs expressed doubts about the language. The new many-volumed epitome of gazetteers and tribes-and-castes compendia, *People of India* (being issued by the Anthropological Survey of India) in its ninth volume, *Languages and Scripts* recognizes (and, on the strength of its own investigations[?], finds) the ‘Nahals’ as speaking – in different regions – Nimari or Korku (but not Bhili), but there is no mention of a spoken Nihali language. What is defective or ‘so-called’ about Nihali? Why is it not just a language, *comme les autres*? For Kuiper it is an argot (of what? or in what multi-lingual package?), and he talks of *Gaunersprachen* (secret languages used by criminals); Koppers seems to doubt that it is a full (complete) native language, the first language of anyone, the assumption being that all Nihali-speaking Nihals (a small minority of those identified in official records as Nihals) are bilingual, their other – full – language presumably being the North Munda language Korku. Kuiper reports Koppers’ mention of a collection of texts collected by Koppers and Fuchs, but Fuchs in a recent book presents a considerable amount of information about the Nihals and says something about their language but makes no mention of any text collection. (Mundlay found and worked with Nihali monolinguals as well as bilinguals,

and did collect texts. It is, as of the sixties and perhaps still, a first language, a home language, and most likely a 'full language', however we choose to define such a phrase. We will come to that later.) Konow was responsible for the data on the Munda languages and Nihali (in volume 4 of the *Linguistic Survey of India (LSI)*) and Koppers quotes his 1908 article: (that there is) 'one tribe, the so-called Nahals of Nimar, who were stated to speak Kurku in 1870, but who now speak (c. 1908) a mixture of Munda, Dravidian and Aryan dialects' – presumably our 'Nihali' or some variant of it. The evidence for their speaking Korku and nothing else in 1870, in any case, is shaky. The progress from monolingual Korku to some 'Nihali' seems unlikely. 'Nihali' has been in and out of the roster of Indian languages several times. Now you see it, now you do not.

Discussions of Nihali presuppose conjectural histories of the 'language', so that, for instance, it is not clear that calling it an 'argot' refers to present day use of Nihali usage (if Mundlay's data on monolinguals are accurate, and I think they are, then at least for some group(s) of Nihals Kuiper is mistaken) or to some earlier stage ('argotization'[?]) in the formation of Nihali. Is an argot 'stage' recognized in the formation of some component of other ('full') languages? If Kuiper is referring to phonological deformation and 'mutilation' – and he does talk of mutilation – does 'argot' have something to do with 'pidgin'?

Reasons adduced for doubting that Nihali – if there is one or a set of closely related dialects that are being consistently referred to in these publications – is a full-fledged language, and the first and/or only language of any speaker are the following: (1) the common association of Nihali-speaking Nihals (the estimate as of 1963 of the Nihal population was c.25,000, of which (Mundlay's estimate) perhaps 10% spoke Nihali) with Korkus and Korku villages. Presumably elsewhere in Nihal territory Nihals (i.e. de Candolle's Zones 2 and 3) they did not speak or know Nihali. (I doubt if anyone has made a careful investigation of Nihali language competence and use throughout the area. Mundlay has useful information of the Nihals she surveyed in the Melghat region.) Korku villagers I worked with when asked about the Nihali language (Mundlay assured me that there were Nihali-speaking Nihals living in the village) told me that the Nihals had no language of their own; they spoke Korku. The few extended descriptions of the Nihals are in books (Fuchs 1998, Hermanns, Koppers) primarily concerned with other groups: the Korkus or the Bhils, and this seems to be the characteristic angle of encountering and viewing the Nihals, when seeing them as anything but a source of civil disturbance and disruption. The few exceptions, papers addressing primarily Nihali matters, include the much quoted brief piece – the one source known to administrators or scholars having to find out something on the Nihals – in the Tribes and Castes of the Central Provinces (1916), and de Candolle's paper). (2) The (more than) apparent secrecy about the language and the general ignorance of its existence suggest to some that it isn't a language, but an argot or jargon used for certain limited purposes, the real language of these people (in Melghat/Nimar) being Korku, or in other areas – is the argot there completely gone? – some form of Nimari – or Bhili or Hindi or Marathi. (3) More interesting is the judgment that the language is limited, defective, perhaps a broken down descendant of an earlier 'full language', or a mixed language, and not adequate to the usual needs of linguistic communication. Thus, the need for Korku, or some other 'full' language. (4) An examination of the structures of the lexicon has led Kuiper to suggest that, in fact, Nihali is an argot (see below).

But the information on which all these generalizations have been made is limited. Schafer and Kuiper independently exhumed Nihali from the brief description in the *LSI* and observed that it was not – as Grierson/Konow implied (although not without reservations) – a North Munda language, probably closely related to Korku, but something else. The *LSI* has a few pages on Nihali. Bhattacharya, on one short field trip, collected a small amount of material on Nihali (see his article, but the further field trips to the Melghat area he intended (personal communication, S. Bhattacharya c.1966)) were not allowed by his superior, the then Director of the Anthropological Survey, Nirmal Kumar Bose. Kuiper's thorough study of Nihali of 1962 was based on the (limited) Nihali materials of the *LSI* and Bhattacharya. That's all that was available. All the surmises about the status of Nihali can be shot down by new and better observational data, and Mundlay provides some of that.

Nihals and Bhils, Nihali and 'Bhili'. For the connection of the Nihals with the Bhils, see Koppers (1948), and also Kuiper (1962) and Fuchs (1988). Koppers quotes Campbell (1880) who wrote that the Nihals 'are the most savage of the Bhils', but this, Koppers says, is not to be taken literally; the Nihal problem is complicated. (Koppers has a few notes about the Nihals in this book and in his *Geheimnisse*.) There is an extensive ethnographic literature on the numerous Bhil groups, and something, but much less, on the language(s). The 'Bhili language' is, apparently, a number of Indo-Aryan dialects of the regional languages in the extensive area of Bhil settlement (Gujarat, Rajasthan, Maharashtra, Madhya Pradesh – see Koppers' map). In some regions Nihals have been long associated with Bhils, have lived with them and still do (see de Candolle). (There are no – or few? – Nihali villages; Nihals live in sections of Korku, Bhil, and other [which?]) villages.) Fuchs also mentions Bhili-speaking Nihals as well as Korku-speaking and Nihali-speaking Nihals. The pre-Indo-Aryan language the Bhils probably spoke is lost, and we don't know its genetic affiliation. West-central India is almost entirely Indo-Aryan-speaking now. Presumably other linguistic families were more strongly represented in these areas in earlier times. There are other – fairly large – groups in central India, the Baiga for one, who now speak a variety of the local 'Hindi dialect', but who probably had their own, non-Indo-Aryan, language earlier. That 'Old Bhili' was related to 'Old Nihali' – that there was an ancient Nihali-Bhili family – is a plausible surmise (this was suggested by Koppers and by Shafer and accepted by a number of others), but as yet there is no linguistic evidence for it, and I have seen no strong claims based on ethnographic materials to support it. I examined one lexicon of Bhili – Thompson (1895) – and found no vocabulary cognate with the Old Nihali vocabulary identified by Kuiper, Shafer, and this author in the data Kuiper used and in Mundlay's data.

2 PROPOSED LINGUISTIC LAYERS IN NIHALI

2.1 Dravidian/Tibeto-Burman

Most of the Dravidian cognates adduced by Kuiper, Shafer, and Bhattacharya seem plausible. Pinnow in his review of Kuiper summarized his (Kuiper's) material on Dravidian influence: there are four strata (*Schichten*) – (Kuiper speaks only of sources of Dravidian words, not strata) with c.47–50 examples (9% of the total). If borrowing from Kurukh – one of the four strata – is relevant to the history of Nihali, it may be possible to date (approximately) some of the borrowing, given a

hypothesis about the location of the Nihals at the time of known movements north of the Kurukhs (Oraons). Burrow wrote a short notice of the book, but had no comments on the Dravidian material. The review provides a concise description of Kuiper's intentions, materials, conclusions; the only comment he allowed himself was to remark Kuiper's 'considerable reserve' – resistance – to accepting Nihali – the lexical remains after the borrowings have been extracted – as 'a language which is in origin quite independent' (of Indo-Aryan, Dravidian, Munda and Austroasiatic, Tibeto-Burman), which to Burrow 'seems most likely', and which, as I too see it, Kuiper's monograph makes a good case for. One would like a Dravidianist to go over the entire corpus (i.e. including Mundlay's material) and comment on the whole picture. I am no Tibeto-Burmanist but the Tibeto-Burman (possible) cognates – of Konow, Shafer, and Kuiper – I find less convincing. Here too, much new material – data and analyzes – is now available, and it would clarify several of the issues if someone familiar with it and the other Nihali-related material were to reevaluate the Tibeto-Burman connection. Certainly a few of the forms (Nihali *sunum* 'oil', from Korku and North Munda *sunum*) seem to have related forms in Tibeto-Burman, and there are certainly old Proto-Munda (and Proto-Austroasiatic) loans – whichever way the borrowing went (e.g. PM **kuXla* 'tiger), but, for instance, the geographical information Kuiper provides on the proximity of a Tibeto-Burman-speaking group, Limbu – 'not greater than about 130 miles', to (present day) Santal groups may not be relevant to Nihali although it is to North Munda since there is no reason to think that the North Munda connection (borrowing or whatever) was not primary, and that Nihali borrowed the form – as it did so many – later, from Korku. Kuiper mentions Konow's views on 'complex pronominalized Himalayan languages' and a Munda substratum as a contributor to their formation. Konow's views on these languages – and the Munda substratum – are not accepted by Tibeto-Burmanists today. Kuiper finds the Tibeto-Burman connection to be 'the most puzzling problem' of Nihali contacts, but, with caveats, goes on to find grammatical morphemes in Himalayan languages as (possibly) connected with Nihali morphemes.

2.2 Austroasiatic (apart from Munda)

The discussion of the possible connections of Nihali with Austroasiatic is based on material in Kuiper (1962) and on various papers of Pinnow's, which are also taken up in some detail by Kuiper (1972). The linguists who have done considerable work on Austroasiatic (primarily Mon-Khmer) in the last 35 years or so, that is, Shorto, Diffloth, and Ferlus, have had nothing to say about Nihali, probably because they don't find it to be (interestingly or at all) Austroasiatic, if, in fact, they find it at all. The identification of West (Munda) and East (Mon-Khmer) Austroasiatic cognates in general (of course, some languages have undergone more obscuring sound changes) has not been difficult. That the establishing of plausible Nihali cognates – the paucity of data making things that much harder – has been difficult and uncertain could be a result of several state of affairs, one being remoteness of relationship. Pinnow (1963) proposes a provisional ('the present state of investigation of the position of Nihali does not permit any definite judgment') western group of Austroasiatic languages which he calls Nahali–Munda, Nahali (now definitely judged not to be Munda) being western Nahali–Munda, and Munda being eastern Nahali–Munda.

He writes that the classification of Nahali is ‘particularly difficult’, in large part due to lack of data. The more interesting and difficult-to-explain connection of Nihali with Munda has to do with its morphology (‘Its morphological system ... is obviously connected with that of the Munda languages.’ See details in 1966a, and some remarks on Nihali verb morphology and its implications below under ‘What Kind of a (Contact-Shaped) Language is Nihali?’). Kuiper quotes Pinnow’s 1963 conclusion, ‘We may perhaps come closest to the truth if we assume that Nihali possesses an isolated non-Austroasiatic stratum that has been partially replaced by an Austroasiatic stratum which has also provided Nihali with its inflection.’ I would ask why ‘Austroasiatic’ here should not be replaced by ‘Munda’. In 1965 in the Austroasiatic pronoun paper Pinnow writes (again I am quoting Kuiper, 1972), ‘the personal pronouns of the disputed language Nihali can be classified with those of the Austroasiatic family, even though they are rather markedly distinguished from the personal pronouns of the other groups.’ The Nihali pronouns don’t look like Munda pronouns, and Pinnow finds a few similarities of individual Nihali pronouns with forms of similar meaning in Austroasiatic languages, for example, Khasi. I don’t find these miscellaneous similarities indicative of genetic relationship, and Pinnow himself expresses doubts in the paper. But in his 1966b review of Kuiper, Pinnow finds himself increasingly persuaded of Nihali’s fundamental Austroasiatic character (*Der grundlegend austroasiatische Charakter des Nahali schalt sich so nach und nach immer mehr heraus*). Kuiper writes that ‘my provisional attempt at an analysis of the case-endings and the pronouns did not confirm this assumption of an Austroasiatic provenance’. I agree with Kuiper in finding little evidence of Austroasiatic provenance. Kuiper’s ‘central problem’ in 1962, ‘how we must conceive the relations between the oldest Austroasiatic stratum and the other unidentified component of the language’ should perhaps now be decentered.

2.3 Munda

Apart from the numerous (transparent) borrowings from Korku, what has Nihali borrowed from Munda, or Munda from Nihali? First, of course, there is no assurance that all the Korku borrowings have been identified. And in the absence of sufficient possibly cognate vocabulary, no setting-up of sound correspondences (Nihali–Kherwarian or Nihali–South Munda(SM)) is possible, so that one goes by one’s own intuitions about relations of words – in one’s own style of negative capability. Examining the sets of words on Kuiper’s page 39, ‘A. More closely connected with North Munda (Kherwarian), and B. More closely connected with Central and South Munda’, I find several of the seven items in set A unacceptable or implausible, most importantly *te-* (Mundlay *te-*) ‘to eat’, which does belong here, but in set B. Of the items in set B, the words for ‘father’, *a-ba*, is pan-Munda, reconstructible – and not a loan – for Proto-Munda. The most interesting – and to my eye the most solid forms – are *be-* ‘to give’ (Mundlay *be-*), **er*, *ier-* ‘to go’, *piy-* ‘to come’, and *te-*. Although we have only these four words, the connection here is more persuasive to me than anything in Kuiper’s Munda alignments and the claims of cognation that go with them. These four do have good parallel forms in one or another branch of SM, and apparently no related forms in North Munda. I don’t accept Kuiper’s Santali *atin* as likely to be connected with Nihali *te-*. The actual forms and their antiquity – sub-family membership – will be discussed elsewhere. We

give here rough reconstructions: Gutob–Remo–Gta? **bəd-* ‘to give’, Gutob–Remo–Gta? **piŋ-* ‘to come’, Koraput Munda **tej-* ‘to serve food’, KM **ir/er-* ‘to run, jump, move.’) What do we make of this? The critical question – crucial to a hypothesis of SM subfamilies having borrowed from Nihali words that have no congeners in Mon-Khmer – of Austroasiatic cognates for these SM forms is as yet unanswered. That Nihali could have come into contact with SM languages is not at all unlikely. Certainly, some of the Koraput Munda languages (e.g. Gutob Gadba, now spoken only in Koraput District, Orissa was spoken further west, in Bastar, ca. 75 years ago, if not more recently), and we have no realistic notions of where and how the Nahals earlier ranged or came from. What sort of contact situations between what sort of groups, Munda and Nihali, could have resulted in the borrowing of basic lexical items? If the power of the Nihals earlier (as it was some time later) was military and their success as marauders (like the (SM) Remo (Bonda) today but in a smaller way, contained by the local district administration, or like the Comanche and other Plains Indians), what sort of linguistic impression would we expect the marauders’ language to make and under what conditions? Was there (intermittent?) occupation of the raided territories? Inter-marriage? The claim that some of the SM languages (Juang, for one) spoke something else before they adopted the ancestor of the Munda language they now speak is not new. That ‘something else’ could have been Nihali, or a sister language of Nihali. There are many possible scenarios to account for the lexical similarities (borrowing, presumably), but I want to affirm the importance of the identification (by Shafer and Kuiper) of these forms; they are less questionable and (therefore) more important than the other putative linkages proposed. In Kuiper’s discussion of Nihali and Austroasiatic he writes ‘the circumstance that the non-Kurku elements of the Nihali vocabulary cannot be attributed to any one of the sub-groups would seem to point to the conclusion that the older Munda stratum in Nihali stands somewhat apart from the sub-groups into which Munda is divided. Berger arrived at the same conclusion.’ If the Kherwarian similarities can be discounted, and I think they can, and the A and B sets are revised and realigned as proposed above, then perhaps (the corpus is still too small, but we can perhaps find more forms supporting this hypothesis) it is precisely one subgroup, SM, or perhaps some subfamily or subfamilies of SM that show(s) these lexical relationships, and it is SM (the SM family and/or one or another of its subfamilies) that has the connection with Nihali. I suggest that SM or KM has borrowed from Nihali, and Kuiper’s and Berger’s conclusion is wrong.

2.4 Argot

In 1962 Kuiper writes ‘In the case of Nihali, it is true, there are no certain indications of an analogous origin’ (he has been talking in the previous paragraph of metonymy and mutilation in speech disguise in various secret languages of the subcontinent) ‘of the names of parts of the body, etc., which categories are also in Nihali etymologically unexplained. Still, it may be useful not to forget that some of the obscure Nihali words may also belong to an argot, and need not necessarily date back to a linguistic pre-history of India.’ This is an interesting and useful warning. He mentions *jiki* ‘eye’ as perhaps connected with Santali *jhiki miki*, *jiki miki* ‘splendid, resplendent, shining, radiant’ – and also notes Ainu *shik(i)*.

(Kuiper in his earlier work on Proto-Munda words in Sanskrit brings in echo pairs of this sort, none of which are, as yet, reconstructable for Proto-Munda). In his 1972 review he writes (in a discussion of my inadequate treatment of Nihali in a survey article on the Austroasiatic languages of India): (Zide's) 'observations contain nothing new except the confirmation that Nihali is actually an argot, as had been suggested (by Kuiper) in 1962.' I did not think or say that Nihali was an argot, but that it was likely that 'Nihali was used as a more or less secret language'; Navajo was used during World War II by the United States military as a 'secret language' because it was unintelligible to the enemy; this did not make it an argot. It seems possible that some of the obscure Nihali words may belong to an argot, but there are – as Kuiper shows – other reasons for obscurity. Despite the interesting and not irrelevant discussion of *Gaunersprachen*, I see no good reason (the *jhiki miki* forms do not convince me) to claim that Nihali is an argot (now? at which previous stage? all of it? some section of the vocabulary? which?). It may be that the phonological distortions, lexical substitutions (rhyming slang, etc.) found in (other) secret languages are responsible for some of the Nihali vocabulary – certainly 'some of the obscure Nihali words' may be argot, that is, the result of speech disguising transformation and substitution, but this is something suggested here, and in no way demonstrated, and if it was something like rhyming slang (as in Cockney) there would be no way of retrieving the baseforms, and thus of proving that there was, in fact, this sort of distortion. Kuiper's reasons for proposing his argot hypothesis seem to be, first, the social position and criminal activities of the Nihals (which don't guarantee their possession and use of an argot), and, second, certain speculations about a few words in the old Nihali lexicon. I find the case unproven. I learn from Hal Fleming about 'jargons' in small, low status hunting and gather groups in East Africa where a small stock (~40 words) of 'jargon' has been recorded and the casual conclusion drawn from this short vocabulary is that the language 'is a jargon'. One needs to see how much and what segment of the lexicon is (speech disguise-derived) 'jargon' and what else 'the language consists of'. Kuiper's case for an 'argot', more explicitly, is the following (1962: 11–16): he first takes up the low status of the Nihals as a 'despised social group' and notes that other low status groups in India have secret languages. He then introduces various kind of phonological 'mutilation' found in such secret languages. He notes that words for body parts are commonly replaced in secret languages by disguised forms, and goes on to give the sources and derivations of some of the replacement forms, and, a bit later, suggests that Nihali *jiki* 'eye' might perhaps (originally) be a descriptive term. All of this is suggestive, but hardly probative, and I don't find it persuasive. In 1972 he is more positive about the argot hypothesis, and adduces some material (e.g. on Vedda) that might suggest analogs for what happened to Nihali, but again with nothing closer to a proof.

The quest for Nihali seems in some subtext to reveal a plot, one that Professor Kuiper most probably did not intend and would not accept: the voyage to Nihali as the grand occasion for wide-ranging exploration of Indo-Aryan, Dravidian, Austroasiatic, and Tibeto-Burman. Then, the winnowing of the Nihali lexicon, and the extraction of contact-derived matter. What is left is a small cache of semi-precious Old Nihali words, but this does not satisfy. A second voyage, on the Argo(t) – the golden fleece was plastic after all – leads to the discounting and discarding – throwing overboard – of some of that old Nihali vocabulary. Apart from these substantial, scholarly souvenirs de voyage, what is left of 'Nihali'? More,

I think, than the argot hypothesis seems to allow. Pinnow (1966b) agrees with Kuiper on the importance of the argot hypothesis (*Der Hinweis auf den möglichen Argot-Charakter des Nihali ist eins der wichtigsten Ergebnisse der Arbeit Kuipers, und seine Bedeutung kann nicht genug unterstrichen werden*).

We need to distinguish a 'functional argot' – that is, the use of a (secret) language for concealment, from an 'argot' (secret) language formed through processes of word-deformation, substitution, etc. (see Guiraud and Mehrotra. Mehrotra isn't aware of Kuiper's monograph. Kuiper is not aware of some of the earlier material mentioned in Mehrotra, for example, Sleeman on the language of the Thugs. The two discussions and bibliographies taken together provide a good survey of secret languages in India through c.1966).

The parallels with Sri Lankan Vedda and Rodiya are interesting (Vedda – Kuiper quoting de Silva – 'is a creole based on an older Vedda language with Sinhalese as the second contributing factor', (Kuiper) 'Rodiya is a secret language in which non-Sinhalese items are used in Sinhalese structures'). Can (our) Nihali be a creole based on an older Nihali language? Individual factors and contexts may be shared by Nihali (under various conjectures) and Rodiya, Vedda, etc. But, as Kuiper's data show, none of these cases is closely parallel to the Nihali situation. Nihali exhibits a wide range of linguistic contacts, many more than were available to Vedda or Rodiya. What the time scale is – in any of these cases – is still unknown. For Nihali, we assume that there was considerable mobility in a fairly extensive multilingual territory, so that such partly similar contact situations as that of Brahui or Vedda or Rodiya with massive borrowing (or deformation?) but less extensive linguistic contacts are only partly similar. Perhaps the language of the Thugs should be more closely examined. The Romani (Gypsy) sociolinguistic situations – one or more of them – seem more like what the Nihali situation(s) may have been, but for Romani we know where the people came from and, roughly, when (it is relevant that earlier speculations about the Gypsies posited a much more ancient exodus than the one scholars later reconstructed), and we know their original – re-exodus and pre-wandering – Indo-Aryan language and a fair amount about the languages they came in contact with, whereas for Nihali the ancestor language is presumably unattested outside the (observed) Nihali lexical corpus, and some of its possible contacts – as proposed by Kuiper – have yet to be more firmly demonstrated. And as Kuiper has shown the proportion of borrowed vocabulary in Nihali is very high, presumably much higher than in any of the Romani dialects (what the corresponding figures for Vedda are I cannot say). As with most everything else about the Nihals, we know little about their social or occupational history. They do not now and did not in the recent past own and cultivate fields (and there is no evidence that they practiced slash and burn agriculture) or cattle. In the Melghat they seem to have been associated with the Korkus (themselves known earlier as freebooters, but now settled agriculturalists), but we do not know how far back the connection goes. They were probably hunters and gatherers, and did more and less raiding of neighboring sedentary communities. Fuchs mentions that they are skilled trainers of dogs, and this skill is appreciated by the Korkus. As 'caterans' one wonders about their mobility. Did they have horses? They don't now, and neither do the Korkus, although (see Fuchs) there are representations of horses commonly on Korku wooden funerary tablets, and a taboo on eating horsemeat. Horses in that area would be expensive to keep, not particularly practical – bullocks are at least as efficient as ploughing and cart animals, and more docile, and healthier in that terrain. The word for 'horse' (see Kuiper) is *māv*, which Kuiper connects

with Dravidian (e.g. Telugu *māvu*) and possibly Tibeto-Burman and Tai forms. The Korku and other Munda forms (old borrowings) are not related. Hermanns has origin stories from Nihali informants according to which in earlier times the Nihals and Korkus were one people. (Not surprising, Korku informants deny this.) This means, I assume that they (all) then spoke a form of Nihali. Later the Korku upgraded themselves (giving up beef, certain low occupations, etc.) and, presumably, adopted a North Munda language from high status invaders/immigrants to the region. The Nihals were downgraded, and the Korkus have maintained the social distance (see Fuchs for Korku–Nihali interactions). There are references to Korku presence and activity in the fifteenth century, and to the Nihals ‘at the time of Akbar’, that is, the latter part of the sixteenth century, in both cases as hill robbers and freebooters. See Fuchs (1963) also on the antiquity of the Korkus (and, by implication, the Nihals who, these writers would claim, probably were in there earlier if in fact they were a different group) in the region, and that of the other North Munda (linguistic) groups in Bihar and adjacent regions. There are, of course, the usual putative identifications with peoples mentioned in the Ramayana (as *rāksasas* – demons, see Fuchs (1988), but also Zide (1972) on *Khara*). The social position and the marginal occupations of the Nihals suggest that they may well have had and used a secret language, early and late. Information on this may still be obtainable. That this is an argot (or that they use an argot – and what relation that argot has with ‘Nihali’) has yet to be proved.

3 WHAT KIND OF A (CONTACT-SHAPED) LANGUAGE IS NIHALI?

A much more thorough treatment of this topic is called for, but I offer here one possible scenario showing schematically how Nihali may have come to its present state (The data on the Nihali verb can be found in Grierson (1906), Pinnow (1966a) Bhattacharya (1957) and Kuiper (1962) as well as in Mundlay and Lynch, sources that were not available to Kuiper but do not describe a system that is significantly different). An examination of what Kuiper meant and might mean by ‘argot’ might introduce various linguistic–sociolinguistic–historical scenarios and tentatively try to place various statements and implications of Kuiper’s in such a scheme. Here is a preliminary attempt at doing just that (all of the assumptions and stages are arguable).

- (i) The (Old) Nihali language – not Austroasiatic – was spoken (perhaps in west-central India, when??) as a first language by, perhaps, a nomadic group, probably not agriculturalists, and probably not pastoralists either. These people may well have been bi- or multilingual. It was a representative of a family no longer found in India, apart from words inherited from that older lexicon in modern Nihali; there may well have been related languages in earlier times.
- (ii) In the course of wandering in or to the eastern parts of central India (I won’t try to break this down into ordered stages), there was borrowing to and from SM (not that we can identify which is which with much assurance, but see the section 2.3). They, Nihals, may have been more powerful at this time, more dominant as raiders. It is possible that they were technologically more advanced in some ways than the SM groups (but obviously not in agricultural techniques). The morphosyntax of Nihali, whatever it was – I am assuming that not much of it is left and/or identifiable – at this stage, whatever ‘natural’ changes it may have undergone, was fairly intact.

- (iii) Some borrowing from Dravidian, massive disruption of the Nihali community/communities, perhaps the decimation of the community on orders of local rulers and chiefs (see Fuchs (1988) or earlier traumatic reduction and breakup of the community (due to unknown causes). (Kuiper quotes Forsyth on the aboriginal races having been compelled to retire to the mountains before Hindu invaders. A few remained in the country occupied by the Hindus, chiefly in the position of agricultural serfs, or village watch-men, a description which, somewhat modified, could apply to the Nihals in the (non-Hindu) Korku villages in the present. Historical speculation about the position of the Nihals, early and late, is stymied by the lack of information on where ‘they’ were when). The old morphosyntax breaks up. Probably the Nihals (some group of them, one that remains, later ‘Nihali-speaking’) are bi- or multilingual and lean more heavily on one. Other language(s) are used for many vital functions, but retain some older Nihali, at least some of the lexicon. (On these massacres, Fuchs (1988) writes that in the latter part of the eighteenth century Mughals were being replaced by Marathas who heavily taxed not only the farmers of the plains but the tribals in the hills to which the tribal leaders responded by waging drawn-out guerilla-style warfare; this led eventually to their nearly complete annihilation (men, women, and children alike) by a group of Arabs in the service of Scindia. According to Russell and Lal, in the late sixteenth century the Raja of Jeetgurh and Mekote conducted a similar slaughter. Now the question remains open as to whether the Nihals referred to in these sources actually designates the Nihali-speaking ancestors of the present Nihals, but it seems likely.
- (iv) The older language is remade and socially reconfigured with a ‘creole-style’ grammar, this perhaps (some) used as a secret language, perhaps with argot-style speech deformation in some of the lexicon. The tense/aspect, etc. markers, indeed most of the morphemes used (as Pinnow 1966a points out) are familiar from a number of languages of the area, but they are not used in the same way, that is, this does not look like a case of ‘natural genetic’ inheritance. The Gutob (SM) language, shows an extreme example of tense suffixes splitting in confusing ways between the positive and the negative, but it still seems like a ‘natural’ development; the Nihali salad of verbal morphemes on the other hand, seems to show a disruption. The language is heavily re-lexified, with Melghat Korku being the lexifier (including words Korku borrowed from Indo-Aryan). This latter observation has relevance to the antiquity and nature of Korku–Nihali connections. If anciently, the Korku and Nihal were one people, and the Korku separated themselves and adopted a North Munda variety, were the Nihal and Korku still in contact in the kind of relationship (with the Nihals as servants of Korkus) we find today, and which is implied in the literature to have been the case for some time? If such was the case, the Nihals would have been – as they are now – bilingual in Korku (and perhaps, as now, familiar with other local languages as well). This would suggest that the borrowing from Korku started very early, and not, as it looks, fairly late. It is simpler to say that while the Korku–Nihali connection may be an old one, the heavy relexification dates to a comparatively recent period when the Nihal community was badly broken up, reduced, scattered, and that then some of the Nihals retreated to the Korku area, and recuperated a social organization and a Nihali language. As to possible influences of or direct contact with non-Munda Austroasiatic or Tibeto-Burman languages, as the above discussion indicates, I have my doubts

about these – they could have been acquired in the wander years of the Nihals, and clearly the Nihals have moved around. This scenario rules out the ‘fundamentally Austroasiatic character’ hypothesis put forth by Pinnow. If Nihali has a fundamentally Austroasiatic character, and acquired it in the usual historical linguistic ways, then my scenario is all wrong – unless the connection is remote indeed, in which case we would want to have heard more from Pinnow about that fundamental Austroasiatic character.

On questions of syntax that have not been examined, and not mentioned in this study, consult Mundlay materials. In this ‘remaking’, was what was *not* borrowed (at one stage or another), and is retrievable and assignable to ‘Old Nihali’ (at one stage or another) just a small set of words? We have said nothing about Nihali syntax and how it compares to Korku or Hindi or Marathi or (some) Dravidian. The SOV word order, use of postpositions, etc. is common to all the languages in the area. But it is not clear, for example what relative-clause-type structures are like in Nihali. How ‘simplified’ or ‘reduced’ or ‘creole-like’ is Nihali grammar? Pinnow has suggested that the verb system is like that of Proto-Munda. I don’t see this.

Younger, Marathi-speaking historians of the region might be encouraged to look at the Nihals and their history. If concentrated subalternity is of interest, are they subaltern’s (Hindi- and Marathi-speakers) subaltern’s (Korku) subalterns? How ‘other’ in the non-urban Indian scene can one get? Perhaps some illumination of the linguistic problems will come out of a better understanding of where the Nihals were and what they were doing and saying to whom.

As a final word, note that the paucity of data on Nihali has not prevented megalocomparativists from finding a home for it. Thus, Bengston (1994) finds a place for Nihali in his Macro-Australic superstock. Most linguists do not accept such proposals.

NOTES

- * EDITOR’S NOTE: This paper is a revised version of a paper first published in *Mother Tongue* in 1998. The author retained the copyright. Many authors use Nihali and Nahali interchangeably. Zide in this article uses Nihali for the language (except when discussing authors who use Nahali).

BIBLIOGRAPHY

- Bengston, J. (1994). ‘Note on Macro-Australic’, *Mother Tongue* 23: 73–75.
 Berger, H. (1959) ‘Deutung einiger alter Stammesnammer der Bhil aus der vorarischen Mythologie des Epos und der Purana’, *Wiener Zeitschrift für die Kunde Süd- und Ostasiens und Archiv für indische Philosophie* III: 34–82.
 Bhattacharya, S. (1957) ‘Field Notes on Nahali’, *Indian Linguistics* 17: 245–258.
 — (1977) ‘The Munda Languages in Ancient India’, *Journal of Indian History* 55: 25–33.
 Burrow, T. (1958) ‘Sanskrit and the Pre-Aryan Tribes and Languages’, *Bulletin of the Ramakrishna Institute of Culture*, February 1958, Transaction 19.
 — (1963) ‘Review of Kuiper’s Nahali’, *Bulletin School of Oriental and African Studies* 26: 436–437.

- De Candolle, B. (1961) 'Contribution a l'étude des Nahals', *Anthropos* 56: 750–788.
- Forsyth, J. (1870) *Report of the Revenue Settlement of British Nimar District of the Central Provinces*.
- (1872) *The Highlands of Central India*, London: Chapman and Hall.
- Fuchs, S. (1963) 'The Aboriginal Races in North-Eastern Central India' in *Festschrift Paul Schebesta*, Studi Instituti *Anthropos* 18, Wien: St. Gabriel Verlag, 349–355.
- (1988) *The Korkus of the Vindhya Hills*, New Delhi: Inter-India Publications.
- Grierson, G. (ed.) (1906) *Linguistic Survey of India IV: Munda and Dravidian*. [Konow's article on Nahali is on pp 185–189] Calcutta: Government printing office.
- Guiraud, P. (1956) *L'Argot Que Sais-je*, No. 700, Paris: Presses Universitaires de France.
- Hermanns, M. (1966) *Die religiose-magische Weltanschauung der Primitivstämme Indiens, Band 2: Die Bhilala, Korku, Gond, Baiga*, Wiesbaden: Franz Steiner Verlag.
- Konow, S. (1908) 'Notes on the Munda family of speech in India', *Anthropos* 3: 68–82.
- Koppers, W. (1947) *Geheimnisse des Dschungels*, Luzern: Josef Stocker Verlag.
- (1948) *Die Bhil in Zentralindien*, Horn-Wien: Ferdinand Berger Verlag.
- Kuiper, F.B.J. (1962) *Nihali: A Comparative Study*, Amsterdam. N.V.Noord-Hollandsche Uitgevers Maatschappij (Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen: Afd. Letterkunde; N.R.25.5).
- (1967) 'The Genesis of a Linguistic Area', *Indo-Iranian Journal* 10 (2/3): 81–102.
- (1972) 'Review of Current Trends in Linguistics', *Indo-Iranian Journal* 14(3/4): 292–293.
- Lynch, O. (1963?) 'Outline of Nihali Grammar', Unpublished typescript.
- Mehrotra, R.R. (1977) *Sociology of Secret Languages*, Simla: Indian Institute of Advanced Study.
- Mufwene, S. (1995) 'Jargons, pidgins, creoles, koinés: What are they?', in A. Spears and D. Winford (eds) *Pidgins and Creoles: Structure and Status*, Amsterdam: John Benjamins, 35–70.
- Mundlay, Asha (1964/1972) 'Nihali (Kaltu-Mandi): Its place among the Munda Languages of India', Unpublished typescript.
- (1979) A Nihali Lexicon, Unpublished typescript.
- Pinnow, H.-J. (1963) 'The position of the Munda languages within the Austroasiatic language family' in H. L. Shorto (ed.) *Linguistic Comparison in Southeast Asia and the Pacific*, London: SOAS, 140–152.
- (1966a) 'A comparative study of the verb in the Munda languages' in N. Zide (ed.) *Studies in Comparative Austroasiatic Linguistics*, (Indo-Iranian Monographs v. 5), The Hague: Mouton, 96–193.
- (1966b) 'Review of Kuiper's Nihali', *Orientalische Literaturzeitung* 61: 492–496.
- Russell, R.L. and H. Lal (1916) *Tribes and Castes of the Central Provinces, volume 4*, London: Macmillan and co.Ltd.
- Shafer, R. (1940) 'Nahali: A Linguistic Study in Paleoethnography', *Harvard Journal of Asiatic Studies* 5: 346–371.
- Singh, K.S. and S. Manohoran (1993) *People of India, National Series Volume 9 Languages and Scripts*, Anthropological Survey of India, New Delhi: Oxford.
- Thomason, S. and T. Kaufman (1988) *Language Contact, Creolization and Genetic Linguistics*, Berkeley: University of California Press.
- Thompson, C.S. (1895) *Rudiments of the Bhili Language*, Ahmedabad.

- Zide, N.H. (1968) 'Some comparative notes on Gata? pronouns', J.C. Heesterman, G.H. Schokker, V.I. Subramoniam (eds) *Pratidānam. Indian, Iranian and Indo-European Studies Presented to Franciscus Bernardus Kuiper on his Sixtieth Birthday*, The Hague / Paris: Mouton, 348–358.
- (1969) 'Munda and Non-Munda Austroasiatic languages', in Thomas Sebeok (ed.) *Current Trends in Linguistics*, The Hague: Mouton, 5: 411–430.
- (1972) 'Some Munda etymological notes on names in the Ramayana', *Journal of the Ganganatha Jha Kendriya Sanskrit Vidyapeeth XXVIII*, 747–763.
- (1978) *Studies in the Munda Numerals*, Mysore: Central Institute of Indian Languages.

INDEX

Note: References such as “178–179” indicate (not necessarily continuous) discussion of a topic across a range of pages. Wherever possible in the case of topics with many references, these have been divided into sub-topics. Because the entire volume is about Munda languages, the term “Munda” has not been used as an entry point.

- ablative: Gta? 692; Ho 216, 234; Juang 515, 518–519, 526; Kharia 436, 454; Kherwarian 216, 234; Korku 263, 285, 287; Santali 11, 34–35, 38, 81; Sora 325
- accusative: Gorum 397; Gta? 695, 721; Gutob 665; Ho 205; Juang 515–516, 529, 543; Kera? Mundari 171, 176–177, 185; Kharia 459; Kherwarian 205; Korku 263, 275, 284–285, 287, 289; Santali 36, 54, 68
- adjectives: Gorum 396–397; Gta? 688, 718, 745; Gutob 644, 652, 665, 671; Ho 204, 209, 214–215, 230, 234; Juang 514, 526, 541–543, 549; Kera? Mundari 172–175; Kharia 453–455, 457; Kherwarian 204, 209, 214–215, 230, 234; Korku 262, 264–266, 269–270, 275, 278, 283, 288, 290; Mundari 106–107, 111, 117–118, 147; Remo 577, 580–581, 586–587, 608, 610–611; Santali 11–12, 48, 50–52, 76; Sora 308, 316, 324–325, 362
- adpositions *see also* postpositions:
Gorum 393–394; Gta? 688, 714;
Gutob 648–650; Ho 212–213; Kera?
Mundari 173–174; Kharia 442–443,
451–452; Kherwarian 212–213;
Korku 268; Remo 573, 575, 582–583, 615;
Santali 48–50; Sora 306, 320–321
- adverbials: Gorum 397, 413–414; Gta? 697,
702, 708, 711, 718–719; Gutob 647,
652, 665; Ho 209–210, 212, 215–216,
234; Juang 526–527, 541–543; Kera?
Mundari 172, 175, 183, 187; Kharia 443,
445, 448, 451, 455, 461, 471, 479, 489–491;
Kherwarian 209–210, 212, 215–216, 234;
Korku 265, 267, 269–270, 275, 282–283, 290;
Mundari 106–107, 109, 111–113, 118–119,
137, 143–144, 146–147; Remo 583, 587–588,
611–612; Santali 12, 47, 49–50, 52–53, 72,
81–82; Sora 306, 310, 325–327, 363–364, 366
- adversative 76, 79–80, 292, 670
- Agariya 2, 195 *see also* Kherwarian
- agreement: Gorum 387–388; Gta? 704–705,
719, 721–723; Ho 205, 207–208, 219,
222–223; Juang 514, 516, 519–520, 527–531;
Kharia 459–461, 492; Kherwarian 205,
207–208, 219, 222–223; Korku 278;
Mundari 120–121, 131–132, 145–146;
Remo 570–571; Santali 11, 32, 54–55,
77, 85; Sora 309, 319, 327–329, 331,
338–339, 355
- allative 11, 18, 34–36, 38, 312, 347, 389
- Allative Korku 263
- allophony: Gorum 385; Gta? 686; Gutob
637–639; Ho 200–202; Kharia 435–436,
438, 500; Kherwarian 200–202; Korku 259;
Mundari 102; Santali 20–21, 26
- analyzed texts: Gorum 423–429; Gta?
756–761; Gutob 671–675; Ho 240–250;
Juang 549–554; Kera? Mundari 188–192;
Kharia 495–498; Kherwarian 240–250;
Korku 293–297; Mundari 156–161;
Remo 623–630; Santali 88–93; Sora 371–378
- Aryanization *see* Indo-Aryan influences
- aspect 7; Gorum 400; Gta? 726–727;
Gutob 653–656; Ho 221–225; Juang
532–534; Kera? Mundari 178–181;
Kharia 462–465; Kherwarian 221–225;
Korku 273–279; Mundari 125–127;
Remo 594–596; Santali 57–63; Sora 337
- Asuri 1–2, 195 *see also* Kherwarian;
coordination 236; distribution 2, 197;
lexicon 238–239; subject 219; tense
222–224; voice 226
- Austroasiatic connection 1, 5, 7 *see also*
Austroasiatic lexical elements
- Austroasiatic lexical elements: Gorum 422;
Gta? 756; Ho 237–240; Juang 548–549;
Kharia 493; Kherwarian 237–240;
Mundari 156; Remo 619; Santali 88;
Sora 371
- auxiliary verb constructions *see also*
compound verb constructions: Asuri 229;
Gorum 410–413; Gta? 726–727, 737–741;
Gutob 654, 662–664; Ho 228–229;
Juang 539; Juray 358; Kharia 481–482;
Kherwarian 228–229; Korku 283–288;
Mundari 134–138; Remo 603–607;
Sora 357–360; Turi 229

- Bhumij *see also* Kherwarian: distribution 2–3, 196–197; Indo-Aryan influences 197, 205, 240; lexicon 238–239; mood 225; numerals 210–211; object types 220; simple sentences 232; subject 218; tense 221, 224–225; voice 226
- bilingualism: Gorum 382, 421; Gutob 634, 670; Juang 508, 544; Kherwarian 250; Korku 256, 269; Nihali 764, 773; Santali 15–17
- Birhor *see also* Kherwarian: distribution 2, 197–198; lexicon 238–239; object types 220; tense 222
- Bonda *see* Remo
- cardinal numbers *see also* numerals:
Gorum 393; Gta? 713–714; Gutob 646–647; Juang 524; Kera? Mundari 173; Kharia 448–450; Mundari 114; Remo 580–582; Santali 45–46
- case: Bhumij 206; Gorum 388–389; Gta? 691–700; Gutob 641; Ho 205–207; Juang 515–519; Kera? Mundari 170–171; Kharia 442–443; Kherwarian 205–207; Korku 263–264; Mundari 108–109; Remo 570–576; Santali 34–38; Sora 309–312
- class (nouns): Gorum 391; Gta? 704; Gutob 642–643; Ho 207–208; Kera? Mundari 171; Kharia 445; Kherwarian 207–208; Korku 264; Mundari 107–108; Remo 577; Santali 39–41; Sora 315
- classification of languages 2–4; Gorum 381; Gta? 682; Gutob 635; Juang 508; Kera? Mundari 167–168; Santali 13; Sora 299
- classifiers: Gutob 647; Juang 515, 524; Kera? Mundari 173; Kharia 450–451; Santali 46–47
- combining forms: Gorum 409–410; Gta? 736–737; Gutob 651–652, 662; Korku 283; Remo 602–603; Santali 72; Sora 351–357
- complement clauses: Gutob 666–667; Juang 547; Kera? Mundari 186; Kharia 486; Remo 613–614; Santali 80–81; Sora 364
- complex predicate types *see also* auxiliary verb constructions: Gorum 410–413; Gta? 737–741; Remo 603–607
- complex sentence structure: Gorum 415–420; Gta? 746–750; Gutob 666–669; Ho 234–236; Juang 544–548; Kera? Mundari 185–187; Kherwarian 234–236; Korku 290–293; Mundari 148–152; Remo 613–617; Santali 78–84; Sora 364–370
- compound verb constructions *see also* auxiliary verb constructions: Gutob 657; Kera? Mundari 183–184; Mundari 134–135; Santali 65
- conditional clauses: Gta? 748; Gutob 668–669; Juang 544, 546, 548; Kera? Mundari 181, 187; Santali 82–83
- conjunctions: Gorum 412, 418; Gta? 695; Gutob 651, 666, 668; Ho 235; Juang 545; Kera? Mundari 186; Kharia 465, 468, 485, 489–490; Mundari 106, 148–149; Remo 616; Santali 12, 78; Sora 369
- consonants 6; Gorum 384–385; Gta? 683–684; Gutob 638–639; Ho 201–203; Juang 510–511; Kera? Mundari 168–169; Kharia 437–439; Kherwarian 201–203; Korku 258–259; Mundari 100; Remo 559–560; Santali 26–29, 31–32; Sora 301–302
- coordination: Asuri 236; Gorum 418–420; Gta? 750–754; Gutob 666; Ho 235–236; Juang 545; Kera? Mundari 186; Kherwarian 235–236; Mundari 148–149; Remo 616–617; Santali 78–80; Sora 368–370
- dative: Gorum 388; Gta? 695, 699, 704, 735; Juang 515, 519, 529, 539; Kera? Mundari 171, 173, 177, 185; Kharia 491; Korku 263, 289; Mundari 122, 156; Santali 34–36, 68, 75; Sora 309, 315, 354
- definiteness: Gorum 390–391; Gta? 702–704; Gutob 642; Ho 207; Juang 520; Kera? Mundari 171; Kharia 445; Kherwarian 207; Korku 264; Remo 576–577; Sora 315
- demonstratives 6; Gorum 392–393; Gta? 710–712; Gutob 645–646; Ho 209–210; Juang 523; Kera? Mundari 172–173; Kharia 447–448; Kherwarian 209–210; Korku 266–267; Mundari 110–114; Remo 578–580; Santali 44–45; Sora 317–318
- derivation: Gorum 394–396, 408–409; Gta? 714–717, 736; Gutob 650–652, 661–662; Ho 213–215, 228; Juang 525, 538–539; Kera? Mundari 174, 182–183; Kharia 452–453, 473–480; Kherwarian 213–215, 228; Korku 268–269, 282–283; Mundari 116–117, 132–133; Remo 583–585, 601–602; Santali 50–52, 68–72; Sora 321–325, 351
- Desia: and Gorum 382–383, 385, 387, 391, 393, 397, 400, 413–414, 420–422; and Gta? 695, 756; and Gutob 634–636, 638–639, 643–647, 664, 669–671; and Remo 557, 580, 596
- dialectal differences: Gorum 382; Gutob 635–636; Kharia 434; Santali 17–19; Sora 299
- diphthongs: Gta? 683; Gutob 638; Ho 199, 244; Juang 509; Kera? Mundari 168; Kharia 436, 440; Remo 558–559; Santali 11, 23–25, 30, 71

- directionality: Gorum 402–403; Gta? 729–730; Gutob 657; Ho 226; Juang 536; Kharia 465; Kherwarian 226; Korku 273–279; Mundari 129; Remo 597–598; Santali 65; Sora 341–342
- discourse: Gorum 420–421; Gta? 755; Gutob 669–670; Ho 236–237; Kera? Mundari 188; Kherwarian 236–237; Korku 293; Mundari 153–155; Remo 617–619; Santali 86–87; Sora 370
- distribution of languages: Asuri 197; Bhumij 196–197; Birhor 197–198; Gorum 381–382; Gta? 682; Gutob 633–635; Ho 196; Juang 508; Kera? Mundari 165–166; Kharia 434–435; Kherwarian 195–198; Korku 256–257; Korwa 197; map 2; Mundari 99; Remo 557–558; Santali 13–15; Sora 299–300; Turi 198
- Dravidian connections 1; Gorum 382, 385, 422; Gta? 682; Gutob 635–636; Juang 549; Kera? Mundari 165, 184; Kharia 434; Korku 256, 291; Nihali 765–767, 772–773; Sora 371
- dual: Gta? 689; Ho 208–209; Juang 514–515; Kera? Mundari 178; Kharia 460; Korku 262, 264–265; Mundari 109; Santali 32–34, 42–43
- echo forms *see* expressives
- expressives 7; Gorum 413; Gta? 741–743; Gutob 665; Ho 230–231; Juang 539; Kera? Mundari 184; Kharia 482–483; Kherwarian 230–231; Korku 288; Mundari 139–145; Remo 607–608; Santali 73–74; Sora 360–362
- finiteness: Gorum 407; Gta? 731–733; Gutob 658–659; Ho 227; Juang 537; Kera? Mundari 181; Kharia 468–471; Kherwarian 227; Korku 273–279; Mundari 131–132; Remo 598–600; Santali 66; Sora 345–346
- gender: Gorum 391; Gta? 704; Gutob 642–643; Ho 207–208; Juang 520–521; Kera? Mundari 171; Kharia 445; Kherwarian 207–208; Korku 264; Remo 577; Santali 39–41; Sora 315
- genitive: Gorum 388–389, 397, 416; Gta? 694–695, 699–703, 705, 718, 732, 745, 750; Gutob 641, 648; Ho 206–207, 234–235; Juang 515–520, 523–524, 539, 549; Kera? Mundari 167, 170, 175, 185; Kharia 440–441, 443, 449, 451–452, 456–457, 469, 488–489; Korku 258, 261–264, 268, 284; Mundari 101, 109–110, 113, 147; Remo 574–577, 585, 611, 615; Santali 11, 13, 18, 34–35, 38, 40, 48, 53, 88; Sora 310, 312, 320–321
- Gorum 381–430; analyzed text 423–429; classification 3–4, 381; and Desia 382–383, 385, 387, 391, 393, 397, 400, 413–414, 420–422; dialects 382; discourse 420–421; distribution 2–3, 381–382; lexicon 422; literary status 382; morphology 387–413; morphophonology 386–387; nominal morphology 387–397; phonology 382–387; semantics 420; and Sora 299; syntax 414–420; verbal morphology 397–413
- Gta? 682–761; analyzed text 756–761; classification 3–4, 682; and Desia 695, 756; dialects 682; discourse 755; distribution 2, 682; expressives 741–743; and Gutob 719; and Kharia 719; lexicon 756; literary status 682; loans 756; morphology 688–743; morphophonology 686–687; nominal morphology 688–719; phonology 682–687; and Remo 557, 682, 719, 731; semantics 754; syntax 743–754; verbal morphology 719–741
- Gutob 633–678; analyzed text 671–675; classification 3–4, 635; and Desia 634–636, 638–639, 643–647, 664, 669–671; dialects 635–636; discourse 669–670; distribution 2–3, 633–635; expressives 665; and Gta? 719; language name derivation 633; lexicon 670–671; literary status 636; loans 670–671; morphology 640–665; morphophonology 640; nominal morphology 640–651; phonology 637–640; and Remo 557, 590; semantics 669; studies 636–637; syntax 665–669; verbal morphology 652–664
- Hasada dialect: Mundari 99
- Ho 1, 195–253; analyzed text 240–250; discourse 236–237; distribution 2–3, 196; lexicon 237–240; loans 240; morphology 204–231; nominal morphology 205–216; orthography 250–252; phonology 198–204; semantics 236; syntax 231–236; verbal morphology 216–229
- honorific pronouns: Santali 42
- hortative: Gta? 727, 729; Gutob 643–644; Juang 535; Sora 338–339
- imperative: Gorum 401–402, 404; Gta? 719, 723, 727–729, 733; Gutob 643, 657, 661; Ho 217, 225, 227, 231; Juang 509, 533–536, 543; Kera? Mundari 181; Kharia 464–465, 472; Korku 273, 277, 285; Mundari 119, 121, 127–128, 132, 151; Remo 563, 596–597, 601, 607, 609–610, 617; Santali 31–32, 39, 63–67, 70; Sora 329, 337–340, 349, 357
- indefinite pronouns: Gta? 708–709; Gutob 645; Kera? Mundari 172; Kharia 447; Mundari 112; Santali 43–44

- Indo-Aryan influences 1; Bhumij 197, 205, 240; Gorum 381–382, 384, 396, 422; Gta? 682, 704, 753; Gutob 635–636, 667, 670–671; Ho 210, 237; Juang 508, 524, 527, 549–550, 553; Kera? Mundari 168, 171–177, 181, 183–184, 186, 188; Kharia 434–436, 438–439, 445, 448, 450, 454, 471, 485, 487, 490, 493–494; Kherwarian 197, 205, 240; Korku 256, 258, 262, 267, 282–283, 290–291; Mundari 101, 108, 114–115, 122, 156; Nihali 765–766, 773; Remo 575, 577, 580–581, 586, 612–613, 619, 623; Santali 15–17, 21, 26–27, 30, 39–40, 45–8, 88; Sora 315, 371
- infixation: Gutob 661–662; Ho 213; Kherwarian 213; Korku 262; Santali 51, 56, 68–69, 71
- interrogative pronouns: Gorum 392; Gta? 707–709; Gutob 644–645; Santali 42–44; Sora 317
- interrogatives: Kera? Mundari 172; Kharia 447; Korku 265–266; Mundari 112
- intonation: Gorum 386; Gta? 686; Gutob 639–640; Ho 204; Juang 512–513; Kera? Mundari 169; Kharia 439; Kherwarian 204; Korku 260; Mundari 104–105; Remo 565; Sora 306
- Juang 508–555; analyzed text 549–554; classification 3–4, 508; distribution 2–3, 508; expressives 539; lexicon 548–549; loans 549; morphology 513–539; morphophonology 513; orthography 508; phonology 509–513; semantics 548; studies 508; syntax 539–548; verbal morphology 527–539
- Juray 299; auxiliary verb constructions 358; mood 340–341; voice 344
- Kera dialect: Mundari 99
- Kera? Mundari 1–2, 165–193; analyzed text 188–192; classification 167–168; discourse 188; lexicon 188; morphology 170–184; morphophonology 169–170; nominal morphology 170–175; phonology 168–170; semantics 188; studies 166–167; syntax 184–187; verbal morphology 175–184
- Kharia 434–507; analyzed text 495–498; classification 3–4; dialects 434; distribution 2–3, 434–435; expressives 482–483; and Gta? 719; lexicon 493–495; loans 494–495; morphology 441–483; morphophonology 439–440; nominal morphology 441–455; phonology 435–440; and Remo 557; semantics 492–493; studies 506–507; syntax 483–492; verbal morphology 455–482
- Khasi 1, 737, 768
- Kherwarian 1, 195–253; analyzed text 240–250; classification 3–4; discourse 236–237; distribution 195–198; lexicon 237–240; loans 240; morphology 204–231; nominal morphology 205–216; phonology 198–204; semantics 236; syntax 231–236; verbal morphology 216–229
- Koda: distribution 2; tense 222–224
- Koraku: distribution 2
- Koraput Munda: classification 3
- Korku 1, 256–297; analyzed text 293–297; classification 3–4; distribution 2, 256–257; lexicon 293; loans 293; morphology 262–288; morphophonology 260–262; and Nihali 764–766; nominal morphology 262–269; phonology 257–262; syntax 288–293; verbal morphology 270–288
- Korwa *see also* Kherwarian: distribution 2, 197; lexicon 238–239; negation 228; object types 220; tense 222–223
- Kurux 165, 171, 188
- language name derivations: Gorum 381; Gutob 633; Kera? Mundari 166; Santali 12–13
- lexicon: Asuri 238–239; Austroasiatic components *see* Austroasiatic lexical elements; Bhumij 238–239; Birhor 238–239; Gorum 422; Gta? 756; Gutob 619–621; Ho 237–240; Juang 548–549; Kera? Mundari 188; Kharia 493–495; Kherwarian 237–240; Korku 293; Korwa 238–239; Mundari 156, 238–239; Remo 619–623; Santali 87–88; Sora 371; Turi 238–239
- loans: Gta? 756; Gutob 670–671; Ho 240; Juang 549; Kherwarian 240; Korku 293; Mundari 156; Remo 619–623; Santali 88; Sora 371
- locative: Juang 517–518; Korku 264; Remo 573; Santali 35, 38; Sora 311
- Mahali: distribution 2–3
- mood: Bhumij 225; Gorum 400–402; Gta? 727–729; Gutob 653–655, 657; Ho 225–226; Juang 534–536; Juray 340–341; Kera? Mundari 178–181; Kharia 462–465; Kherwarian 225–226; Korku 273–279; Mundari 127–129; Remo 596–597; Santali 63–64; Sora 337–341
- morphology: expressives *see* expressives; Gorum 387–413; Gta? 688–743; Gutob 640–665; Ho 204–231; Juang 513–539; Kharia 441–483; Kherwarian 204–231; Korku 262–288; Mundari 106–145; nominal *see* nominal morphology; Remo 568–608; Santali 18, 32–74; Sora 306–362; verbal *see* verbal morphology

- morphophonology: Gorum 386–387;
 Gta? 686–687; Gutob 640; Ho 204;
 Juang 513; Kera? Mundari 169–170;
 Kharia 439–440; Kherwarian 204;
 Korku 260–262; Mundari 105–106;
 Remo 565–568; Santali 31–32;
 Sora 306
- Munda *see Note and detailed topics*
- Mundari 1, 99–162; analyzed text 156–161;
 classification 3–4; discourse 153–155;
 distribution 2–3, 99; Hasada dialect 99;
 Kera dialect 99; lexicon 156; morphology
 106–145; morphophonology 105–106;
 Naguri dialect 99; nominal morphology
 107–119; phonology 100–106; semantics
 153; syntax 145–152; Tamararia dialect 99;
 verbal morphology 119–138; word
 classes 106–107
- Muwasi *see* Korku
- Naguri dialect: Mundari 99
- negation 7; Gorum 407–408; Gta? 708,
 733–736; Gutob 654–655, 659–661;
 Ho 227–228; Juang 537–538; Kera?
 Mundari 181–182; Kharia 471–473;
 Kherwarian 227–228; Korku 279–281;
 Korwa 228; Mundari 132; Remo 600–601;
 Santali 66–68; Sora 346–350; Turi 228
- Nicobarese 1, 410, 737
- Nihali 764–774; as argot 769–772;
 Austroasiatic (non-Munda) connections
 767–768; Dravidian connections 766–767;
 and Korku 764–766; linguistic layers
 766–772; Munda connections 768–769;
 proposed history 772–774; Tibeto-Burman
 connections 766–767
- nominal attributes *see* adjectives
- nominal morphology: adjectives *see* adjectives;
 adpositions *see* adpositions; adverbials
see adverbials; case *see* case; class
 (nouns) *see* class (nouns); definiteness
see definiteness; demonstratives *see*
 demonstratives; derivation *see* derivation;
 gender *see* gender; Gorum 387–397;
 Gta? 688–719; Gutob 640–651; Ho
 205–216; Juang 514–527; Kera? Mundari
 170–175; Kharia 441–455; Kherwarian
 205–216; Korku 262–269; number *see*
 number; numerals *see* numerals; person
see person; postpositions *see* postpositions;
 pronouns *see* pronouns; Remo 568–588;
 Santali 32–53; Sora 307–327
- nominalizers *see* derivation
- non-finiteness *see* finiteness
- North Munda group 2 *see also under*
individual languages; classification 3–4; and
 Santali 13
- noun class *see* class (nouns)
- noun incorporation: Gorum 409–410;
 Gta? 736–737; Gutob 662; Ho 228;
 Juang 539; Kera? Mundari 183;
 Kharia 480–481; Kherwarian 228;
 Korku 283; Remo 602–603; Santali 72;
 Sora 351–357
- nouns *see* nominal morphology
- number: Bhumij 205; Gorum 387–388;
 Gta? 689–691; Gutob 640–641; Ho 205;
 Juang 514–515; Kera? Mundari 170;
 Kharia 442; Kherwarian 205; Korku 262;
 Mundari 107–108; Remo 568–570;
 Santali 32–34; Sora 307–309
- numerals: Bhumij 210–211; Gorum 393;
 Gta? 713–714; Gutob 646–647; Ho
 210–212; Juang 524; Kera? Mundari 173;
 Kharia 448–451; Kherwarian 210–212;
 Korku 267–268; Mundari 114–115;
 Remo 580–582; Santali 45–48; Sora 318–320
- object types: Bhumij 220; Birhor 220;
 Gorum 398–399; Gta? 723; Gutob 653;
 Ho 219–221; Juang 529–531; Kera?
 Mundari 177–178; Kharia 461–462;
 Kherwarian 219–221; Koda 221; Korku
 271–273; Korwa 220; Mundari 121–125;
 Remo 590; Santali 55–56; Sora 329–330
- Oraons 99, 165–167, 181, 188, 767
- ordinal numbers *see* numerals
- orientation: Gorum 402–403; Gta? 729–730;
 Gutob 657; Ho 226; Juang 536; Kharia
 465; Kherwarian 226; Korku 273–279;
 Mundari 129; Remo 597–598; Santali 65;
 Sora 341–342
- orthography: Ho 250–252; Juang 508;
 Santali 19–20; Sora 299–301
- person 7; Gorum 389–390; Gta? 700–702;
 Gutob 642; Ho 207; Juang 519–520;
 Kera? Mundari 171; Kharia 443–444;
 Kherwarian 207; Korku 264; Remo 576;
 Santali 38–39; Sora 312–315
- personal pronouns: Gorum 391–392
see also pronouns; Gta? 704–707; Gutob
 643–644; Ho 208–209; Juang 521–523;
 Kera? Mundari 171–172; Kharia 445–447;
 Kherwarian 208–209; Korku 264–
 265; Mundari 109; Remo 577–578;
 Santali 41–42; Sora 316
- phonemic inventory *see* phonology
- phonology: Gorum 382–387; Gta? 682–687;
 Gutob 637–640; Ho 198–204; Juang
 509–513; Kera? Mundari 168–170;
 Kharia 435–440; Kherwarian 198–204;
 Korku 257–262; Mundari 100–106;
 Remo 558–568; Santali 17–18, 20–32;
 Sora 301–306
- phonotactics: Gta? 684–686; Gutob 639;
 Ho 203–204; Kera? Mundari 169;
 Kherwarian 203–204; Korku 259;
 Mundari 102–103; Remo 560–565;
 Sora 302–305

- plural: Gorum 387–388; Gta? 689–691; Gutob 640–641; Ho 208–209; Juang 514–515; Kera? Mundari 178; Kharia 460; Korku 262, 264–265; Mundari 109; Remo 568–570; Santali 32–34; Sora 307–309, 327–328, 330
- possessive: Gorum 389–390; Gta? 698–699, 701; Juang 519–520; Kharia 443–444; Mundari 109–110, 113; Remo 574–576; Sora 312–315
- postpositions *see also* adpositions: Gorum 393–394; Gta? 714; Gutob 648–650; Ho 212–213; Juang 524–525; Kera? Mundari 170–171, 173–174; Kharia 451–452; Kherwarian 212–213; Korku 263–264, 268; Mundari 109, 115–116, 150–151; Remo 582–583; Santali 48–50; Sora 320–321
- predicative morphology *see* verbal morphology
- prefixation: Gutob 662; Santali 51–52, 68
- prepositions *see also* adpositions: Kera? Mundari 173; Kharia 451
- pronouns: Gorum 391–392; Gta? 704–710; Gutob 643–645; Ho 208–209; Juang 521–523; Kera? Mundari 171–172; Kharia 445–447; Kherwarian 208–209; Korku 264–266; Mundari 109–110; Remo 577–578; Santali 41–44; Sora 316–317
- Proto-Munda 3–4
- reduplication 6–7; Gorum 394–395, 413; Gta? 701, 715, 718, 726, 733, 738, 747, 755; Gutob 640, 651, 661, 670; Ho 214–215, 230–231; Juang 525, 527–528, 532, 534, 537–538, 544; Kera? Mundari 181; Kharia 442, 453–454, 469, 483; Kherwarian 214–215; Korku 258, 261, 270, 274, 279–280, 282; Mundari 115, 119, 132–133, 139–142; Remo 567, 583, 585, 587, 596, 598–599, 602, 604, 607–608, 627; Santali 47, 68, 71, 73; Sora 307, 321–322, 346, 360–361
- register *see* suprasegmental phenomena
- relative clauses: Gorum 416; Gta? 747–748; Gutob 667–668; Ho 234; Juang 545; Kera? Mundari 186–187; Kharia 486–489; Kherwarian 234; Korku 290–293; Mundari 149; Remo 614; Santali 83–84; Sora 365
- relative pronouns *see also* pronouns: Gutob 644; Kera? Mundari 172
- Remo 557–631; analyzed text 623–630; classification 3–4; and Desia 557, 580, 596; discourse 617–619; distribution 2–3, 557–558; expressives 607–608; and Gta? 557, 682, 719, 731; and Gutob 557, 590; and Kharia 557; lexicon 619–623; loans 619–623; morphology 568–608; morphophonology 565–568; nominal morphology 568–588; phonology 558–568; semantics 617; syntax 608–617; verbal morphology 588–608
- Sadani 165–166, 171; and Kharia 494
- Santali 1, 11–94, 188; adpositions 48–50; classification 3; demonstratives 44–45; derivation of name 12–13; dialectal differences 17–19; discourse 86–87; distribution 2–3, 13–15; expressives 73–74; language name 12–13; lexicon 18–19, 87–88; loans 88; morphology 18, 32–74; morphophonology 31–32; nominal morphology 32–34; orthography 19–20; phonology 17–18, 20–32; semantics 85–86; studies 97–98; syntax 74–84
- scripts 6 *see also* orthography
- semantics: Gorum 420; Gta? 754; Gutob 669; Ho 236; Juang 548; Kera? Mundari 188; Kharia 492–493; Kherwarian 236; Korku 293; Mundari 144–145, 153; Remo 617; Santali 85–86; Sora 370
- sentence structure *see* complex sentence structure; simple sentences
- serial verb constructions *see* auxiliary verb constructions
- simple sentences: Bhumij 232; Gorum 414–415; Gta? 743–746; Gutob 665; Ho 231–234; Juang 539–544; Kera? Mundari 184–185; Kharia 483–485; Kherwarian 231–234; Korku 289–290; Mundari 145–146; Remo 608–613; Santali 74–78; Sora 362–364
- singular: Gorum 387–388; Gta? 689; Gutob 640–641; Ho 208–209; Juang 514–515; Kera? Mundari 178; Kharia 460; Korku 264–265; Mundari 109; Santali 32–34; Sora 328, 330
- Sora 299–379; analyzed text 371–378; Austroasiatic components 371; classification 3–4, 299; discourse 370; distribution 2–3, 299–300; expressives 360–362; and Gorum 299; lexicon 371; loans 371; morphology 306–362; morphophonology 306; nominal morphology 307–327; orthography 299–301; phonology 301–306; semantics 370; syntax 362–370; verbal morphology 327–360
- South Munda group 2 *see also under individual languages*; classification 3–4; and Santali 13
- speakers *see* distribution of languages
- statistics *see* distribution of languages
- stress: Gorum 386; Gta? 686; Gutob 639–640; Ho 204; Juang 512–513; Kera? Mundari 169; Kherwarian 204; Korku 260; Mundari 104–105; Remo 565; Sora 306
- studies 5–6; Gorum 382; Gutob 636–637; Juang 508; Mundari 99–100; Santali 97–98
- subgroups 1–2
- subject: Asuri 219; Bhumij 218; Gorum 397–398; Gta? 719–723; Gutob 653; Ho 216–219; Juang 528–529; Kera?

- Mundari 176–177; Kharia 459–461; Korku 271–273; Mundari 120–121; Remo 589; Santali 54–55; Sora 327–329
- subordinate clauses: Gorum 416–418; Gta? 748–750; Gutob 668–669; Ho 234–235; Kera? Mundari 187; Kharia 491–492; Kherwarian 234–235; Mundari 150–152; Remo 614–616; Santali 81–83; Sora 366–368
- suprasegmental phenomena: Bhumij 200–201; Gorum 383–384; Gta? 683; Gutob 638; Ho 200–201; Juang 509–510; Kera? Mundari 168; Kharia 436–437; Kherwarian 200–201; Korku 257–258; Remo 558–559; Santali 29–30; Sora 301
- switch reference: Gorum 418–420; Gta? 750–754; Ho 235–236; Kherwarian 235–236; Mundari 152; Remo 616–617
- syllable structure: Gorum 385–386; Gta? 684–686; Gutob 639; Ho 203–204; Juang 511–512; Kera? Mundari 169; Kharia 439; Kherwarian 203–204; Korku 259; Mundari 102–103; Remo 560–565; Santali 30–31; Sora 302–305
- syntax: Gorum 414–420; Gutob 665–669; Ho 231–236; Juang 539–548; Kera? Mundari 184–187; Kharia 483–492; Kherwarian 231–236; Korku 288–293; Mundari 145–152; Remo 608–617; Santali 74–84; Sora 362–370
- tag forms *see* expressives
- TAM (tense, aspect, mood) *see* aspect; mood; tense
- Tamaria dialect: Mundari 99
- taxonomy *see* classification of languages
- tense 7; Asuri 222–224; Bhumij 221, 224–225; Birhor 222; Gorum 399–400; Gta? 724–726; Gutob 653–655; Ho 221–225; Juang 531–532; Kera? Mundari 178–181; Kharia 462–465; Kherwarian 221–225; Koda 222–224; Korku 273–279; Korwa 222–223; Mundari 125–127; Remo 590–594; Santali 56–57; Sora 330–337; Turi 222, 224
- tone *see* suprasegmental phenomena
- Turanian *see* Santali
- Turi *see also* Kherwarian: auxiliary verb constructions 229; distribution 2–3, 198; lexicon 238–239; negation 228; tense 222, 224
- typological features: Gta? 745–746; Gutob 665; Juang 542–544; Remo 610–613
- valence 7 *see also* voice; Gutob 657–658; Kharia 465–468; Mundari 129–131; Sora 343–344
- verbal morphology: aspect *see* aspect; auxiliary verb constructions *see* auxiliary verb constructions; combining forms *see* combining forms; derivation *see* derivation; directionality *see* directionality; finiteness *see* finiteness; Gta? 719–741; Gutob 652–664; Ho 216–229; Juang 527–539; Kera? Mundari 175–184; Kharia 455–482; Kherwarian 216–229; Korku 270–288; mood *see* mood; Mundari 119–138; negation *see* negation; noun incorporation *see* noun incorporation; object types *see* object types; orientation *see* orientation; Remo 588–608; Santali 53–73; Sora 327–360; subject *see* subject; tense *see* tense; valence *see* valence; version *see* version; voice *see* voice
- version *see also* voice: Gorum 403–407; Kharia 465–468; Sora 342–344
- vocabulary *see* lexicon
- voice 7; Asuri 226; Bhumij 226, 404; Gorum 403–407; Gta? 730–731; Gutob 657–658; Ho 226–227; Juang 536–537; Juray 344; Kera? Mundari 181; Kharia 465–468; Kherwarian 226–227; Korku 273–279; Mundari 129–131, 404; Remo 598; Santali 65–66; Sora 342–344
- vowels 6; Gorum 382–383; Gta? 683; Gutob 637–638; Ho 198–200; Juang 509; Kera? Mundari 168; Kharia 435–436; Kherwarian 198–200; Korku 257; Korwa 199; Mundari 100; Remo 558; Santali 20–25; Sora 301

This page intentionally left blank

eBooks

eBooks – at www.eBookstore.tandf.co.uk

A library at your fingertips!

eBooks are electronic versions of printed books. You can store them on your PC/laptop or browse them online.

They have advantages for anyone needing rapid access to a wide variety of published, copyright information.

eBooks can help your research by enabling you to bookmark chapters, annotate text and use instant searches to find specific words or phrases. Several eBook files would fit on even a small laptop or PDA.

NEW: Save money by eSubscribing: cheap, online access to any eBook for as long as you need it.

Annual subscription packages

We now offer special low-cost bulk subscriptions to packages of eBooks in certain subject areas. These are available to libraries or to individuals.

For more information please contact
webmaster.ebooks@tandf.co.uk

We're continually developing the eBook concept, so keep up to date by visiting the website.

www.eBookstore.tandf.co.uk